RESEARCH CENTER
Bordeaux - Sud-Ouest

FIELD

Activity Report 2019

Section New Results

Edition: 2020-03-21

LFANT Project-Team

6. New Results

6.1. Cryptographic Protocols

Participants: Guilhem Castagnos, Ida Tucker.

In [20], G. Castagnos, D. Catalano, F. Laguillaumie, F. Savasta and I. Tucker propose a new cryptographic protocol to compute ECDSA signatures with two parties.

ECDSA (Elliptic Curves Digital Signature Algorithm) is a widely adopted standard for electronic signatures. For instance, it is used in the TLS (Transport Layer Security) protocol and in many cryptocurrencies such as Bitcoin. For cryptocurrencies, ECDSA is used in order to sign the transactions: if Alice wants to give n bitcoins to Bob, she uses her secret key to sign with ECDSA a bit string encoding this information.

As a result, if the secret key of Alice is stolen, for example if her computer is compromised, an attacker can stole all her bitcoins. A common solution to this problem is to share the key on multiple devices, for example a laptop and a mobile phone. Both devices must collaborate in order to issue a signature, and if only one device is compromised, no information on the key is leaked. This setting belongs to the area of secure multiparty computation.

There have been recent proposals to construct 2 party variants of ECDSA signatures but constructing efficient protocols proved to be much harder than for other signature schemes. The main reason comes from the fact that the ECDSA signing protocol involves a complex equation compared to other signatures schemes. Lindell recently managed to get an efficient solution using the linearly homomorphic cryptosystem of Paillier. However his solution has some drawbacks, for example the security proof resorts to a non-standard interactive assumption.

By using another approach based on hash proofs systems we obtain a proof that relies on standard assumptions. Moving to concrete constructions, we show how to instantiate our framework using class groups of imaginary quadratic fields. Our implementations show that the practical impact of dropping such interactive assumptions is minimal. Indeed, while for 128-bit security our scheme is marginally slower than Lindell's, for 256-bit security it turns out to be better both in key generation and signing time. Moreover, in terms of communication cost, our implementation significantly reduces both the number of rounds and the transmitted bits without exception.

This paper was presented at the CRYPTO Conference 2019, and is part of the ALAMBIC project.

6.2. Coding Theory

Participants: Xavier Caruso, Aurel Page.

In [29], Xavier Caruso developed a theory of residues for skew rational functions (which are, by definition, the quotients of two skew polynomials), proving in particular a skew analogue of the residue formula and a skew analogue of the classical formula of change of variables for residues. He then used his theory to define and study a linearized version of Goppa codes. He showed that these codes meet the Singleton bound (for the sum-rank metric) and are the duals of the linearized Reed–Solomon codes defined recently by Martínez-Peñas. Efficient encoding and decoding algorithms are also designed.

C. Maire and A. Page updated the preprint *Error-correcting codes based on non-commutative algebras* [33] according to the comments of referees.

6.3. Number fields

Participants: Razvan Barbulescu, Jean-Marc Couveignes, Jean-Paul Cerri, Pierre Lezowski.

In [30], Jean-Marc Couveignes constructs small models of number fields and deduces a better bound for the number of number fields of given degree n and discriminant bounded by H. This work improves on previous results by Schmidt and Ellenberg-Venkatesh. Schmidt obtains a bound $H^{\frac{n+2}{4}}$ times a function of n. Ellenberg and Venkatesh obtain a bound $H^{\exp(O(\sqrt{\log n}))}$ times a function of n. The new idea is to combine geometry of numbers and interpolation theory to produces small projective models and lower the exponent of H down to $O(\log^3 n)$. A key point is to look for local equations rather than a full set of generators of the ideal of these models.

In [12], Razvan Barbulescu in a joint work with Jishnu Ray (University of British Columbia, Vancouver) brings elements to support Greenberg's p-rationality conjecture. On the theoretical side, they propose a new family proven to be p-rational. On the algorithmic side, the compare the tools to enumerate number fields of given abelian Galois group and of computing class numbers, and extend the experiments on the Cohen-Lenstra-Martinet conjectures.

In collaboration with Pierre Lezowski, Jean-Paul Cerri has studied in [15] norm-Euclidean properties of totally definite quaternion fields over number fields. Building on their previous work about number fields, they have proved that the Euclidean minimum and the inhomogeneous minimum of orders in such quaternion fields are always equal. Additionally, they are rational under the hypothesis that the base number field is not quadratic. This single remaning open case corresponds to the similar open case remaining for real number fields.

They also have extended Cerri's algorithm for the computation of the upper part of the norm-Euclidean spectrum of a number field to this non-commutative context. This algorithm has allowed to compute the exact value of the norm-Euclidean minimum of orders in totally definite quaternion fields over a quadratic number field. This has provided the first known values of this minimum when the base number field has degree strictly greater than 1.

6.4. Modular forms and L-functions

Participant: Henri Cohen.

Members of the team have taken part in an international autumn school on computational number theory at the Izmir Institute of Technology (IZTECH) in 2017. Henri Cohen has transformed his two lectures in book chapters. The text on modular forms [23] presents the (of course extremely condensed) view of the book [6] he has coauthored. The chapter on L-functions [24] is closely related to new developments in PARI/GP.

In [25] the same author explains how to compute Fourier expansions at all cusps of any modular form of integral or half-integral weight thanks to a theorem of Borisov–Gunnells and explicit expansions of Eisenstein series at all cusps. Using this, he gives a number of methods for computing arbitrary Petersson products. Implementations in our PARI/GP software are also described.

A complementary approach using modular symbols is used in [14] by Karim Belabas, Dominique Bernardi and Bernadette Perrin-Riou to compute Manin's constant and the modular degree of elliptic curves defined over $\mathbb{Q}$.

6.5. p-adic rings and geometry

Participant: Xavier Caruso.

In [19], Xavier Caruso, Tristan Vaccon and Thibaut Verron laid the foundations of an algorithmic treatment of rigid p-adic geometry by introducting and studing Gröbner bases over Tate algebras. In addition, they designed a Buchberger-like and a F4-like algorithm for computing such Gröbner bases.

In [22], Xavier Caruso presents a survey on Fontaine's theory of p-adic period rings. These notes are based on a course given jointly by Laurent Berger and Xavier Caruso in Rennes in 2014; their aim is to detail the construction of the rings B_{crys} and B_{dR} (and some of their variants) and state several comparison theorems between étale and crystalline or de Rham cohomologies for p-adic algebraic varieties.

6.6. Geometry

Participant: Aurel Page.

The paper [13], *Can you hear the homology of 3-dimensional drums?* by A. Bartel and A. Page was published in Commentarii Mathematici Helvetici.

6.7. Complex multiplication of abelian varieties and elliptic curves

Participants: Razvan Barbulescu, Sorina Ionica, Chloe Martindale, Enea Milio, Damien Robert.

In [16], Sorina Ionica, former postdoc of the team, and Emmanuel Thomé look at the structure of isogeny graphs of genus 2 Jacobians with maximal real multiplication. They generalise a result of Kohel's describing the structure of the endomorphism rings of the isogeny graph of elliptic curves. Their setting considers genus 2 jacobians with complex multiplication, with the assumptions that the real multiplication subring is maximal and has class number 1. Over finite fields, they derive a depth first search algorithm for computing endomorphism rings locally at prime numbers, if the real multiplication is maximal.

Antonin Riffaut examines in [18] whether there are relations defined over $\mathbb{Q}$ that link (additively or multiplicatively) different singular moduli $j(\tau)$, invariants of elliptic curves with complex multiplication by different quadratic rings.

In [34], Chloe Martindale presents an algorithm to compute higher dimensional Hilbert modular polynomials. She also explains applications of this algorithm to point counting, walking on isogeny graphs, and computing class polynomials.

In [28], Razvan Barbulescu and Sudarshan Shinde (Sorbonne Université) make a complete list of the 1525 infinite families of elliptic curves without CM which have a particular behaviour in the ECM factoring algorithm, the 20 previously known families having been found by ad-hoc methods. The new idea was to use the characterisation of ECM-friendly families in terms of their Galois image and to use the recent progress in the topic of Mazur's program. In particular, for some of the families mentioned theoretical in the literature the article offers the first publication of explicite equations.

E. Milio and D. Robert updated their paper [35] on computing cyclic modular polynomials.

6.8. Pairings

Participant: Razvan Barbulescu.

In [27], Razvan Barbulescu in a joint work with Nadia El Mrabet (École des Mines de Saint-Étienne) et Loubna Ghammam (Bosch) makes a review of the families of elliptic curves for pairing-based cryptology. This was necessary after the invention of a new variant of the NFS algorithm in 2016 by Barbulescu and Taechan Kim, which showed that the previously used key sizes for pairings were insecure. The novelty of this review article is double : first they consider a large number of families, some of which were never analysed in the literature because they were not likely to be the best and secondly they combine in the same article the security analysis of each family with a non-optimized implementation. This allows the industry to select a different family for each type of utilisation of pairings.

6.9. Multiprecision arithmetic

Participant: Fredrik Johansson.

In [17], F. Johansson and I. Blagouchine devise an efficient algorithm to compute the generalized Stieltjes constants $\gamma_n(a)$ to arbitrary precision with rigorous error bounds, for the first time achieving this with low complexity with respect to the order n. The algorithm consists of locating an approximate steepest descent contour and then evaluating the integral numerically in ball arithmetic using the Petras algorithm with a Taylor expansion for bounds near the saddle point. An implementation is provided in the Arb library.

In [26], F. Johansson describes algorithms to compute elliptic functions and their relatives (Jacobi theta functions, modular forms, elliptic integrals, and the arithmetic-geometric mean) numerically to arbitrary precision with rigorous error bounds for arbitrary complex variables. Implementations in ball arithmetic are available in the Arb library. This overview article discusses the standard algorithms from a concrete implementation point of view, and also presents some improvements.

In [21], Fredrik Johansson develops algorithms for real and complex dot product and matrix multiplication in arbitrary-precision floating-point and ball arithmetic. The new methods are implemented in Arb and significantly speed up polynomial operations and linear algebra in high precision.

CAGIRE Project-Team

7. New Results

7.1. A parameter free pressure based approach for simulating flows at all Mach

Participant: Pascal Bruel.

The latest version of the pressure-correction algorithm developed in the last years in close partnership with Prof. E. Dick (Ghent University, Belgium) and Dr. Y. Moguen (UPPA, France) has been published this year [14]. Beyond its promotion [25], future efforts will be directed towards its extension to reacting single phase flows.

7.2. Simulation of a jet in a supersonic cross-flow

Participant: Pascal Bruel.

In our joint work, published this year with our Kazakh colleagues [10], we have focused on detailing the vortical patterns present in the structure of a sonic jet injected normally into a supersonic crossflow. Such a generic flow configuration is considered to feature some prominent characteristics encountered in propulsion system based on supersonic combustion (scramjet). A critical pressure ratio value beyond which new vortical structures are appearing has been evidenced. Contacts have been taken with ONERA to have access to their experimental database which is much more developed than the one we had access to so far. Some specific adaptation will have to be conducted though, in order to adapt the simulation to the slightly different flow configuration considered by ONERA.

7.3. Experiments and simulations related to flows over aerial tanks

Participant: Pascal Bruel.

These results have been obtained in the framework of the cooperation with the National University of Córdoba (Argentina).

- Predicting the pressure loads produced by the atmospheric wind flow over a cylindrical vertical tank

  Pressure distributions obtained using different RANS turbulence models were compared with experimental data obtained in wind tunnel tests for a tank with a closed roof. We have considered a flat shape and a conical roof of 25 degrees. Combinations of aspect ratios equal to 0.5, 1 and 2, and Reynolds numbers equal to 250000, 290000 and 340000 were simulated. The results of the numerical model were compared with those obtained in experimental tests in a wind tunnel of the atmospheric boundary layer using a rigid tank model. We have worked to obtain a stable atmospheric boundary layer in the complete domain using the correct boundary conditions for the implemented RANS models. After that, we have developed numerical simulations for the flow around the tanks inside the atmospheric boundary layer. These results have been published in [17].
- Studying the interaction of a wall with the flow around two cylinders arranged in tandem

For this configuration, the cylinders were immersed in a flow with a boundary layer profile at a subcritical Reynolds number (Re=10000). The three-dimensional transient turbulent flow around the cylinders was simulated numerically using the SAS turbulence model. The effects of wake interference due to both the proximity between the cylinders and their position with respect to the wall were examined through the values of drag, lift and pressure coefficients. The details of the flow fields in the near wake of the cylinders were also studied. The results were compared with experimental and numerical results reported in the literature, and with the case of a single cylinder near a wall. These results have been published in [15]. In parallel, a specific test section for the team's Maveric test facility has been developed. It gives the possibility to accommodate wall mounted cylinder(s) that represent a scaled down version of real horizontal tanks. The objective here is to generate validation data. Particle image velocimetry (PIV) measurements have been carried out during the 1-month stay of Mauro Grioni in Pau in September 2019.

- Simulating the effects of explosions on liquid fuel storage tanks.

  The fast release of energy in explosive processes produces intense shock waves (blast waves). The interaction of these waves with obstacles such as tanks can be extremely destructive. As a first step towards the full simulation of a blast wave with a tank, we have studied the capabilities of OpenFOAM to simulate a blast wave. The numerical results were compared in a cylindrical configuration with the analytical solution provided by the Sedov theory. A special attention has been paid to evaluate the influence of the reconstruction functions in the Euler flux (Kurganov scheme) on the numerical results. The predicted position and velocity of the generated shock wave as well as the pressure jump and its evolution behind the shock were in good agreement with their theoretical counterparts. These results have been published in [16].

7.4. A density-based numerical flux for high order simulation of low Mach flows with acoustics

Participants: Pascal Bruel, Jonathan Jung, Vincent Perrier.

The topic dealt with concerns acoustic computations in low Mach number flows with density based solvers. For ensuring a good resolution of the low Mach number base flow, a scheme able to deal with stationary low Mach number flows is necessary. Previously proposed low Mach number fixes have been tested with acoustic computations. Numerical results prove that they are not accurate for acoustic computations. The issues raised with acoustic computations with low Mach number fixes were studied and a new scheme has been developed, in order to be accurate not only for steady low Mach number flows, but also for acoustic computations. These results have been published in [12].

7.5. Conjugate heat transfer with different fluid-solid physical properties: a new elliptic blending second-moment closure

Participants: Rémi Manceau, Gaëtan Mangeon.

Our approach to model the wall/turbulence interaction, based on Elliptic Blending, was successfully applied to flows with standard thermal boundary conditions at the walls [11]. However, Conjugate Heat Transfer, which couples fluid and solid domains, are particularly challenging for turbulence models. We have developed an innovative model, the Elliptic Blending Differential Flux Model, to account for the influence of various wall thermal boundary conditions on the turbulent heat flux and the temperature variance. An assessment of this new model in Conjugate Heat Transfer has been performed for several values of fluid-solid thermal diffusivity and conductivity ratios. A careful attention is paid to the discontinuity of the dissipation rate associated with the temperature variance at the fluid-solid interface. The analysis is supported by successful comparisons with Direct Numerical Simulations [30].

7.6. Eddy-viscosity models sensitized to buoyancy effects for predicting natural convection flows

Participants: Rémi Manceau, Saad Jameel.

Eddy-viscosity turbulence models have been sensitized to the effects of buoyancy, in order to improve the prediction in natural convection flows. The approach extends in a linear way the constitutive relations for the Reynolds stress and the turbulent heat flux, in order to account for the anisotropic influence of buoyancy. The novelty of this work involves the buoyancy extension applied to two very different eddy-viscosity models, which leads to encouraging results for the highly challenging case of the differentially heated vertical channel [27], [22].

7.7. Development and validation of an improved HTLES approach

Participants: Rémi Manceau, Vladimir Duffal, Hassan Afailal, Franck Mastrippolito, Pascal Bruel.

The HTLES (hybrid temporal LES) approach, developed by the team, has been improved by introducing shielding functions and an internal consistency constraint to enforce the RANS behavior in the near-wall regions [18]. The influence of the underlying closure model was studied by applying HTLES to two RANS models: the k-ω SST and the BL-v2/k. The validation, carried out using two different solvers, Code_Saturne (collaboration with EDF) and Converge-CFD (collaboration with IFPEN), encompassed different type of internal flows: channel, periodic hill, internal combustion engines (in fixed position and with a moving piston), jet in crossflow [21], [9]. The robustness of HTLES to grid coarsening and the accuracy of the predictions at a reduced numerical cost compared to LES was demonstrated. Notably, the capacity of HTLES to provide information on velocity and pressure temporal fluctuations at the wall was assessed, offering a cost-saving alternative to LES to predict unsteady loads. One of the major limitation of HTLES, i.e., the fact that it is based on statistically averaged quantities, which are challenging to estimate in non-stationary flows, has been removed, by approximating the statistical average by a Dynamic Temporal Filter.

7.8. Compact WENO, positivity preserving stabilization of discontinuous Galerkin methods

Participant: Vincent Perrier.

Jointly with Alireza Mazaheri (NASA Langley) and Chi Wang Shu, we have developed a compact WENO stabilization that moreover ensures the positivity of physical quantities. The work was published in [13].

CARDAMOM Project-Team

7. New Results

7.1. Modelling of free surface flows

- Participants: Umberto Bosi, Mathieu Colin, Maria Kazolea, Mario Ricchiuto
- Corresponding member: Maria Kazolea
  This year we continued our work on free surface flow modelling. We can dived our work on four main axis. First, we presented a depth-integrated Boussinesq model for the efficient simulation of nonlinear wave-body interaction [4]. The model exploits a ÂÂunifiedÂÂ Boussinesq framework, i.e. the fluid under the body is also treated with the depth-integrated approach. The unified Boussinesq approach was initially proposed by Jiang [60] and recently analysed by Lannes [67]. The choice of Boussinesq-type equations removes the vertical dimension of the problem, resulting in a wave-body model with adequate precision for weakly nonlinear and dispersive waves expressed in horizontal dimensions only. The framework involves the coupling of two different domains with different flow characteristics. Inside each domain, the continuous spectral/hp element method is used to solve the appropriate flow model since it allows to achieve high-order, possibly exponential, convergence for non-breaking waves. Flux-based conditions for the domain coupling are used, following the recipes provided by the discontinuous Galerkin framework. The main contribution of this work is the inclusion of floating surface-piercing bodies in the conventional depth-integrated Boussinesq framework and the use of a spectral/hp element method for high-order accurate numerical discretization in space. The model is verified using manufactured solutions and validated against published results for wave-body interaction. The model is shown to have excellent accuracy and is relevant for applications of waves interacting with wave energy devices. The outcome of this work is the phd thesis of Uberto Bosi.[1].

  Second, a detailed analysis of undular bore dynamics in channels of variable cross-section is performed and presented in [5] . Two undular bore regimes, low Froude number (LFN) and high Froude number (HFN), are simulated with a Serre Green Naghdi model, and the results are compared with the experiments by Treske (1994). We show that contrary to Favre waves and HFN bores, which are controlled by dispersive non-hydrostatic mechanisms, LFN bores correspond to a hydrostatic phenomenon. The dispersive-like properties of the LFN bores is related to wave refraction on the banks in a way similar to that of edge waves in the near shore. A fully hydrostatic asymptotic model for these dispersive-like bores is derived and compared to the observations, confirming our claim.

  An other part of our last year's work was focused on wave breaking for Boussineq-type equations [3]. The aim of this work was to develop a model able to represent the propagation and transformation of waves in nearshore areas. The focus is on the phenomena of wave breaking, shoaling and run-up. These different phenomena are represented through a hybrid approach obtained by the coupling of non-linear Shallow Water equations with the extended Boussinesq equations of Madsen and Sorensen. The novelty is the switch tool between the two modelling equations: a critical free surface Froude criterion. This is based on a physically meaningful new approach to detect wave breaking, which corresponds to the steepening of the wave's crest which turns into a roller. To allow for an appropriate discretization of both types of equations, we consider a finite element Upwind Petrov Galerkin method with a novel limiting strategy, that guarantees the preservation of smooth waves as well as the monotonicity of the results in presence of discontinuities.We provide a detailed discussion of the implementation of the newly proposed detection method, as well as of two other well known criteria which are used for comparison. An extensive benchmarking on several problems involving different wave phenomena and breaking conditions allows to show the robustness of the numerical method proposed, as well as to assess the advantages and limitations of the different detection methods.

We also continue to work on the modelling of free surface flows by investigating a new family of models, derived from the so-called Isobe-Kakinuma models. The Isobe–Kakinuma model is a system of Euler–Lagrange equations for a Lagrangian approximating Luke's Lagrangian for water waves. In [23], We consider the Isobe–Kakinuma model for two-dimensional water waves in the case of the flat bottom. We show theoretically the existence of a family of small amplitude solitary wave solutions to the Isobe–Kakinuma model in the long wave regime. We have also performed numerical computations for a toy system included large amplitude solitary wave solutions. Our computations suggest the existence of a solitary wave of extreme form with a sharp crest. This models seems very promising for future research.

7.2. Modelling of icing and de-icing of aircrafts

- Participants: Héloïse Beaugendre, Mathieu Colin and Francois Morency
- Corresponding member: Héloïse Beaugendre
  In-flight icing on an aircrafts surface can be a major hazard in aeronautics's safety. Numerical simulations of ice accretion on aircraft is a common procedure to anticipate ice formation when flying in a supercooled water droplets cloud. Numerical simulations bring a better understanding of ice accretion phenomena, performance degradations and lead to even more efficient thermal de-icing systems designs. Such simulations imply modelling the phase change of water and the mass and energy transfers. The Messinger model developed in the 1950 is still used today as a reliable basis for new models development. This model estimates the ice growth rate using mass and energy balances coupled to a runback water flow. The main parameter introduced with this approach is the freezing fraction, denoting the fraction of incoming water that effectively freezes on the airfoil.

  In-flight ice accretion code predictions depend on heat loss over rough surfaces. The equivalent sand grain roughness models the friction coefficient, but an additional model is needed for heat transfer prediction. The turbulent Prandtl number correction and the sublayer Stanton-based models are commonly used. This year, both models have been used with the Spalart-Allmaras turbulence model to predict heat transfer over rough surfaces typical of ice accretion. The objective here is to compare the results of the two models. First, the sublayer Stanton-based model is rewritten in the context of a turbulent Prandtl number correction formulation. Then, the two models are implemented in the open source software, SU2, and then verified and validated for flow over rough flat plates, airfoils, and wings. The two models' predictions evolve differently with the local Reynolds number, but are always within the experimental 20% error margin. The paper related to those developments are under revision.

  In the work [17], the objective is to model an ice accretion on an airfoil using a Messinger-based approach and to make a sensitivity analysis of roughness models on the ice shape. The test case is performed on a 2D NACA0012 airfoil. A typical test case on a NACA0012 airfoil under icing conditions is run and confronted with the literature for verification prior to further investigations. Ice blocks profiles comparisons will highlight the differences implied by the choice of the roughness correction, which impact the heat transfer coefficient.

  Numerical simulation of separated flow around an iced airfoil is still a challenge. Predictions of post-stall aerodynamic performance by RANS models are unsatisfactory. Recent hybrid RANS/LES methods, based on modified DDES models, have shown promising results for separated flow. However, questions still arise about the best compromise between computation time and accuracy for the unsteady 3D simulations. The span width of the domain and the best grid practice to obtain accurate results still has to be investigated, especially taking into account the recent method improvements. A recent method such as shear-layer adapted DDES should give acceptable flow prediction with a relatively coarse mesh. In the paper [18], we further study the effects of span width length on predicted aerodynamic coefficients and on the pressure coefficient. The study is done using the open-source software SU2. The backward facing step and a stalled NACA0012 is used to validate the numerical results. Then, the numerical flow around an iced Model 5-6 is studied, especially the

flow within the separation bubble behind the ice. The accuracy of the CFD results is discussed and a recommendation is made about the span width of the computation domain and the grid size.

In order to spare time and resources while increasing the results' accuracy of the stalled wing configuration's aerodynamic coefficients, the following study [16] offers a parametric grid study for the DDES model. For three different grid refinements, characteristics of lift and eddy phenomena are presented and compared to determine, for an infinite wing, the best compromise between time and resources' consumption, and results's accuracy. Using the open software SU2 6.1 (Stanford University Unstructured), we generate three different types of grid refinements around an airfoil, developed spanwise to obtain a straight wing. On the same stalled configuration for each mesh, CFD solutions are ran with the DDES model, and the raw data are post processed with the open software ParaView 5.6. We then compare the aerodynamic coefficients' distributions obtained by the three mesh. The general modelling of vortex shedding's topology and turbulence viscosity are compared with the literature to ensure the right rendering of vortex structures. Chordwise pressure and friction coefficients' distributions as well as the spanwiselift coefficient are also compared. We conclude with the optimum mesh in term of results and resources' consumption.

7.3. High order embedded and immersed boundary methods

- Participants: Héloïse Beaugendre, Mirco Ciallella, Benjamin Constant and Mario Ricchiuto
- Corresponding member: Héloïse Beaugendre
  In the last years the team has invested some effort in developing the high order embedded method known as "shifted boundary method". In this method, geometrical boundaries are not meshed exactly but embedded in the mesh. Boundary conditions are imposed on a surrogate boundary, toughly defined by the collection of mesh faces “closest” to the true boundary. To recover high order of accuracy, the boundary condition imposed on this surrogate boundary is modified to account for the distance from the true boundary.

  This year's work has focused on two aspects. First, we proposed an efficient extension of the method to elliptic diffusion equations in mixed form (e.g., Darcy flow, heat diffusion problems with rough coefficients, etc.) [10] Our aim is to obtain an improved formulation that, for linear finite elements, is at least second-order accurate for both flux and primary variable, when either Dirichlet or Neumann boundary conditions are applied. Following previous work of Nishikawa and Mazaheri in the context of residual distribution methods, we consider the mixed form of the diffusion equation (i.e., with Darcy-type operators), and introduce an enrichment of the primary variable. This enrichment is obtained exploiting the relation between the primary variable and the flux variable, which is explicitly available at nodes in the mixed formulation. The proposed enrichment mimics a formally quadratic pressure approximation, although only nodal unknowns are stored, similar to a linear finite element approximation. We consider both continuous and discontinuous finite element approximations and present two approaches: a non-symmetric enrichment, which, as in the original references, only improves the consistency of the overall method; and a symmetric enrichment, which enables a full error analysis in the classical finite element context. Combined with the shifted boundary method, these two approaches are extended to high-order embedded computations, and enable the approximation of both primary and flux (gradient) variables with second-order accuracy, independently on the type of boundary conditions applied. We also show that the the primary variable is third-order accurate, when pure Dirichlet boundary conditions are embedded.

  Second, using the same ideas underlying the shifted boundary method, a novel approach to handle shock waves has been proposed. In this method shocks are seen as embedded boundaries on which appropriately shifted jump conditions are imposed, allowing to connect the upstream and downstream domains. This new technique, named "shifted shock-fitting", has been implemented on two-dimensional unstructured grids to deal with shocks by treating them as they were immersed boundary. The new algorithm is aimed at coupling a floating shock-fitting technique with the shifted

boundary method, so far introduced only to simulate flows with embedded boundaries (see [15], full length paper in revision on *J.Comput.Phys.*).
A new PhD in collaboration with ONERA has started (Benjamin Constant 's thesis) involving the numerical simulation of unsteady flows around complex geometries in aeronautics. In the CFD simulation process, mesh generation is the main bottleneck when one wishes to study realistic configurations, such as an aircraft landing gear. A mesh can represent a month to several months of engineer time for a specialist, which is prohibitive in the pre-design phase where several geometries are evaluated in a very short time frame. For this reason, Inria and Onera have been interested for several years in the development of an immersed boundary method, which does not require representing obstacles by mesh conforming to the wall, thus simplifying the generation of mesh. The consideration of the wall is carried out by the introduction of a forcing term at certain points in the vicinity of obstacles. In our approach, this technique is combined with a method of generating adaptive octree cartesian mesh. This allows us to exploit the advantages of Cartesian mesh (generation and rapid adaptation, performance gains of a dedicated Cartesian solver). In order to model the boundary layer, a wall model is used to avoid an extra cost. This method has been implemented for the simulation of steady turbulent flows around geometries studied in compressible aerodynamics (winged fuselage, engine air intake, helicopter fuselage...), providing a very good compromise between the quality of the aerodynamic solution and the time it takes to return the solution from the definition of geometry. However, the quality of the solution obtained by steady simulations is not sufficient to predict the acoustics satisfactorily. Indeed, oscillations appear for certain sizes of interest (such as turbulent viscosity or pressure fluctuations) in the vicinity of the wall. In addition, the passage of grids of different levels causes reflections, which can greatly degrade the prediction of the acoustic solution. The objective of this thesis is to solve these two problems in order to be able to perform unsteady simulations around complex geometries, such as a landing gear. On the one hand, we will study an algorithm to regularize the solution at the points of interest of the IBM method located near the wall. Together, we will also be interested in improving the wall model, in collaboration with modeling specialists from the department. We will do a study by error estimators to analyze the impact of these improvements on the solution. Validations on academic test cases will be carried out. A second step is to improve the transfer of the solution between grids of different levels (for which the mesh size double). We will propose an algorithm to regularize this passage, by geometrically modifying the mesh and by modifying the transfer formula of the solution at the passage of the fitting. A validation will be carried out on an unsteady case of LEISA profile. Finally, a demonstrative application that is both geometrically complex and of acoustic interest will be performed, typically a LAGOON or Gulfstream landing gear.

7.4. Composites Materials

- Participants: Giulia Bellezza, Mathieu Colin and Mario Ricchiuto
- Corresponding member: Mario Ricchuito
  Self-healing is an important phenomenon in new-generation refractory ceramic-matrix composites, obtained by the oxidation of a glass-forming phase in the composite. The dynamics of oxygen diffusion, glass formation and flow are the basic ingredients of a self-healing model that has been developed here in 2D in a trans-verse crack of a mini composite [11]. The presented model can work on a realistic image of the material section and is able to simulate healing and quantify the exposure of the material to oxygen, a prerequisite for its lifetime prediction. Crack reopening events are handled satisfactorily, and secondary healing can be simulated. This papers describes and discusses a typical case in order to show the model potentialities.

  Additional work involve two main topics. The first one is dedicated to the modeling of the propagation of a self-healing oxyde in a crack. The aim here is to introduce new models which describe both the self-healing behavior and oxygen diffusion towards fibers. In general, the evolution of an incompressible fluid can be described by the Navier-Stokes equations. However, we observe that a direct numerical method applied to these equations will induce a significant computational

cost, especially in our case, due to the long lifespan of the material. Thus, alternatively, we derive several asymptotic models obtained by performing a dimensional analysis on the Navier-Stokes equations : we focus here on shallow water models and thin film models. The second aspect under study is more theoretical. We propose in the full generality a link between the BD entropy introduced by D. Bresch and B. Desjardins for the viscous shallow-water equations and the Bernis-Friedman (called BF) dissipative entropy introduced to study the lubrications equations. Different dissipative entropies are obtained playing with the drag terms on the viscous shallow water equations. It helps for instance to prove global existence of nonnegative weak solutions for the lubrication equations starting from the global existence of nonnegative weak solutions for appropriate viscous shallow-water equations.

7.5. Adaptation techniques

- Participants: Nicolas Barral, Héloïse Beaugendre, Luca Cirrottola, Algiane Froehly, Mario Ricchiuto.
- Corresponding member: Nicolas Barral
  In [14] we presented an algorithm to perform PDE-based r-adaptation in three-dimensional numerical simulations of unsteady compressible flows on unstructured meshes. A Laplacian-based model for the moving mesh is used to follow the evolving shock-wave patterns in the fluid flow, while the finite volume ALE formulation of the flow solver is employed to implicitly perform a conservative re-mapping of the solution from the previous to the current mesh, at each time step of the simulation.We show the application of this method to compressible flows on three-dimensional geometries. To this aim, an improved relaxation scheme has been developed in order to preserve the validity of the mesh throughout the time simulation in three dimensions, where the geometrical constraints typically restrict the allowable mesh motion.

  Similar adaptive strategies have been investigate for shallow water flows in the context of space-time residual distribution methods [7]. In the scalar case, these schemes can be designed to be unconditionally (w.r.t. the time step) positive, even on the distorted space-time prisms which arise from moving the nodes of an unstructured triangular mesh. Consequently, a local increase in mesh resolution does not impose a more restrictive stability constraint on the time-step, which can instead be chosen according to accuracy or physical requirements. Moreover, schemes of this type are analogous to conservative ALE formulations and automatically satisfy a discrete geometric conservation law. For shallow water flows over variable bed topography, the so-called C-property (retention of hydrostatic balance between flux and source terms, required to maintain the steady state of still, flat, water) can also be satisfied by considering the mass balance equation in terms of free surface level instead of water depth, even when the mesh is moved. Combined with a simple implementation of Laplacian based r-adaptation, this technique has been shown to allow up to 60% CPU time savings for a given error.

  We also extended these methods to curvilinear coordinates to do shallow water simulations on the sphere for oceanographic applications [2]. To provide enhanced resolution of moving fronts present in the flow we consider adaptive discrete approximations on moving triangulations of the sphere. To this end, we re- state all Arbitrary Lagrangian Eulerian (ALE) transport formulas, as well as the volume transformation laws, for a 2D manifold. Using these results, we write the set of ALE-SWEs on the sphere. We then propose a Residual Distribution discrete approximation of the governing equations. Classical properties as the DGCL and the C-property (well balancedness) are reformulated in this more general context. An adaptive mesh movement strategy is proposed. The discrete framework obtained is thoroughly tested on standard benchmarks in large scale oceanography to prove their potential as well as the advantage brought by the adaptive mesh movement.

CQFD Project-Team

6. New Results

6.1. Power-of-d-Choices with Memory: Fluid Limit and Optimality

Abstract: In multi-server distributed queueing systems, the access of stochastically arriving jobs to resources is often regulated by a dispatcher, also known as load balancer. A fundamental problem consists indesigning a load balancing algorithm that minimizes the delays experienced by jobs. During the last twodecades, the power-of-d-choice algorithm, based on the idea of dispatching each job to the least loadedserver out ofdservers randomly sampled at the arrival of the job itself, has emerged as a breakthroughin the foundations of this area due to its versatility and appealing asymptotic properties. In this paper,we consider the power-of-d-choice algorithm with the addition of a local memory that keeps track of thelatest observations collected over time on the sampled servers. Then, each job is sent to a server withthe lowest observation. We show that this algorithm is asymptotically optimal in the sense that the load balancer can always assign each job to an idle server in the large-system limit. Our results quantify and highlight the importance of using memory as a means to enhance performance in randomized load balancing.

Authors: J. Anselmi (CQFD); F. Dufour (CQFD).

6.2. Hamilton-Jacobi-Bellman Inequality for the Average Control of Piecewise Deterministic Markov Processes

Abstract : The main goal of this work is to study the infinite-horizon long run average continuous-time optimal control problem of piecewise deterministic Markov processes (PDMPs) with the control acting continuously on the jump intensity λ and on the transition measure Q of the process.

Authors : O.L.V. Costa; F. Dufour (CQFD)

6.3. Approximation of discounted minimax Markov control problems and zero-sum Markov games using Hausdorff and Wasserstein distances

Abstract : This work is concerned with a minimax control problem (also known as a robust Markov Decision Process (MDP) or a game against nature) with general state and action spaces under the discounted cost optimality criterion. We are interested in approximating numerically the value function and an optimal strategy of this general discounted minimax control problem. To this end, we derive structural Lipschitz continuity properties of the solution of this robust MDP by imposing suitable conditions on the model, including Lipschitz continuity of the elements of the model and absolute continuity of the Markov transition kernel with respect to some probability measure μ. Then, we are able to provide an approximating minimax control model with finite state and action spaces, and hence computationally tractable, by combining these structural properties with a suitable discretization procedure of the state space (related to a probabilistic criterion) and the action spaces (associated to a geometric criterion). Finally, it is shown that the corresponding approximation errors for the value function and the optimal strategy can be controlled in terms of the discretization parameters. These results are also extended to a two-player zero-sum Markov game.

Authors : F. Dufour (CQFD); T. Prieto-Rumeau

6.4. Combining clustering of variables and feature selection using random forests: the CoV/VSURF procedure

Abstract : Standard approaches to tackle high-dimensional supervised classification problem often include variable selection and dimension reduction procedures. The novel methodology proposed in this paper combines clustering of variables and feature selection. More precisely, hierarchical clustering of variables procedure allows to build groups of correlated variables in order to reduce the redundancy of information and summarizes each group by a synthetic numerical variable. Originality is that the groups of variables (and the number of groups) are unknown a priori. Moreover the clustering approach used can deal with both numerical and categorical variables (i.e. mixed dataset). Among all the possible partitions resulting from dendrogram cuts, the most relevant synthetic variables (i.e. groups of variables) are selected with a variable selection procedure using random forests. Numerical performances of the proposed approach are compared with direct applications of random forests and variable selection using random forests on the original p variables. Improvements obtained with the proposed methodology are illustrated on two simulated mixed datasets (cases n>p and n<p, where n is the sample size) and on a real proteomic dataset. Via the selection of groups of variables (based on the synthetic variables), interpretability of the results becomes easier.

Authors : Marie Chavent (CQFD), Robin Genuer (SISTM), Jerome Saracco (CQFD)

6.5. Statistical model choice including variable selection based on variable importance: A relevant way for biomarkers selection to predict meat tenderness

Abstract : In this work, we describe a new computational methodology to select the best regression model to predict a numerical variable of interest Y and to select simultaneously the most interesting numerical explanatory variables strongly linked to Y. Three regression models (parametric, semi-parametric and non-parametric) are considered and estimated by multiple linear regression, sliced inverse regression and random forests. Both the variables selection and the model choice are computational. A measure of importance based on random perturbations is calculated for each covariate. The variables above a threshold are selected. Then a learning/test samples approach is used to estimate the Mean Square Error and to determine which model (including variable selection) is the most accurate. The R package modvarsel (MODel and VARiable SELection) implements this computational approach and applies to any regression datasets. After checking the good behavior of the methodology on simulated data, the R package is used to select the proteins predictive of meat tenderness among a pool of 21 candidate proteins assayed in semitendinosus muscle from 71 young bulls. The biomarkers were selected by linear regression (the best regression model) to predict meat tenderness. These biomarkers, we confirm the predominant role of heat shock proteins and metabolic ones.

Authors : Marie-Pierre Ellies-Oury, Marie Chavent, Alexandre Conanec, Jérôme Saracco

6.6. Genome sequencing for rightward hemispheric language dominance

Abstract : Most people have left-hemisphere dominance for various aspects of language processing, but only roughly 1 % of the adult population has atypically reversed, rightward hemispheric language dominance (RHLD). The genetic-developmental program that underlies leftward language laterality is unknown, as are the causes of atypical variation. We performed an exploratory whole-genome-sequencing study, with the hypothesis that strongly penetrant, rare genetic mutations might sometimes be involved in RHLD. This was by analogy with situs inversus of the visceral organs (left-right mirror reversal of the heart, lungs etc.), which is sometimes due to monogenic mutations. The genomes of 33 subjects with RHLD were sequenced, and analysed with reference to large population-genetic datasets, as well as thirty-four subjects (14 left-handed) with typical language laterality. The sample was powered to detect rare, highly penetrant, monogenic effects if they would be present in at least 10 of the 33 RHLD cases and no controls, but no individual genes had mutations in more than 5 RHLD cases while being un-mutated in controls. A hypothesis derived from invertebrate mechanisms of left-right axis formation led to the detection of an increased mutation load, in

RHLD subjects, within genes involved with the actin cytoskeleton. The latter finding offers a first, tentative insight into molecular genetic influences on hemispheric language dominance.

Authors : Amaia Carrion-castillo, Lise van der Haegen, Nathalie Tzourio-mazoyer, Tulya Kavaklioglu, Solveig Badillo, Marie Chavent, Jérôme Saracco, Marc Brysbaert, Simon Fisher, Bernard Mazoyer, Clyde Francks

6.7. An Original Methodology for the Selection of Biomarkers of Tenderness in Five Different Muscles

Abstract : For several years, studies conducted for discovering tenderness biomarkers have proposed a list of 20 candidates. The aim of the present work was to develop an innovative methodology to select the most predictive among this list. The relative abundance of the proteins was evaluated on five muscles of 10 Holstein cows: gluteobiceps, semimembranosus, semitendinosus, Triceps brachii and Vastus lateralis. To select the most predictive biomarkers, a multi-block model was used: The Data-Driven Sparse Partial Least Square. Semimembranosus and Vastus lateralis muscles tenderness could be well predicted (R^2= 0.95 and 0.94 respectively) with a total of 7 out of the 5 times 20 biomarkers analyzed. An original result is that the predictive proteins were the same for these two muscles: μ-calpain, m-calpain, h2afx and Hsp40 measured in m. gluteobiceps and μ-calpain, m-calpain and Hsp70-8 measured in m. Triceps brachii. Thus, this method is well adapted to this set of data, making it possible to propose robust candidate biomarkers of tenderness that need to be validated on a larger population.

Authors : Ellies-Oury, M.-P., Lorenzo, H., Denoyelle, C., Conanec, A., Saracco, J., Picard B.

6.8. New Approach Studying Interactions Regarding Trade-Off between Beef Performances and Meat Qualities

Abstract : The beef cattle industry is facing multiple problems, from the unequal distribution of added value to the poor matching of its product with fast-changing demand. Therefore, the aim of this study was to examine the interactions between the main variables, evaluating the nutritional and organoleptic properties of meat and cattle performances, including carcass properties, to assess a new method of managing the trade-off between these four performance goals. For this purpose, each variable evaluating the parameters of interest has been statistically modeled and based on data collected on 30 Blonde d'Aquitaine heifers. The variables were obtained after a statistical pre-treatment (clustering of variables) to reduce the redundancy of the 62 initial variables. The sensitivity analysis evaluated the importance of each independent variable in the models, and a graphical approach completed the analysis of the relationships between the variables. Then, the models were used to generate virtual animals and study the relationships between the nutritional and organoleptic quality. No apparent link between the nutritional and organoleptic properties of meat (r = -0.17) was established, indicating that no important trade-off between these two qualities was needed. The 30 best and worst profiles were selected based on nutritional and organoleptic expectations set by a group of experts from the INRA (French National Institute for Agricultural Research) and Institut de l'Elevage (French Livestock Institute). The comparison between the two extreme profiles showed that heavier and fatter carcasses led to low nutritional and organoleptic quality.

Authors : Conanec ,A., Picard, B., Cantalapiedra-Hijar, G., Chavent, M., Denoyelle, C., Gruffat, D., Normand, J., Saracco, J., Ellies-Oury M.P.

6.9. Impact of Speller Size on a Visual P300 Brain-Computer Interface (BCI) System under Two Conditions of Constraint for Eye Movement

Abstract : The vast majority of P300-based brain-computer interface (BCI) systems are based on the well-known P300 speller presented by Farwell and Donchin for communication purposes and an alternative to people with neuromuscular disabilities, such as impaired eye movement. The purpose of the present work is to study the effect of speller size on P300-based BCI usability, measured in terms of effectiveness, efficiency, and

satisfaction under overt and covert attention conditions. To this end, twelve participants used three speller sizes under both attentional conditions to spell 12 symbols. The results indicated that the speller size had, in both attentional conditions, a significant influence on performance. In both conditions (covert and overt), the best performances were obtained with the small and medium speller sizes, both being the most effective. The speller size did not significantly affect workload on the three speller sizes. In contrast, covert attention condition produced very high workload due to the increased resources expended to complete the task. Regarding users' preferences, significant differences were obtained between speller sizes. The small speller size was considered as the most complex, the most stressful, the less comfortable, and the most tiring. The medium speller size was always considered in the medium rank, which is the speller size that was evaluated less frequently and, for each dimension, the worst one. In this sense, the medium and the large speller sizes were considered as the most satisfactory. Finally, the medium speller size was the one to which the three standard dimensions were collected: high effectiveness, high efficiency, and high satisfaction. This work demonstrates that the speller size is an important parameter to consider in improving the usability of P300 BCI for communication purposes. The obtained results showed that using the proposed medium speller size, performance and satisfaction could be improved.

Authors :Ron-Angevin, R., Garcia, L., Fernandez-Rodriguez, A., Saracco, J., André, J.-M., Lespinet-Najib, V.

6.10. High-Dimensional Multi-Block Analysis of Factors Associated with Thrombin Generation Potential

Abstract : The identification of novel biological factors associated with thrombin generation, a key biomarker of the coagulation process, remains a relevant strategy to disentangle pathophysiological mechanisms underlying the risk of venous thrombosis (VT). As part of the MARseille THrombosis Association Study (MARTHA), we measured whole blood DNA methylation levels, plasma levels of 300 proteins, 3 thrombin generation biomarkers (endogeneous thrombin potential, peak and lagtime), clinical and genetic data in 700 patients with VT. The application of a novel high-dimensional multi-levels statistical methodology we recently developed, the data driven sparse Partial Least Square method (ddsPLS), on the MARTHA datasets enabled us 1/ to confirm the role of a known mutation of the variability of endogenous thrombin potential and peak, 2/ to identify a new signature of 7 proteins strongly associated with lagtime.

Authors : Lorenzo, H., Razzaq, M., Odeberg, J., Saracco, J., Tregouet, D.-A., Thiébaut, R.

6.11. Multiple-output quantile regression through optimal quantization

Abstract : A new nonparametric quantile regression method based on the concept of optimal quantization was developed recently and was showed to provide estimators that often dominate their classical, kernel-type, competitors. In the present work, we extend this method to multiple-output regression problems. We show how quantization allows approximating population multiple-output regression quantiles based on halfspace depth. We prove that this approximation becomes arbitrarily accurate as the size of the quantization grid goes to infinity. We also derive a weak consistency result for a sample version of the proposed regression quantiles. Through simulations, we compare the performances of our estimators with (local constant and local bilinear) kernel competitors. The results reveal that the proposed quantization-based estimators, which are local constant in nature, outperform their kernel counterparts and even often dominate their local bilinear kernel competitors. The various approaches are also compared on artificial and real data.

Authors : Charlier, I., Paindaveine, D., Saracco, J.

6.12. Artificial evolution, fractal analysis and applications

Abstract :

This document contains a selection of research works to which I have contributed. It is structured around two themes, artificial evolution and signal regularity analysis and consists of three main parts: Part I: Artificial evolution, Part II: Estimation of signal regularity and Part III: Applications, combination of signal processing, fractal analysis and artificial evolution. In order to set the context and explain the coherence of the rest of the document, this manuscript begins with an introduction, Chapter 1, providing a list of collaborators and of the research projects carried out. Theoretical contributions focus on two areas: evolutionary algorithms and the measurement of signal regularity and are presented in Part I and Part II respectively. These two themes are then exploited and applied to real problems in Part III. Part I, Artificial Evolution, consists of 8 chapters. Chapter 2 contains a brief presentation of various types of evolutionary algorithms (genetic algorithms, evolutionary strategies and genetic programming) and presents some contributions in this area, which will be detailed later in the document. Chapter 3, entitled Prediction of Expected Performance for a Genetic Programming Classifier proposes a method to predict the expected performance for a genetic programming (GP) classifier without having to run the program or sample potential solutions in the research space. For a given classification problem, a pre-processing step to simplify the feature extraction process is proposed. Then the step of extracting the characteristics of the problem is performed. Finally, a PEP (prediction of expected performance) model is used, which takes the characteristics of the problem as input and produces the predicted classification error on the test set as output. To build the PEP model, a supervised learning method with a GP is used. Then, to refine this work, an approach using several PEP models is developed, each now becoming a specialized predictors of expected performance (SPEP) specialized for a particular group of problems. It appears that the PEP and SPEP models were able to accurately predict the performance of a GP-classifier and that the SPEP approach gave the best results. Chapter 4, entitled A comparison of fitness-case sampling methods for genetic programming presents an extensive comparative study of four fitness-case sampling methods, namely: Interleaved Sampling, Random Interleaved Sampling, Lexicase Selection and the proposed Keep-Worst Interleaved Sampling. The algorithms are compared on 11 symbolic regression problems and 11 supervised classification problems, using 10 synthetic benchmarks and 12 real-world datasets. They are evaluated based on test performance, overfitting and average program size, comparing them with a standard GP search. The experimental results suggest that fitness-case sampling methods are particularly useful for difficult real-world symbolic regression problems, improving performance, reducing overfitting and limiting code growth. On the other hand, it seems that fitness-case sampling cannot improve upon GP performance when considering supervised binary classification. Chapter 5, entitled Evolving Genetic Programming Classifiers with Novelty Search, deals with a new and unique approach towards search and optimization, the Novelty Search (NS), where an explicit objective function is replaced by a measure of solution novelty. This chapter proposes a NS-based GP algorithm for supervised classification. Results show that NS can solve real-world classification tasks, the algorithm is validated on real-world benchmarks for binary and multiclass problems. Moreover, two new versions of the NS algorithm are proposed, Probabilistic NS (PNS) and a variant of Minimal Criteria NS (MCNS). The former models the behavior of each solution as a random vector and eliminates all of the original NS parameters while reducing the computational overhead of the NS algorithm. The latter uses a standard objective function to constrain and bias the search towards high performance solutions. This chapter also discusses the effects of NS on GP search dynamics and code growth. The results show that NS can be used as a realistic alternative for supervised classification, and specifically for binary problems the NS algorithm exhibits an implicit bloat control ability. In Chapter 6, entitled Evaluating the Effects of Local Search in Genetic Programming, a memetic GP that incorporates a local search (LS) strategy to refine GP individuals expressed as syntax trees is studied in the context of symbolic regression. A simple parametrization for GP trees is proposed, by weighting each function with a parameter (unique for each function used in the construction of a tree). These parameters are then optimized using a trust region optimization algorithm which is therefore used here as a local search method. Then different heuristic methods are tested over several benchmark and real-world problems to determine which individuals from the tree population should be subjected to a LS. The results show that the best performances (in term of both quality of the solution and bloat control) was achieved when LS is applied to all of the solutions or to random individuals chosen from the top percentile (with respect to fitness) of the population. Chapter 7, entitled A Local Search Approach to Genetic Programming for Binary Classification, proposes a memetic GP, tailored for binary classification problems, extending the work on symbolic regression presented in the

previous chapter. In particular, a small linear subtree is added on the top of the root node of the original tree and each node in a tree is weighted by a real-valued parameter, which is then numerically optimized using the trust-region algorithm used as a local search method. Experimental results show that potential classifiers produced by GP are improved by the local searcher, and hence the overall search is improved achieving substantial performance gains. Application on well-known benchmarks provided results competitive with state-of-the-art. Chapter 8, entitled RANSAC-GP: Dealing with Outliers in Symbolic Regression with Genetic Programming, presents a hybrid methodology based on the RAndom SAmpling Consensus (RANSAC) algorithm and GP, called RANSAC-GP. RANSAC is an approach to deal with outliers in parameter estimation problems, widely used in computer vision and related fields. This work presents the first application of RANSAC to symbolic regression with GP. The proposed algorithm is able to deal with extreme amounts of contamination in the training set, evolving highly accurate models even when the amount of outliers reaches 90Part II, Estimation of signal regularity consists of 3 chapters. Chapter 9, entitled Hölderian Regularity, provides some reminders and some theoretical contributions on the estimation of Hölderian regularity. Some details are given on the estimation of the Hölder exponent using oscillation method or a wavelet transform. These approaches and improved versions are compared on synthetic signals. The FracLab software, were all the above methods have been integrated, is also presented at the end of this chapter. The work proposed in Chapter 10, entitled Theoretical comparison of the DFA and variants for the estimation of the Hurst exponent, involves a theoretical and numerical comparison between the Detrended Fluctuation Analysis (DFA) and its variants, namely DMA, AFA, RDFA and the proposed Continuous DFA method, in which the trend is constrained to be continuous. The DFA is a well-established method to detect long-range correlations in time series. It has been used in a wide range of applications, from biomedical applications to signal denoising. It allows the Hurst exponent of a pure mono-fractal time series to be estimated. It operates as follows: after integration, the signal is split into segments. Using a least-squares criterion, local trends are deduced. The resulting piecewise linear trend is then subtracted to the whole signal. The power of the residual is computed for different segment lengths and its log-log representation allows the Hurst exponent to be deduced. The comparison performed in this chapter is based on a new common matrix writing formalism of the square of the fluctuation function from the instantaneous correlation function of the process for all these methods. In the case where the process under study is stationary in the broad sense, the statistical mean of the square of the fluctuation function is thus expressed as a weighted sum of the terms of the autocorrelation function, and this without any approximation. More precisely, the mathematical expectation of the square of the fluctuation function can be seen for each method as the autocorrelation function of the output of a filter dependent on this method and calculated for a lag equal to zero, i.e. the power of the filter output. In the general case, this analytical framework provides a means of comparing the DFA and its variants that is different from a traditional synthetic signal performance study, and explains the different behaviours of these regularity estimation methods, using the proposed filter analysis. Chapter 11 contains two patents with THALES AVS related to the work presented in the previous chapter. Part III of this manuscript, Applications, combination of signal processing, fractal analysis and artificial evolution, contains contributions combining the tools previously mentioned in order to develop new tools such as in the Chapters 12 and 13 or contributions on the resolution of real problems in the biomedical field, such as in the Chapters 14, 15 and 16. Chapter 12, entitled "The Estimation of Hölderian Regularity using Genetic Programming", presents a GP approach to synthesize estimators for the pointwise Hölder exponent in 2D signals. The optimization problem to solve is to minimize the error between a prescribed regularity and the estimated regularity given by an image operator. The search for optimal estimators is then carried out using a GP algorithm. Experiments confirm that the GPoperators produce a good estimation of the Hölder exponent in images of multifractional Brownian motions. In fact, the evolved estimators significantly outperform a traditional method by as much as one order of magnitude. These results provide further empirical evidence that GP can solve difficult problems of applied mathematics. In Chapter 13, entitled "Optimization of the Hölder Image Descriptor using a Genetic Algorithm", a local descriptor based on the Hölder exponent is studied. The proposal is to find an optimal number of dimensions for the descriptor using a genetic algorithm (GA). To guide the GA search, fitness is computed based on the performance of the descriptor when applied to standard region matching problems. This criterion is quantified using the F-Measure, derived from recall and precision analysis. Results show that it is possible to reduce the size of the canonical Hölder descriptor without degrading the quality of its performance. In fact, the best descriptor found through the GA search is nearly 70performance

on standard tests. Chapter 14, entitled "Interactive evolution for cochlear implants fitting", presents a study that intends to make cochlear implants more adaptable to environment and to simplify the process of fitting, by designing and using a specific interactive evolutionary algorithm combined with signal processing. Real experiments on volunteer implanted patients are presented, that show the efficiency of interactive evolution for this purpose. In Chapter 15, entitled "Feature extraction and classification of EEG signals. The use of a genetic algorithm for an application on alertness prediction", the development of computer systems for the automatic analysis and classification of mental states of vigilance; i.e., a person's state of alertness is studied. Such a task is relevant to diverse domains, where a person is expected or required to be in a particular state. For instance, pilots, security personnel or medical staffs are expected to be in a highly alert state, and a brain computer interface could help confirm this or detect possible problems. In this chapter, a combination of an evolutionary algorithm and signal processing is used. The purpose of this algorithm was to select an electrode and a frequency range to use in order to discriminate between the two states of vigilance. This approach determined the most useful electrode for the classification task. Using the recording of this electrode, the prediction obtained has a reliability rate of 89.33

In Chapter 16, entitled "Regularity and Matching Pursuit Feature Extraction for the Detection of Epileptic Seizures", a novel methodology for feature extraction on EEG signals that allows to perform a highly accurate classification of epileptic states is presented. Specifically, Hölderian regularity and the Matching Pursuit algorithm are used as the main feature extraction techniques, and are combined with basic statistical features to construct the final feature sets. These sets are then delivered to a Random Forests classification algorithm to differentiate between epileptic and non-epileptic readings. Several versions of the basic problem are tested and statistically validated producing perfect accuracy in most problems and 97.6a well known database, reveals that the proposal achieves state-of-the-art performance.The experimental results suggest that using a feature extraction methodology composed of regularity analysis, a Matching Pursuit algorithm and time-domain statistic measures together with a classifier produces a system that can predict epileptic states with competitive performance that matches or even surpass other novel methods. Finally the last chapter concludes this manuscript and provides perspectives for future work.

Authors : Pierrick Legrand

6.13. Self-affinity of an Aircraft Pilot's Gaze Direction as a Marker of Visual Tunneling

Abstract : For the last few years, a great deal of interest has been paid to crew monitoring systems in order to tackle potential safety problems during a flight. They aim at detecting any degraded physiological and/or cognitive state of an aircraft pilot, such as attentional tunneling or excessive focalization. Indeed, they might have a negative impact on his performance to pursue his mission with adequate flight safety levels. One of the usual approaches consists in using sensors to collect physiological signals which are analyzed in real-time. Two main families exist to process the signals. The first one combines feature extraction and machine learning whereas the second is based on deep-learning approaches but may require a large amount of labelled data. Here, we focused on the first family. In this case, various features can be deduced from the data by different approaches: spectrum analysis, a priori modelling and nonlinear dynamical system analysis techniques including the estimation of the self-affinity of the signals. In this paper, our purpose was to analyze whether the self-affinity of the pilot gaze direction can be related to his cognitive state. To this end, an experiment was carried out on 18 subjects in a representative aircraft environment based on a modified version of the software MATB-II. The scenarii were designed to elicit different levels of mental workload eventually associated to attentional tunneling. A database to train the machine learning step was first created by recording the directions of gaze of the subjects with an eye-tracker. The self-affinities of these signals were extracted with the Detrended Fluctuation Analysis method. They constituted the inputs of the classifier based on a Support Vector Machine. Then, new signals were analyzed and classified. Preliminary results showed promising abilities to detect attentional tunnelling episodes for different levels of mental workload.

Authors : Bastien Berthelot, Patrick Mazoyer, Sarah Egea, Jean-Marc André, Eric Grivel

6.14. Filtering-based Analysis Comparing the DFA with the CDFA for Wide Sense Stationary Processes

Abstract : The detrended fluctuation analysis (DFA) is widely used to estimate the Hurst exponent. Although it can be outperformed by wavelet based approaches, it remains popular because it does not require a strong expertise in signal processing. Recently, some studies were dedicated to its theoretical analysis and its limits. More particularly, some authors focused on the so-called fluctuation function by searching a relation with an estimation of the normalized covariance function under some assumptions. This paper is complementary to these works. We first show that the square of the fluctuation function can be expressed in a similar matrix form for the DFA and the variant we propose, called Continuous-DFA (CDFA), where the global trend is constrained to be continuous. Then, using the above representation for wide-sense-stationary processes, the statistical mean of the square of the fluctuation function can be expressed from the correlation function of the signal and consequently from its power spectral density, without any approximation. The differences between both methods can be highlighted. It also confirms that they can be seen as ad hoc wavelet based techniques.

Authors : Bastien Berthelot, Eric Grivel, Pierrick Legrand, Jean-Marc André, Patrick Mazoyer, et al

6.15. 2D Fourier Transform Based Analysis Comparing the DFA with the DMA

Abstract : Even if they can be outperformed by other methods, the detrended fluctuation analysis (DFA) and the detrended moving average (DMA) are widely used to estimate the Hurst exponent because they are based on basic notions of signal processing. For the last years, a great deal of interest has been paid to compare them and to better understand their behaviors from a mathematical point of view. In this paper, our contribution is the following: we first propose to express the square of the socalled fluctuation function as a 2D Fourier transform (2D-FT) of the product of two matrices. The first one is defined from the instantaneous correlations of the signal while the second, called the weighting matrix, is representative of each method. Therefore, the 2D-FT of the weighting matrix is analyzed in each case. In this study, differences between the DFA and the DMA are pointed out when the approaches are applied on non-stationary processes

Authors : Bastien Berthelot, Eric Grivel, Pierrick Legrand, Marc Donias, Jean-Marc André, et al.

6.16. Interpréter les Fonctions de Fluctuation du DFA et du DMA comme le Résultat d'un Filtrage

Abstract : The detrended uctuation analysis (DFA) and the detrending moving average (DMA) are often used to estimate the regularity of the signal, since they do not require a strong expertise in the eld of signal processing while providing good results. In this paper, our contribution is twofold. We propose a framework that allows these approaches to be compared. It is based on a matrix form of the square of the uctuation function. Using the above representation for wide-sense-stationary processes, we show that the statistical mean of the square of the uctuation function can be expressed from the correlation function of the signal and consequently from its power spectral density, without any approximation. The dierences between both methods can be highlighted. It also conrms that they can be seen as ad hoc wavelet based techniques to estimate the Hurst exponent.

Authors : Bastien Berthelot, Eric Grivel, Pierrick Legrand, Jean-Marc Andre, Patrick Mazoyer, et al.

6.17. A perturbation analysis of stochastic matrix Riccati diffusions

Abstract : Matrix differential Riccati equations are central in filtering and optimal control theory. The purpose of this article is to develop a perturbation theory for a class of stochastic matrix Riccati diffusions. Diffusions of this type arise, for example, in the analysis of ensemble Kalman-Bucy filters since they describe the flow of certain sample covariance estimates. In this context, the random perturbations come from the fluctuations of a mean field particle interpretation of a class of nonlinear diffusions equipped with an interacting sample covariance matrix functional. The main purpose of this article is to derive non-asymptotic

Taylor-type expansions of stochastic matrix Riccati flows with respect to some perturbation parameter. These expansions rely on an original combination of stochastic differential analysis and nonlinear semigroup techniques on matrix spaces. The results here quantify the fluctuation of the stochastic flow around the limiting deterministic Riccati equation, at any order. The convergence of the interacting sample covariance matrices to the deterministic Riccati flow is proven as the number of particles tends to infinity. Also presented are refined moment estimates and sharp bias and variance estimates. These expansions are also used to deduce a functional central limit theorem at the level of the diffusion process in matrix spaces.

Authors : Adrian N. Bishop, Pierre Del Moral, Angele Niclas

6.18. On the stability of matrix-valued Riccati diffusions

Abstract : The stability properties of matrix-valued Riccati diffusions are investigated. The matrix-valued Riccati diffusion processes considered in this work are of interest in their own right, as a rather prototypical model of a matrix-valued quadratic stochastic process. Under rather natural observability and controllability conditions, we derive time-uniform moment and fluctuation estimates and exponential contraction inequalities. Our approach combines spectral theory with nonlinear semigroup methods and stochastic matrix calculus. This analysis seem to be the first of its kind for this class of matrix-valued stochastic differential equation. This class of stochastic models arise in signal processing and data assimilation, and more particularly in ensemble Kalman-Bucy filtering theory. In this context, the Riccati diffusion represents the flow of the sample covariance matrices associated with McKean-Vlasov-type interacting Kalman-Bucy filters. The analysis developed here applies to filtering problems with unstable signals.

Authors : Adrian N. Bishop, Pierre Del Moral

6.19. A variational approach to nonlinear and interacting diffusions

Abstract : The article presents a novel variational calculus to analyze the stability and the propagation of chaos properties of nonlinear and interacting diffusions. This differential methodology combines gradient flow estimates with backward stochastic interpolations, Lyapunov linearization techniques as well as spectral theory. This framework applies to a large class of stochastic models including non homogeneous diffusions, as well as stochastic processes evolving on differentiable manifolds, such as constraint-type embedded manifolds on Euclidian spaces and manifolds equipped with some Riemannian metric. We derive uniform as well as almost sure exponential contraction inequalities at the level of the nonlinear diffusion flow, yielding what seems to be the first result of this type for this class of models. Uniform propagation of chaos properties w.r.t. the time parameter are also provided. Illustrations are provided in the context of a class of gradient flow diffusions arising in fluid mechanics and granular media literature. The extended versions of these nonlinear Langevin-type diffusions on Riemannian manifolds are also discussed.

Authors : Marc Arnaudon (IMB), Pierre Del Moral (CMAP, CQFD)

6.20. An explicit Floquet-type representation of Riccati aperiodic exponential semigroups

Abstract : The article presents a rather surprising Floquet-type representation of time-varying transition matrices associated with a class of nonlinear matrix differential Riccati equations. The main difference with conventional Floquet theory comes from the fact that the underlying flow of the solution matrix is aperiodic. The monodromy matrix associated with this Floquet representation coincides with the exponential (fundamental) matrix associated with the stabilizing fixed point of the Riccati equation. The second part of this article is dedicated to the application of this representation to the stability of matrix differential Riccati equations. We provide refined global and local contraction inequalities for the Riccati exponential semigroup that depend linearly on the spectral norm of the initial condition. These refinements improve upon existing results and are a direct consequence of the Floquet-type representation, yielding what seems to be the first results of this type for this class of models.

Authors : Adrian N. Bishop, Pierre Del Moral

6.21. Uniform propagation of chaos and creation of chaos for a class of nonlinear diffusions

Abstract : We are interested in nonlinear diffusions in which the own law intervenes in the drift. This kind of diffusions corresponds to the hydrodynamical limit of some particle system. One also talks about propagation of chaos. It is well-known, for McKean-Vlasov diffusions, that such a propagation of chaos holds on finite-time interval. We here aim to establish a uniform propagation of chaos even if the external force is not convex, with a diffusion coefficient sufficiently large. The idea consists in combining the propagation of chaos on a finite-time interval with a functional inequality, already used by Bolley, Gentil and Guillin, see [BGG12a, BGG12b]. Here, we also deal with a case in which the system at time t = 0 is not chaotic and we show under easily checked assumptions that the system becomes chaotic as the number of particles goes to infinity together with the time. This yields the first result of this type for mean field particle diffusion models as far as we know.

Authors : Pierre Del Moral and Julian Tugaut

6.22. Stability Properties of Systems of Linear Stochastic Differential Equations with Random Coefficients

Abstract : This work is concerned with the stability properties of linear stochastic differential equations with random (drift and diffusion) coefficient matrices, and the stability of a corresponding random transition matrix (or exponential semigroup). We consider a class of random matrix drift coefficients that involves random perturbations of an exponentially stable flow of deterministic (time-varying) drift matrices. In contrast with more conventional studies, our analysis is not based on the existence of Lyapunov functions, and it does not rely on any ergodic properties. These approaches are often difficult to apply in practice when the drift/diffusion coefficients are random. We present rather weak and easily checked perturbation-type conditions for the asymptotic stability of time-varying and random linear stochastic differential equations. We provide new log-Lyapunov estimates and exponential contraction inequalities on any time horizon as soon as the fluctuation parameter is sufficiently small. These seem to be the first results of this type for this class of linear stochastic differential equations with random coefficient matrices.

Authors : Adrian N. Bishop, Pierre Del Moral

6.23. On One-Dimensional Riccati Diffusions

Abstract : This article is concerned with the fluctuation analysis and the stability properties of a class of one-dimensional Riccati diffusions. These one-dimensional stochastic differential equations exhibit a quadratic drift function and a non-Lipschitz continuous diffusion function. We present a novel approach, combining tangent process techniques, Feynman-Kac path integration, and exponential change of measures, to derive sharp exponential decays to equilibrium. We also provide uniform estimates with respect to the time horizon, quantifying with some precision the fluctuations of these diffusions around a limiting deterministic Riccati differential equation. These results provide a stronger and almost sure version of the conventional central limit theorem. We illustrate these results in the context of ensemble Kalman-Bucy filtering. To the best of our knowledge, the exponential stability and the fluctuation analysis developed in this work are the first results of this kind for this class of nonlinear diffusions.

Authors: Adrian N. Bishop, Pierre Del Moral, Kengo Kamatani, Bruno Remillard

6.24. Adaptive Approximate Bayesian Computational Particle Filters for Underwater Terrain-Aided Navigation.

Authors : C. Palmier, K. Dahia, N. Merlinge, P. Del Moral, D. Laneuville & C. Musso

6.25. Inference for conditioned Galton-Watson trees from their Harris path

Tree-structured data naturally appear in various fields, particularly in biology where plants and blood vessels may be described by trees, but also in computer science because XML documents form a tree structure. This paper is devoted to the estimation of the relative scale parameter of conditioned GaltonWatson trees. New estimators are introduced and their consistency is stated. A comparison is made with an existing approach of the literature. A simulation study shows the good behavior of our procedure on finite-sample sizes and from missing or noisy data. An application to the analysis of revisions of Wikipedia articles is also considered through real data.

Authors: Romain Azaïs, Alexandre Genadot and Benoit Henry

GEOSTAT Project-Team

7. New Results

7.1. Excitable systems

Participants: G. Attuel , E. Gerasimova-Chechkina , F. Argoul , H. Yahia , A. Arnéodo .

In a companion paper (I. Multifractal analysis of clinical data), we used a wavelet-based multiscale analysis to reveal and quantify the multifractal intermittent nature of the cardiac impulse energy in the low frequency range 2Hz during atrial fibrillation (AF). It demarcated two distinct areas within the coronary sinus (CS) with regionally stable multifractal spectra likely corresponding to different anatomical substrates. The electrical activity also showed no sign of the kind of temporal correlations typical of cascading processes across scales, thereby indicating that the multifractal scaling is carried by variations in the large amplitude oscillations of the recorded bipolar electric potential. In the present study, to account for these observations, we explore the role of the kinetics of gap junction channels (GJCs), in dynamically creating a new kind of imbalance between depolarizing and repolarizing currents. We propose a one-dimensional (1D) spatial model of a denervated myocardium, where the coupling of cardiac cells fails to synchronize the network of cardiac cells because of abnormal transjunctional capacitive charging of GJCs. We show that this non-ohmic nonlinear conduction 1D modeling accounts quantitatively well for the "multifractal random noise" dynamics of the electrical activity experimentally recorded in the left atrial posterior wall area. We further demonstrate that the multifractal properties of the numerical impulse energy are robust to changes in the model parameters.

Publications: Frontiers in Physiology, Frontiers, 2019, 10, pp.480 (1-18), HAL

Second international Summer Institute on Network Physiology, Jul 2019 HAL

7.2. Differential diagnosis between atypical Parkinsonian syndromes

Participants: B. Das , K. Daoudi , J. Kemplir , J. Rusz .

In the early stage of disease, the symptoms of Parkinson's disease (PD) are similar to atypical Parkinsonian syndromes (APS). The early differential diagnosis between PD and APS and within APS is thus a very challenging task. It turns out that speech disorder is an early and common symptom to PD and APS. The goal of research is to develop a digital marker based on speech analysis in order to assist the neurologists in their diagnosis. We addressed the problem of differential diagnosis between two APS, Progressive Supranuclear Palsy (PSP) and Multiple System Atrophy (MSA). Using linear discriminant analysis we designed an hypokinetic and an ataxic dysarthria measure which yield a discrimination between PSP and MSA with a high accuracy (88%). This result indicates that hypokinetic and ataxic dysarthria convey valuable discriminative information when mutually considered.

Publication: ICASSP 2019 HAL

7.3. Turbulent dynamics of ocean upwelling

Participants: A. El Aouni , V. Garçon , J. Sudre , H. Yahia , K. Daoudi , K. Minaoui .

The region along the NorthWest African coast (20° N to 36° N and 4° W to 19° W) is characterized by a persistent and variable upwelling phenomenon almost all year round. In this article, the upwelling features are investigated using an algorithm dedicated to delimit the upwelling area from thermal and biological satellite observations. This method has been developed specifically for sea-surface temperature (SST) images, since they present a high latitudinal variation, which is not present in chlorophyll-a concentration images. Developing on the proposed approach, the spatial and temporal variations of the main physical and biological upwelling patterns are studied. Moreover, a study on the upwelling dynamics, which explores the interplay between the upwelling spatiotemporal extents and intensity, is presented, based on a 14-year time archive of weekly SST and chlorophyll-a concentration data.

Publication: IEEE Transactions on Geoscience and Remote Sensing, HAL

7.4. Fourier approach to Lagrangian vortex detection

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

We study the transport properties of coherent vortices over a finite time duration. Here we reveal that such vortices can be identified based on frequency-domain representation of Lagrangian trajectories. We use Fourier analysis to convert particles' trajectories from their time domain to a presentation in the frequency domain. We then identify and extract coherent vortices as material surfaces along which particles' trajectories share similar frequencies. Our method identifies all coherent vortices in an automatic manner, showing high vortices' monitoring capacity. We illustrate our new method by identifying and extracting Lagrangian coherent vortices in different two-and three-dimensional flows.

Publication: Chaos, American Institute of Physics, HAL

7.5. Surface mixing and biological activity

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

Near-shore water along the NorthWest African margin is one of the world's major upwelling regions. It is associated with physical structures of oceanic fronts which influence the biological productivity. The study of these coherent structures in connection with chlorophyll concentration data is of fundamental importance for understanding the spatial distributions of the plankton. In this work, we study the horizontal stirring and mixing in different upwelling areas using Lagrangian coherent structures (LCSs). These LCSs are calculated using the recent geodesic theory of LCSs. We use these LCSs to study the link between the chlorophyll fronts concentrations and surface mixing, based on 10 years of satellite data. These LCSs move with the flow as material lines, thus the horizontal mixing is calculated from the intersection of these LCSs with the finite time Lyapunov exponents (FTLEs) maps. We compare our results with those of a recent study conducted over the same area, but based on Finite Size Lyapunov Exponents (FSLEs) whose output is a plot of scalar distributions. We discuss the differences between FSLE and geodesic theory of LCS. The latter yields analytical solutions of LCSs, while FSLEs can only provide LCSs for sharp enough ridges of nearly constant height.

Publication: Chaos, American Institute of Physics, HAL

7.6. Contribution and influence of coherent mesoscale eddies off the North-West African upwelling on the open ocean

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

Eastern Boundary Upwelling zones include some of the most productive ecosystems in the world, particularly the NorthWest (NW) African upwelling which presents one of the world's major upwelling regions. This latter is forced by the equator-ward trade winds which are known to exhibit mesoscale instabilities; thus, in addition to upwelled cold and nutrient-rich deep waters, significant energy is transferred into mesoscale fronts and eddies in the upper ocean. Oceanic structures of type eddies are well know to stir and mix surrounding water masses. However, they can also carry and transport organic matter and marine in a coherent manner. Here, we are interested in those that remain coherent. The Aim of this work, is to understand the impact and the contribution of the mesoscale eddies off the NW African margin on the open ocean. Our approach to analyze such coherent eddies is based upon the use of our recently developed technique from nonlinear dynamics theory, which is capable of identifying coherent vortices and their centers in an automatic manner. The role of these mesoscale eddies is investigated based on a statistical study of eddies properties off NW African margin (cyclone/anticyclone, their lifetimes, traveled distance, translational speeds, quantity of water masses transported to the open ocean...). This statistical study is carried out over a sets of 24 years (spans from January 1993 to December 2016) of sea surface velocity field derived from satellite surface altimetry under the geostrophic approximation.

Publication: SIAM Conference on Mathematics of Planet Earth, (MPE18) HAL

7.7. Defining coherent vortices from particles trajectories

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

Tracer patterns in the ocean, such as sea surface temperature, chlorophyll concentration and salinity suggest the emergence of coherence even in the ocean, typically with fluxes dominated by advective transport over diffusion. Mesoscale eddies are known to govern advective transport in the ocean, with typical horizontal scales $\mathcal{O}(100km)$ and timescales of $\mathcal{O}(weeks)$. These oceanic structures are omnipresent in the ocean and usually exhibit different properties to their surroundings. They are known to stir and mix surroundings water masses as well as by their ability to trap and carry fluid properties in a coherent manner. As the effect of these mesoscale eddies on the global circulation is remarkable, we focused on studying and understanding the dynamic transport properties of these coherent oceanic structures. For this reason, we have developed a Lagrangian method to identify and extract Lagrangian coherent vortices. The method analyzes particles' trajectories to identify vortices' boundaries and their attractor centers. Indeed, it is based on a decomposition of particle trajectory into two parts: closed curves which give information about uniformly rotating flow, and one that describes the mean displacement. The former part yields an objective measure of material rotation. We define Lagrangian coherent vortex as closed material lines in which fluid parcels complete the same polar rotation. This turns out to be filled with outward-increasing closed contours of the Lagrangian Averaged Closed Curve Length.

Publications: 3 articles submitted.

7.8. Soft hydrogel particle packings

Participants: J. Bares , N. Brodu , H. Zheng , J. A. Dijksman .

We provide the raw data from several years worth of effort on index matching experiments on soft hydrogel particle packings exposed to various loading conditions. Many of the data sets have not before been used, described explicitly in scientific publications or even analyzed. We will present a general overview of methods used to perform the imaging experiments. We provide particle-level information extracted from the 3D images, including contact forces and particle shapes, along with complementary boundary force measurements. We also provide a final working version of the code used to extract these microscopic features.

Publication: *Releasing data from mechanical tests on three dimensional hydrogel sphere packings*, Granular Matter, HAL

7.9. Feature based reconstruction model for fluorescence microscopy image denoising

Participants: S. K. Maji , H. Yahia .

: the advent of Fluorescence Microscopy over the last few years have dramatically improved the problem of visualization and tracking of specific cellular objects for biological inference. But like any other imaging system, fluorescence microscopy has its own limitations. The resultant images suffer from the effect of noise due to both signal dependent and signal independent factors, thereby limiting the possibility of biological inferencing. Denoising is a class of image processing algorithms that aim to remove noise from acquired images and has gained wide attention in the field of fluorescence microscopy image restoration. In this paper, we propose an image denoising algorithm based on the concept of feature extraction through multifractal decomposition and then estimate a noise free image from the gradients restricted to these features. Experimental results over simulated and real fluorescence microscopy data prove the merit of the proposed approach, both visually and quantitatively.

Publication: Scientific Reports, Nature Publishing Group, HAL

7.10. InnovationLab with I2S, sparse signals & optimisation

Participants: M. Martin, H. Yahia, A. Zebadua, S. Sakka, N. Brodu, K. Daoudi, A. Cherif [I2S], J. L. Vallancogne [I2S], A. Cailly [I2S], A. Billy [I2S], B. Lebouill [I2S].

During 2019:

Linear inverse problems in image processing, denoising, deconvolution, non-convex optimization, plug, and play algorithms.

HDR (High Dynamic Range) : C++ implementation with OpenCV library; code deposited on gitlab, sharing with I2S company, enhancement of the code architecture. GPU implementation.

Reflecting improvement: development and C++ implementation.

3D reconstruction: implementation and test with openMVG and openMVS libraries. Depth map enhancement through non-convex optimization.

Image smoothing implementation [57].

CPU and GPU patchmatch implementation [55].

MEMPHIS Project-Team

7. New Results

7.1. DGDD Method for Reduced-Order Modeling of Conservation Laws

Reduced-order models are attractive method to decrease significantly the computational cost of the simulations. However, the ability of reduced-order models to accurately approximate solutions containing strong convection, sharp gradients or discontinuities can be challenging. The discontinuous Galerkin domain decomposition (DGDD) reduced model for systems of conservation laws couples at the discrete level sub-domains of high-fidelity polynomial approximation to regions of low-dimensional resolution as shown in figure 8 .

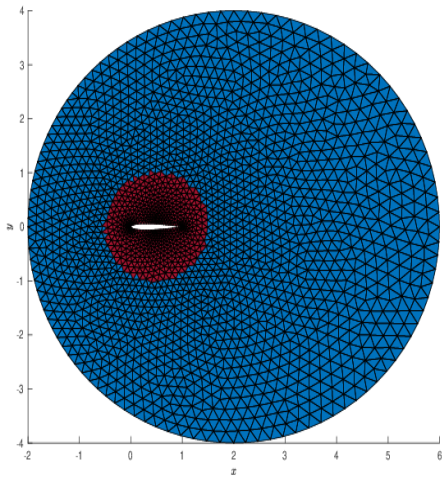


Figure 8. Decomposition of the domain: high-dimensional model (red), reduced-order model (blue).

In this approach, the high-dimensional model solves the equations where a given degree of accuracy is required, while the reduced-order model approximates the solution elsewhere. Since the high-dimensional model is used in a small part of the domain, the computational cost is significantly reduced. To perform the coupling, we develop a reduced-order model based on Proper Orthogonal Decomposition in the offline stage and on discontinuous Galerkin method in the online stage instead of the standard Galerkin method. In this way, the domain decomposition is applied transparently thought the numerical fluxes. We investigate the prediction of unsteady flows over a NACA 0012 airfoil. The results demonstrate the accuracy of the proposed method and the significant reduction of the computational cost. In the figures (9 and 10) we show examples of predictions obtained by the low-order model compared to the actual solutions as a function of the Mach number at infintiy and the angle of attack. In the last figure we present the overall space-time L^2 errors as a function of the low-dimensional space size.

7.2. Segmentation of aortic aneurism: collaboration with Nurea

Starting from February 2019, AMIES granted a one-year contract engineer in close collaboration with the Nurea start-up.

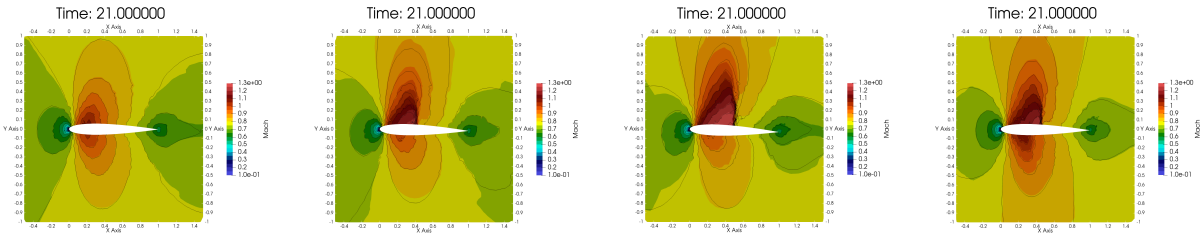


Figure 9. From left to right: $M_\infty = 0.754$ *and* $\alpha = 0.2$, $M_\infty = 0.763$ *and* $\alpha = 1.1$, $M_\infty = 0.776$ *and* $\alpha = 1.8$, $M_\infty = 0.784$ *and* $\alpha = 0.6$

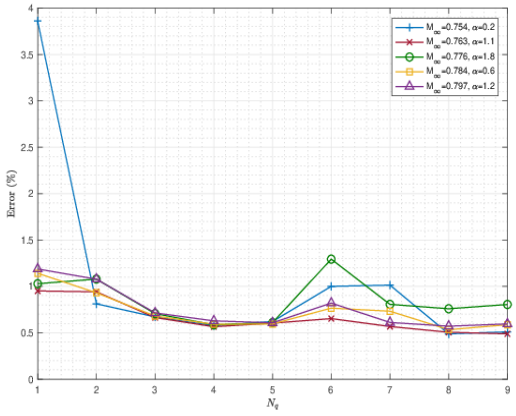


Figure 10. With $N_q = 9$ *basis functions,* ***the approximation error is less than 1% and the run time is reduced by approximately 72%*** *with respect to the high-fidelity solutions.*

The main objective of the project is to improve the quality and the robustness of automatic segmentation of aortic aneurism. An important part of the work was to decide if patient data needed to be pre-processed or not. To do so a criteron was developed to apply or not a smoothing filter. Several filters were tested and their performance were compared in order to choose the filter the more appropriate to our problem.
Another issue was the bones wrongly taken into the segmentation. A cleaning function was created to deal with it and remove the bones from the segmentation. An usual issue when working with medical images is to deal with the gradient of intensity. Existing tools need to be adapted to take into account these variations within images. This is currently worked out.
The code is implemented in C++ and mostly relies on itk and vtk libraries. An example in figure 11 .

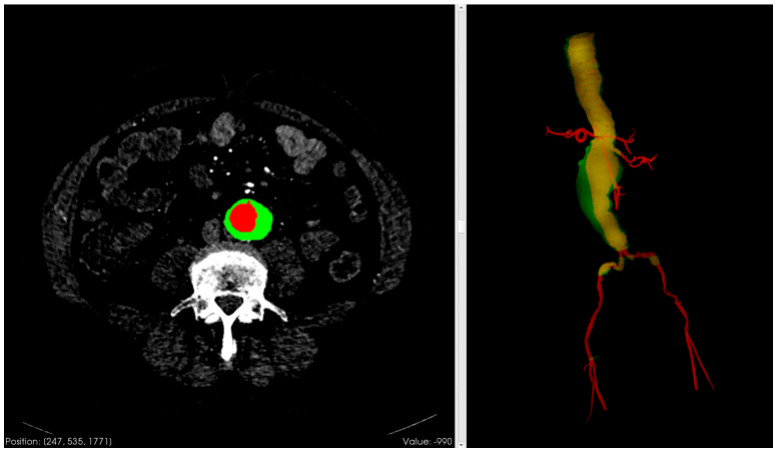


Figure 11. Example of aortic aneurism segmentation. Left: CT scan image; Right: 3D visualization. Red: blood; green: aortic wall

7.3. Fluid-structure interactions on AMR enabled quadree grids

We develop a versatile fully Eulerian method for the simulation of fluid-structure interactions. In the context of a monolithic approach, the whole system is modeled through a single continuum model. The equations are numerically solved using a finite-volume scheme with a compact stencil on AMR enabled quadtree grids where the dynamic refinement is adapted in time to the fluid-structure system.

The geometry is followed using a level-set formulation. In the Eulerian representation, a smooth Heaviside function is defined according to a level-set function on the cartesian mesh to distinguish between fluid and elastic phases. The temporal deformation of the structure is described according to the backward characteristics which are employed to express the Cauchy stress of a two-parameter hyperelastic Mooney-Rivlin material. This model is particularly adapted to elastomeric materials undergoing large deformations, see figure 12 .

7.4. Overset grids

One of the difficulties in the simulation of a fluid flow problem is the representation of the computational domain with a static mesh. As a matter of fact, not only the geometry could be particularly complex in itself, but it could change during the simulation and this necessary involves an *in itinere* geometrical adaptation of the mesh, with a consequent high computational cost. One of the ways to overcome this problem is to use multiple overlapping mesh blocks that together define a *Chimera* or *overset* grid. Once the different mesh blocks are generated, they are properly composed by the creation of holes (*hole cutting*) and, consequently,

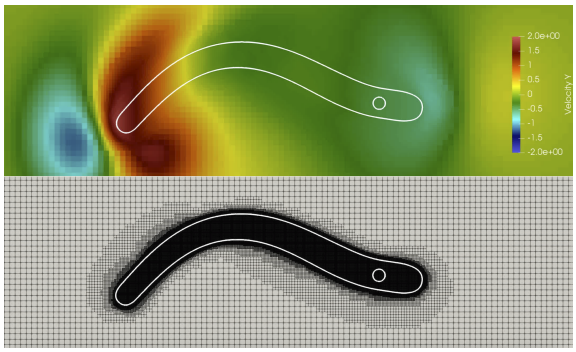


Figure 12. Y-component of the velocity for an oscillating elastomeric membrane actuated by a rigid holder at the tip, immersed in glycerin after 3 periods of oscillations. The criterion used for the dynamic AMR mesh is based on the level-set function.

an *overlapping zone* between two overlapping blocks is defined. Figure 13 shows a Chimera grid in the computational domain $[-\pi, \pi]^2$; in black there is the background mesh, in blue the foreground mesh. In particular, the foreground mesh can move and deform (consequently, the hole in the background mesh can change its configuration). The overlapping zone is necessary for the the communication and data transfer from one mesh to another. These operations are possible through an appropriate definition of local stencils of cells, both within and at the border of the individual blocks.
The Navier-Stokes equations for incompressible flows are going to be approximated through a projection method (*Chorin-Temam*), for this reason we have studied two Finite Volume (FV) solvers, for the pure diffusive equation and the unsteady convective-diffusive equation, on Chimera configurations. The first numerical experiments were conducted on 2D problems. The order of convergence of the error of the mismatch between the exact solution and its FV approximation in L^∞- and L^2-norms is 2.

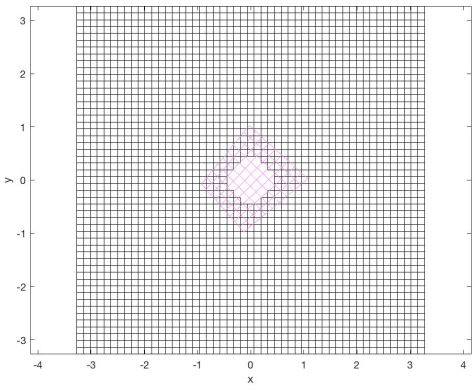


Figure 13. A Chimera grid configuration for the computational domain $[-\pi, \pi]^2$.

7.5. Collective propulsion: collaboration with ONERA

Motivated by recent studies and the locomotion of animal groups for robotics, we investigated the influence of hydrodynamic interactions on the collective propulsion of flapping wings. We studied the horizontal locomotion of an infinite array of flapping wings separated by a constant gap using unsteady non-linear simulations. Two control parameters were explored: the flapping frequency and the gap between the wings. Results obtained for different gaps at a fixed frequency are shown in Figure 14 . We first observe that for a very large spacing between the wings — greater than 20 times the chord of a wing — the interaction effects are no longer present (Figure 14 (b)) and the average speed of the system tends to the speed of a single wing. For lower gaps, the average speed may become lower or higher than that of a single wing. For certain gaps, one can find two different stable solutions: one at higher propulsion speed and the other lower than a single wing. This phenomenon has already been observed for a fixed spacing and different frequencies of movement. The stability of these solutions is linked to the interaction between the vortex wake generated by the previous wings and the vortex ejected at the leading or trailing edge of the considered wing (Figure 14 (a)). We remark that propulsive efficiency is higher for the collective case both in faster and slower solutions. Understanding the key mechanisms responsible for the stable solutions will provide directions to control strategies aiming to optimize the wing horizontal speed.

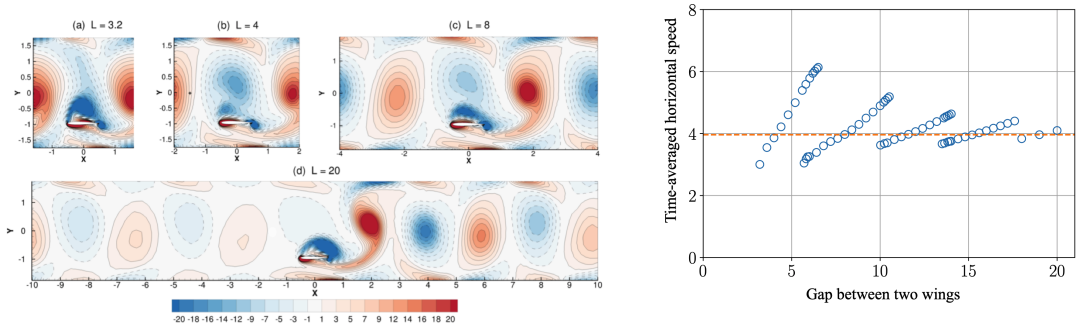


Figure 14. Vorticity contours (a) and time-averaged horizontal speed (b) for a flapping wing interacting in an infinite array. In (b) the average speed of a non interacting flapping wing is represented by a dashed orange line.

7.6. Automatic registration for model reduction

As part of the ongoing team effort on ROMs, we work on the development of automatic registration procedures for model reduction. In computer vision and pattern recognition, registration refers to the process of finding a spatial transformation that aligns two datasets; in our work, registration refers to the process of finding a parametric transformation that improves the linear compressibility of a given parametric manifold. For advection-dominated problems, registration is motivated by the inadequacy of linear approximation spaces due to the presence of parameter-dependent boundary layers and travelling waves.

In [48], we proposed and analysed a computational procedure for stationary PDEs and investigated performance for two-dimensional model problems. In Figure 15 , we show slices of the parametric solution for three different parameters before (cf. Left) and after (cf. Right) registration: we observe that the registration procedure is able to dramatically reduce the sensitivity of the solution to the parameter value μ. In Figure 16 , we show the behaviour of the normalised POD eigenvalues (cf. Left) and of the relative L^2 error of the corresponding POD-Galerkin ROM (cf. Right) for an advection-reaction problem: also in this case, the approach is able to improve the approximation properties of linear approximation spaces and ultimately simplify the reduction task.

We aim to extend the approach to a broad class of non-linear steady and unsteady PDEs: in October 2019, we funded a 16-month postdoc to work on the reduction of hyperbolic systems of PDEs. We are also collaborating with EDF (departments PERICLES and LNHE) to extend the approach to the Saint-Venant (shallow water) equations.

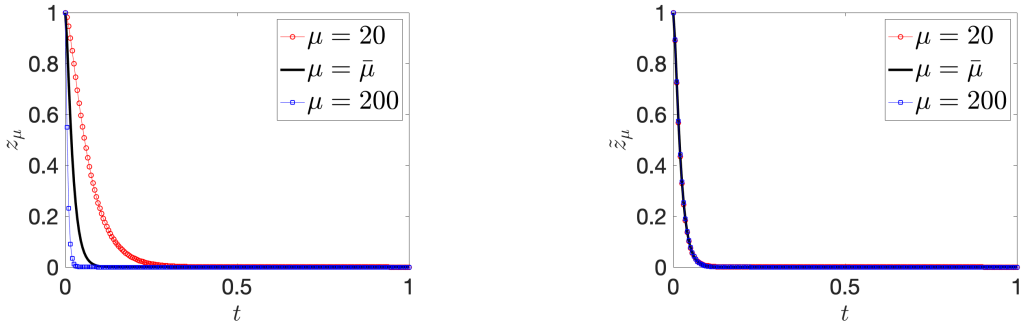


Figure 15. automatic registration for model reduction. Registration of a boundary layer ($\bar{\mu} = \sqrt{20 \cdot 200}$).

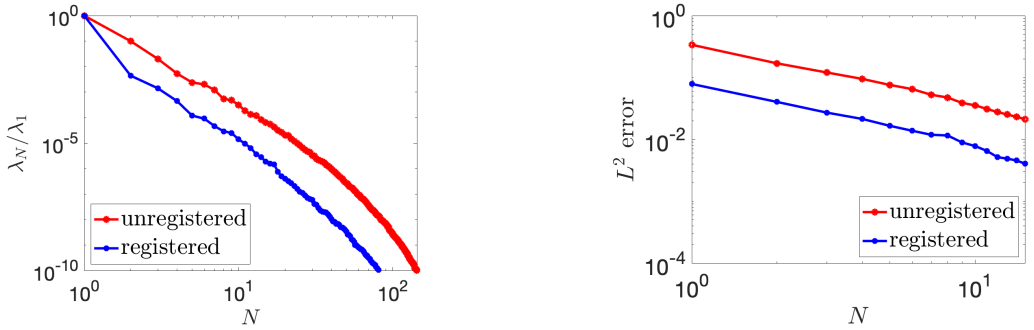


Figure 16. automatic registration for model reduction. Registration of a parameterized advection-reaction problem.

7.7. Modeling and numerical simulation of ellipsoidal particle-laden flows and self propelled swimmers in a porous enclosure

Despite being relevant in many natural and industrial processes, suspensions of non-spherical particles have been largely under-investigated compared to the extensive analyses made on the gravity-driven motions of spherical particles. One of the main reasons for this disparity is the difficulty of accurately correcting the short-range hydrodynamic forces and torques acting on complex particles. These effects, also known as lubrication, are essential to the suspension of the particles and are usually poorly captured by direct numerical simulation of particle-laden flows. We have proposed a partitioned VP-DEM (Volume Penalization method - Discrete Element Method) solver which estimates the unresolved hydrodynamic forces and torques. Corrections are made locally on the surface of the interacting particles without any assumption on the particle global geometry.

This is an extension of our previous work [39]. Numerical validations have been made using ellipsoidal particles immersed in an incompressible Navier-Stokes flow.

Self organization of groups of several swimmers is of interest in biological applications. One of the main question is to determine if the possible organization comes from an uncontrolled or a controlled swimming behavior. This work has been motivated by the recent studies of Hamid Kellay (LOMA). Hamid Kellay has presented his results during a small workshop we organized earlier this year between members of the Memphis team (modeling and numerical methods), the MRGM (experimental zebra-fishes swimming), the LOMA (interaction between self-propelled particles) and the ONERA (flapping wings).

The collision model developed in the previous section has been developed for concave interactions like sphere-sphere or in the limit of sphere-plane wall. For non concave interactions, we have derived a simple approximation considering locally convexity as being a plane wall. We have performed numerical simulations of the interaction of several self propelled swimmers in a porous enclosure (see figure 17). Fishes are organized in small groups that are able to put into motion the enclosure. This behavior is similar to the one observed in the experimental set-up of Hamid Kellay.

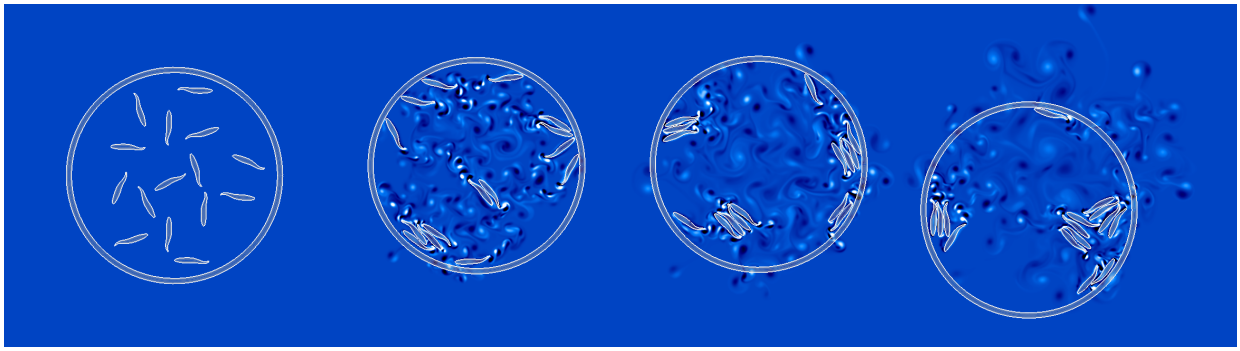


Figure 17. Organization of small self propelled fish groups in a porous enclosure (time evolution is in the usual reading direction). Colormap is the vorticity field.

REALOPT Project-Team

7. New Results

7.1. An iterative dynamic programming approach for the temporal knapsack problem

We have developed an approach to solve the temporal knapsack problem (TKP) based on a very large size dynamic programming formulation [28]. In this generalization of the classical knapsack problem, selected items enter and leave the knapsack at fixed dates. We solve the TKP with a dynamic program of exponential size, which is solved using a method called Successive Sublimation Dynamic Programming (SSDP). This method starts by relaxing a set of constraints from the initial problem, and iteratively reintroduces them when needed. We show that a direct application of SSDP to the temporal knapsack problem does not lead to an effective method, and that several improvements are needed to compete with the best results from the literature.

7.2. A Hypergraph Model for the Rolling Stock Rotation Planning and Train Selection Problem

We investigated an integrated optimization approach for timetabling and rolling stock rotation planning in the context of passenger railway traffic [27]. Given a set of possible passenger trips, service requirement constraints, and a fleet of multiple heterogeneous self-powered railcars, our method aims at producing a timetable and solving the rolling stock problem in such a way that the use of railcars and the operational costs are minimized. To solve this hard optimization problem, we design a mixed-integer linear programming model based on network-flow in an hypergraph. We use this models to handle effectively constraints related to coupling and decoupling railcars. To reduce the size of the model, we use an aggregation and disaggregation technique combined with reduced-cost filtering. Computational experiments based on several French regional railway traffic case studies show that our method scales successfully to real-life problems.

7.3. Decomposition-based approaches for a class of two-stage robust binary optimization problems

We have studied a class of two-stage robust binary optimization problems with objective uncertainty where recourse decisions are restricted to be mixed-binary [25]. For these problems, we present a deterministic equivalent formulation through the convexification of the recourse feasible region. We then explore this formulation under the lens of a relaxation, showing that the specific relaxation we propose can be solved using the branch-and-price algorithm. We present conditions under which this relaxation is exact, and describe alternative exact solution methods when this is not the case. Despite the two-stage nature of the problem, we provide NP-completeness results based on our reformulations. Finally, we present various applications in which the methodology we propose can be applied. We compare our exact methodology to those approximate methods recently proposed in the literature under the name K-adaptability. Our computational results show that our methodology is able to produce better solutions in less computational time compared to the K-adaptability approach, as well as to solve bigger instances than those previously managed in the literature.

7.4. Primal heuristic for Branch-and-Price

Primal heuristics have become an essential component in mixed integer programming (MIP) solvers. Extending MIP based heuristics, our study outlines generic procedures to build primal solutions in the context of a branch-and-price approach and reports on their performance. Our heuristic decisions carry on variables of the Dantzig-Wolfe reformulation, the motivation being to take advantage of a tighter linear programming relaxation than that of the original compact formulation and to benefit from the combinatorial structure embedded

in these variables. In [9], we focus on the so-called diving methods that use re-optimization after each LP rounding. We explore combinations with diversification-intensification paradigms such as limited discrepancy search, sub-MIPing, local branching, and strong branching. The dynamic generation of variables inherent to a column generation approach requires specific adaptation of heuristic paradigms. We manage to use simple strategies to get around these technical issues. Our numerical results on generalized assignment, cutting stock, and vertex coloring problems sets new benchmarks, highlighting the performance of diving heuristics as generic procedures in a column generation context and producing better solutions than state-of-the-art specialized heuristics in some cases.

7.5. Generic solver for vehicle routing problems

Major advances were recently obtained in the exact solution of Vehicle Routing Problems (VRPs). Sophisticated Branch-Cut-and-Price (BCP) algorithms for some of the most classical VRP variants now solve many instances with up to a few hundreds of customers. However, adapting and re-implementing those successful algorithms for other variants can be a very demanding task. In [8], [12], [11], we propose a BCP solver for a generic model that encompasses a wide class of VRPs. It incorporates the key elements found in the best existing VRP algorithms: ng-path relaxation, rank-1 cuts with limited memory, path enumeration, and rounded capacity cuts; all generalized through the new concepts of "packing set" and "elementarity set". The concepts are also used to derive a branching rule based on accumulated resource consumption and to generalize the Ryan and Foster branching rule. Extensive experiments on several variants show that the generic solver has an excellent overall performance, in many problems being better than the best specific algorithms. Even some non-VRPs, like bin packing, vector packing and generalized assignment, can be modeled and effectively solved.

The solver is available for download and free academic use in http://vrpsolver.math.u-bordeaux.fr.

7.6. New algorithms for different vehicle routing problems

In [7], we propose a branch-cut-and-price algorithm for the two-echelon capacitated vehicle routing problem in which delivery of products from a depot to customers is performed using intermediate depots called satellites. We introduce a new route based formulation for the problem which does not use variables to determine product flows in satellites. Second, we introduce a new branching strategy which significantly decreases the size of the branch-and-bound tree. Third, we introduce a new family of satellite supply inequalities, and we empirically show that it improves the quality of the dual bound at the root node of the branch-and-bound tree. Finally, extensive numerical experiments reveal that our algorithm can solve to optimality all literature instances with up to 200 customers and 10 satellites for the first time and thus double the size of instances which could be solved to optimality.

In [22], [21], we are interested in the exact solution of the vehicle routing problem with backhauls (VRPB), a classical vehicle routing variant with two types of customers: linehaul (delivery) and backhaul (pickup) ones. We propose two branch-cut-and-price (BCP) algorithms for the VRPB. The first of them follows the traditional approach with one pricing subproblem, whereas the second one exploits the linehaul/backhaul customer partitioning and defines two pricing subproblems. The methods incorporate elements of state-of-the-art BCP algorithms, such as rounded capacity cuts, limited-memory rank-1 cuts, strong branching, route enumeration, arc elimination using reduced costs and dual stabilization. Computational experiments show that the proposed algorithms are capable of obtaining optimal solutions for all existing benchmark instances with up to 200 customers, many of them for the first time. It is observed that the approach involving two pricing subproblems is more efficient computationally than the traditional one. Moreover, new instances are also proposed for which we provide tight bounds. Also, we provide results for benchmark instances of the heterogeneous fixed fleet VRPB and the VRPB with time windows.

In [20], we consider the standard Capacitated Location-Routing Problem (LRP), which is the combination of two canonical combinatorial optimization problems : Facility Location Problem (FLP), and Vehicle Routing Problem (VRP). We have extended the Branch-and-Cut-and-Price Algorithm from [11] to solve a Mixed

Integer Programming (MIP) formulation with and exponential number of varialbes. A new family of Route Load Knapsack valid inequalites is proposed to strengthen the formulation. Preliminary results showed that our algorithm could solve to optimality, for the first time, 12 open instances of the most difficult classes of LRP instances.

In the first echelon of the two-echelon stochastic multi-period capacitated location-routing problem (2E-SM-CLRP), one has to decide the number and location of warehouse platforms as well as the intermediate distribution platforms for each period; while fixing the capacity of the links between them. The system must be dimensioned to enable an efficient distribution of goods to customers under a stochastic and time-varying demand. In the second echelon of the 2E-SM-CLRP, the goal is to construct vehicle routes that visit customers from operating distribution platforms. The objective is to minimize the total expected cost. We model this hierarchical decision problem as a two-stage stochastic program with integer recourse. The first-stage includes location and capacity decisions to be fixed at each period over the planning horizon, while routing decisions of the second echelon are determined in the recourse problem. In [16], [26], we propose a Benders decomposition approach to solve this model. In the proposed approach, the location and capacity decisions are taken by solving the Benders master problem. After these first-stage decisions are fixed, the resulting subproblem is a capacitated vehicle-routing problem with capacitated multi-depot (CVRP-CMD) that is solved by a branch-cut-and-price algorithm. Computational experiments show that instances of realistic size can be solved optimally within reasonable time, and that relevant managerial insights are derived on the behavior of the design decisions under the stochastic multi-period characterization of the planning horizon.

Much of the existing research on electric vehicle routing problems (E-VRPs) assumes that the charging stations (CSs) can simultaneously charge an unlimited number of electric vehicles, but this is not the case. In [29], we investigate how to model and solve E-VRPs taking into account these capacity restrictions. In particular, we study an E-VRP with non-linear charging functions, multiple charging technologies, en route charging, and variable charging quantities, while explicitly accounting for the capacity of CSs expressed in the number of chargers. We refer to this problem as the E-VRP with non-linear charging functions and capacitated stations (E-VRP-NL-C). This problem advances the E-VRP literature by considering the scheduling of charging operations at each CS. We first introduce two mixed integer linear programming formulations showing how CS capacity constraints can be incorporated into E-VRP models. We then introduce an algorithmic framework to the E-VRP-NL-C, that iterates between two main components: a route generator and a solution assembler. The route generator uses an iterated local search algorithm to build a pool of high-quality routes. The solution assembler applies a branch-and-cut algorithm to select a subset of routes from the pool. We report on computational experiments comparing four different assembly strategies on a large and diverse set of instances. Our results show that our algorithm deals with the CS capacity constraints effectively. Furthermore, considering the well-known uncapacitated version of the E-VRP-NL-C, our solution method identifies new best-known solutions for 80 out of 120 instances.

7.7. Packing problems

In the two-dimensional guillotine cutting-stock problem, the objective is to minimize the number of large plates used to cut a list of small rectangles. We consider a variant of this problem, which arises in glass industry when different bills of order (or batches) are considered consecutively. For practical organisation reasons, leftovers are not reused, except the large one obtained in the last cutting pattern of a batch, which can be reused for the next batch. The problem can be decomposed into an independent problem for each batch. In [6] we focus on the one-batch problem, the objective of which is to minimize the total width of the cutting patterns used. We propose a diving heuristic based on column generation, in which the pricing problem is solved using dynamic programming (DP). This DP generates so-called non-proper columns, i.e. cutting patterns that cannot participate in a feasible integer solution of the problem. We show how to adapt the standard diving heuristic to this “non-proper” case while keeping its effectiveness. We also introduce the partial enumeration technique, which is designed to reduce the number of non-proper patterns in the solution space of the dynamic program. This technique strengthens the lower bounds obtained by column generation and improves the quality of the solutions found by the diving heuristic. Computational results are reported

and compared on classical benchmarks from the literature as well as on new instances inspired from glass industry data. According to these results, variants of the proposed diving heuristic outperform constructive and evolutionary heuristics.instances than those previously managed in the literature.

The bin packing problem with generalized time lags (BPGL) consists of a set of items, each having a positive weight, and a set of precedence constraints with lags between pairs of items, allowing negative and non-negative lags. The items must be packed into the minimum possible number of bins with identical capacity, and the bins must be assigned to time periods satisfying the precedence constraints with lags on the items. In [18] we show a solution strategy using a generic branch-and-price algorithm (implemented in the software platform BaPCod) and applying some problem specific cuts. Our approach outperformed the compact Mixed Integer Programming (MIP) formulation solved by the MIP solver Cplex.

7.8. Covering problems

In [5], we study a covering problem where vertices of a graph have to be covered by rooted sub-trees. We present three mixed-integer linear programming models, two of which are compact while the other is based on Dantzig-Wolfe decomposition. In the latter case, we focus on the column generation sub-problem, a knapsack problem with precedence constraints on trees, for which we propose several algorithms among them a branch and bound algorithm and various dynamic programs. Numerical results are obtained using instances from the literature and instances based on a real-life districting application. Experiments show that the branch-and-price algorithm is able to solve much bigger instances than the compact model, which is limited to very small instance sizes.

7.9. Software Defined Networking

ISP networks are taking a leap forward thanks to emerging technologies such as Software Defined Networking (SDN) and Network Function Virtualization (NFV). Efficient algorithms considered too hard to be put in practice on legacy networks now have a second chance to be considered again. In this context, we rethink the ISP network dimensioning problem with protection against Shared Risk Link Group (SLRG) failures. In [30], [15], [23], we consider a path-based protection scheme with a global rerouting strategy, in which, for each failure situation, we may have a new routing of all the demands. Our optimization task is to minimize the needed amount of bandwidth. After discussing the hardness of the problem, we develop two scalable mathematical models that we handle using both Column Generation and Benders Decomposition techniques. Through extensive simulations on real-world IP network topologies and on random generated instances, we show the effectiveness of our methods. Finally, our implementation in OpenDaylight demonstrates the feasibility of the approach and its evaluation with Mininet shows that technical implementation choices may have a dramatic impact on the time needed to reestablish the flows after a failure takes place.

CARMEN Project-Team

7. New Results

7.1. Modelling, direct simulation and prediction of cardiac phenomena

Using high-performance simulations on a detailed model of the human atria [58] we investigated several aspects of atrial fibrillation (AF). We showed that AF initiation by rapid pacing is sensitive to very small changes in parameter values [70] [32], [38] , and investigated effects of antiarrhythmic drugs and interventions [59] [34] and pathologies [35]. An example movie is available online at https://www.potse.nl/papers/potse/potse-fimh19.html.

High-performance simulations of human ventricular activity have contributed to the testing of new electrocardiographic mapping methods ("inverse models") [73], [36].

We are also developing new methods to help with the treatment of these patients by rapidly guiding an ablation catheter to the origin (strictly speaking: the exit site) of an arrhythmia. These methods are also being tested with simulated data. [33].

We contributed to work by Prof. Michel Haissaguerre and his team in which it is argued that many patients who are now believed to suffer from abnormalities in the genes for specific cardiac ion channels are in reality affected by structural diseases of the heart muscle [62] [26], [41]. These influential publications represent an important change in thinking about these patients and their treatment.

Simple mitochondrial model based on thermodynamic fluxes (PhD work of B. Tarraf): Mitochondria are involved in the regulation of calcium which plays a crucial role in the propagation of cardiac action potentials. However, they are not taken into account in the ionic models that are used to perform simulations at the tissue level. In the framework of the ANR MITOCARD project, we wrote a simple model of mitochondrial calcium regulation based on an extensive review of models of the litterature, which are not suited for further calibration due to their excessive complexity [39]. Now that the equations are written down, we are performing a parameter analysis on the whole model before including other key biological mechanisms.

In a collaboration with Jeremy Darde (IMT Toulouse), we have developed a numerical method on a cartesian grid to solve the direct problem of Electrical Impedance Tomography (EIT) in complex geometries, with first-order convergence. The objective is to solve then the inverse problem of EIT to identify heterogeneities of conductivities on the torso volume.

7.2. Inverse problems: parameter estimation, data assimilation and ECGi

- Data assimilation: In A. Gérard PhD thesis, it is showed that accounting for the anisotropy in the atria is crucial to reconstruct correctly activation maps compatible with a mono-domain model from sparse punctuals activation times. To this purpose, we have developed a new data assimilation method for the mono-domain model, using a Luenberger filter and a Kalman-type filter (ROUKF), based on the dissimilarity measure introduced in A. Collin PhD thesis.
- Parameter estimation: We have been working on the following theoretical question: What are the condition under which the parameters, in the mathematical codomain and bidomain models, are identifiables. Then we proposed an algorithm capable of estimating different ion-channels conductance parameters. mettre ref?
- ECGI: Several approaches have been investigated to improve the resolution of ECGi.
  - development of a new algorithm to choose the regularization coefficient for the resolution of ECGi with the Method of Fondamental Solutions (MFS).
  - a study using a parametrized model of action potentials, showing that accounting for the endocardium can improve the resolution of ECGi.

- in joint work with Laura Bear (IHU Liryc), development of a new method for improving the resolution of ECGI by combining several solutions obtained with various numerical methods (FEM and MFS). The method is based on the selection of the smallest residuals on the torso surface.
- Development of new methods for the ECGI problem, based on machine learning methods. The idea is to learn activation maps from body surface signals.
- collaborative work on the set up of an experimental platform for the experimental non-invasive validation of the reconstruction of cardiac signals

7.3. Numerical schemes

- Very-high order Finite Volume methods: We have showed the very good behavior of a specifically devised domain decomposition technique: the communications are minimized without impacting the accuracy or the order of convergence of the scheme. The total amount of communications dos not increase significantly between the second and the 6th order. The 6th-order Finite-Volume scheme is thus the most performing scheme.
- Numerical analysis of a cartesian method for elliptic problems with immersed interfaces: We have studied the convergence of a cartesian method for elliptic problems with immersed interfaces previously published [54]. The convergence is proved for the original second-order method in one-dimension and for a first-order version in two dimensions. The proof uses a discrete maximum principle to obtain estimates of the coefficients of the inverse matrix.

7.4. Miscellanous

- Refactoring of the CEPS software. Our CEPS software underwent an important refactoring phase so that future students spend less time getting hands on the code. Also, several features developped by previous PhD students were merged into the version that we intend to distribute : high order numerical schemes suited for ionic models, volume fraction due to tissue heterogeneities. We are currently in the process of licensing with Inria and University of Bordeaux.
- Deep brain stimulation procedure: We have been working on three different learning methods for the prediction of the optimal pacing sites in the Deep brain stimulation procedure. We also compared the found position to the position of the stimulation sites to the anatomical geometries in the Ewert brain atlas. Results show the robustness of the methods in founding the stimulation regions.
- Generation of boundary conditions for the Boussinesq system: In a collaboration with David Lannes (IMB, Bordeaux) we have developed a new method for the numerical implementation of generating boundary conditions for a one dimensional Boussinesq system. The method is based on a reformulation of the equations and a resolution of the dispersive boundary layer that is created at the boundary.
- Simulation of solid suspensions in incompressible fluids: In a collaboration with B. Lambert and M. Bergmann (IMB, Bordeaux) we have extended a previously published local lubrication model to non-Brownian suspensions of ellipsoidal solid particles in incompressible flows. This lubrication model used virtual spheres to evaluate local lubrication corrections instead of the global corrections found in the classical lubrication theory.

MAGIQUE-3D Project-Team

7. New Results

7.1. High-order numerical methods for modeling wave propagation in complex media: development and implementation

7.1.1. High order discretization of seismic waves-problems based upon DG-SE methods

Participants: Hélène Barucq, Julien Diaz, Aurélien Citrain.

Accurate wave propagation simulations require selecting numerical schemes capable of taking features of the medium into account. In case of complex topography, unstructured meshes are the most adapted and in that case, Discontinuous Galerkin Methods (DGm) have demonstrated great performance. Off-shore exploration involves propagation media which can be well represented by hybrid meshes combining unstructured meshes with structured grids that are best for representing homogeneous media like water layers. Then it has been shown that Spectral Element Methods (SEm) deliver very accurate simulations on structured grids with much lower computational costs than DGms.

We have developed a DG-SEm numerical method for solving time-dependent elasto- acoustic wave problems. We consider the first-order coupled formulation for which we propose a DG-SEm formulation which turns out to be stable.

While the 2D case is almost direct, the 3D case requires a particular attention on the coupling boundary on which it is necessary to manage the possible positions of the faces of the tetrahedrons with respect to that of the neighboring hexaedra.

In the framework of this DG-SEm coupling, we are also interested in the Perfectly Matched Layer (PML) in particular the use of the SEm inside it to stabilize it in cases where the use of DGm leads to instabilities.

These results have been obtained in collaboration with Henri Calandra (TOTAL) and Christian Gout (INSA Rouen) and have been presented at Journées Ondes Sud-Ouest (JOSO) in Le Barp, the 14th International Conference on Mathematical and Numerical Aspects of Wave Propagation (WAVES) in Vienna (Austria) and MATHIAS conference in Paris [21], [26]

7.1.2. Isogeometric analysis of sharp boundaries in full waveform inversion

Participants: Hélène Barucq, Julien Diaz, Stefano Frambati.

Efficient seismic full-waveform inversion simultaneously demands a high efficiency per degree of freedom in the solution of the PDEs, and the accurate reproduction of the geometry of sharp contrasts and boundaries. Moreover, it has been shown that the stability constant of the FWI minimization grows exponentially with the number of unknowns. Isogeometric analysis has been shown to possess a higher efficiency per degree of freedom, a better convergence in high energy modes (Helmholtz) and an improved CFL condition in explicit-time wave propagation, and it seems therefore a good candidate for FWI.

In the first part of the year, we have focused on a small-scale one-dimensional problem, namely the inversion over a multi-step velocity model using the Helmholtz equation. By exploiting a relatively little-known connection between B-splines ad Dirichlet averages, we have added the knot positions as degrees of freedom in the inversion. We have shown that arbitrarily-placed discontinuities in the velocity model can be recovered using a limited amount of degrees of freedom, as the knots can coalesce at arbitrary positions, obviating the need for a very fine mesh and thus improving the stability of the inversion.

In order to reproduce the same results in two and three dimensions, the usual tensor-product structure of B-splines cannot be used. We have therefore focused on the construction of (unstructured) multivariate B-spline bases. We have generalized a known B-spline basis construction through the language of oriented matroids, showing that multivariate spline bases can be easily constructed with repeated knots and that the construction algorithm can be extended to three dimensions. This gives the freedom to locally reduce the regularity of the basis functions and to place internal boundaries in the domain. The resulting mass matrix is block-dagonal, with adjustable block size, providing an avenue for a simple unstructured multi-patch DG-IGA scheme that is being investigated. With this goal in mind, more efficient quadrature schemes for multivariate B-splines exploiting the connection to oriented matroids are also being investigated.

A research report is in preparation.

7.1.3. Seismic wave propagation in carbonate rocks at the core scale

Participants: Julien Diaz, Florian Faucher, Chengyi Shen.

Reproduction of large-scale seismic exploration at lab-scale with controllable sources is a promising approach that could not only be applied to study small-scale physical properties of the medium, but also contribute to significant progress in wave-propagation understanding and complex media imaging at exploration scale via upscaling methods. We propose to apply a laser-generated seismic point source for core-scale new geophysical experiments. This consists in generating seismic waves in various media by well-calibrated pulsed-laser impacts and measuring precisely the wavefield (displacement) by Laser Doppler Vibrometer (LDV). The point-source-LDV configuration is convenient to model numerically. It can also favor the incertitude estimate of the source and receiver locations. Parallel 2D/3D simulations featuring the Discontinuous Galerkin discretization method with Interior Penalties (IPDG) are done to match the experimental data. The IPDG method is of particular interest when it comes to solve wave propagation problems in highly heterogeneous media, such as the limestone cores that we are studying.

Current seismic data allowed us to retrieve V_p tomography slices. Further more, qualitative/quantitative comparisons between simulations and experimental data validated the experiment protocol and vice-versa the high-order FEM schemes, opening the possibility of performing FWI on dense, high frequency and large band-width data.

This work is in collaboration with Clarisse Bordes, Daniel Brito, Federico Sanjuan and Deyuan Zhang (LFCR, UPPA) and with Stéphane Garambois (ISTerre). Ii is one of the topic of the PhD. thesis of Chengyi Shen.

7.1.4. Simulation of electro-seismic waves using advanced numerical methods

Participants: Hélène Barucq, Julien Diaz, Ha Howard Faucher, Rose-Cloé Meyer.

We study time-harmonic waves propagation in conducting poroelastic media, in order to obtain accurate images for complex media with high-order methods. In these kind of media, we observe the coupling between electromagnetic and seismic wave fields, which is called seismokinetic effect. The converted waves are very interesting because they are heavily sensitive to the medium properties, and the modeling of seismo-electric conversion can allow to detect interfaces in the material where the seismic field would be blind. To the best of our knowledge, the numerical simulation of this phenomenon has never been achieved with high-order finite element methods. Simulations are difficult to perform in time domain, because the time step and the mesh size have to be adapted to the huge variations of wave velocities. To ease the numerical implementation, we work in the frequency domain. We can then include physical parameters that depend non-linearly on the frequencies. Then, we have developed a new Hybridizable Discontinuous Galerkin method for discretizing the equations. This allows us to reduce the computational costs by considering only degrees of freedom on the skeleton of the mesh. We have validated the numerical method thanks to comparison with analytical solutions. We have obtained numerical results for 2D realistic poroelastic media and conducting poroelastic media is under investigation.

Results on analytical solutions for poroelasticity are presented in the research report [44].

7.1.5. Quasinormal mode expansion of electromagnetic Lorentz dispersive materials

Participants: Marc Duruflé, Alexandre Gras.

We have studied the electromagnetic scattering of optical waves by dispersive materials governed by a Drude-Lorentz model. The electromagnetic fields can be decomposed onto the eigenmodes of the system, known as quasinormal modes. In [51], a common formalism is proposed to obtain different formulas for the coefficients of the modal expansion. In this paper, it is also explained how to handle dispersive Perfectly Matched Layers and degenerate eigenvectors. Lately, we have investigated the use of an interpolation method in order to compute quickly the diffracted field for a large number of frequencies.

7.1.6. A Hybridizable Galerkin Discontinuous formulation for elasto-acoustic coupling in frequency-domain

Participants: Hélène Barucq, Julien Diaz, Vinduja Vasanthan.

We are surrounded by many solid-fluid interactions, such as the seabed or red blood cells. Indeed, the seabed represents the ocean floors immersed in water, and red blood cells are coreless hemoglobin-filled cells. Hence, when wanting to study the propagation of waves in such domains, we need to take into account the interactions at the solid-fluid interface. Therefore, we need to implement an elasto-acoustic coupling. Many methods have already tackled with the elasto-acoustic coupling, particularly the Discontinuous Galerkin method. However, this method needs a large amount of degrees of freedom, which increases the computational cost. It is to overcome this drawback that the Hybridizable Discontinuous Galerkin (HDG) has been introduced. The implementations of HDG for the elastic wave equations, as well as partially for the acoustic ones, have been done previously. Using these, we have performed in this work the elasto-acoustic coupling for the HDG methods in 1D, 2D and 3D. The results are presented in Vinduja Vasanthan's master's thesis [54].

7.1.7. Absorbing Radiation Conidition in elongated domains

Participants: Hélène Barucq, Sébastien Tordeux.

We develop and analyse a high-order outgoing radiation boundary condition for solving three-dimensional scattering problems by elongated obstacles. This Dirichlet-to-Neumann condition is constructed using the classical method of separation of variables that allows one to define the scattered field in a truncated domain. It reads as an infinite series that is truncated for numerical purposes. The radiation condition is implemented in a finite element framework represented by a large dense matrix. Fortunately, the dense matrix can be decomposed into a full block matrix that involves the degrees of freedom on the exterior boundary and a sparse finite element matrix. The inversion of the full block is avoided by using a Sherman–Morrison algorithm that reduces the memory usage drastically. Despite being of high order, this method has only a low memory cost. This work has been published in [13].

7.1.8. Discontinuous Galerkin Trefftz type method for solving the Maxwell equations

Participants: Margot Sirdey, Sébastien Tordeux.

Trefftz type methods have been developed in Magique 3D to solve Helmholtz equation and it has been presented in [25].These methods reduce the numerical dispersion and the condition number of the linear system. This work aims in pursuing this development for electromagnetic scattering. We have adapted and tested the method for an academical 2D configuration. This is the topic of the PhD thesis of Margot Sirdey.

7.1.9. Reduced models for multiple scattering of electromagnetic waves

Participants: Justine Labat, Victor Péron, Sébastien Tordeux.

In this project, we develop fast, accurate and efficient numerical methods for solving the time-harmonic scattering problem of electromagnetic waves by a multitude of obstacles for low and medium frequencies in 3D. First, we consider a multi-scale diffraction problem in low-frequency regimes in which the characteristic length of the obstacles is small compared to the incident wavelength. We use the matched asymptotic expansion method which allows for the model reduction. Then, small obstacles are no longer considered as geometric constraints and can be modelled by equivalent point-sources which are interpreted in terms of electromagnetic multipoles. Second, we justify the Generalized Multiparticle Mie-solution method (Xu, 1995) in the framework of spherical obstacles at medium-frequencies as a spectral boundary element method based on the Galerkin discretization of a boundary integral equation into local basis composed of the vector spherical harmonics translated at the center of each obstacle. Numerically, a clever algorithm is implemented in the context of periodic structures allowing to avoid the global assembling of the matrix and so, reduce memory usage. The reduced asymptotic models of the first problem can be adapted for this regime by incorporating non-trivial corrections appearing in the Mie theory. Consequently, a change in variable between the two formulations can be made explicit, and an inherent advantage of the asymptotic formulation is that the basis and the shape can be separated with a semi-analytical expression of the polarizability tensors. A comparison of these different methods in terms of their accuracy has been carried out. Finally, for both methods and in the context of large numbers of obstacles, we implement an iterative resolution with preconditionning in a GMRES framework.

These results have been presented at Journées Ondes Sud-Ouest (JOSO) in Le Barp (France) and the 14th International Conference on Mathematical and Numerical Aspects of Wave Propagation (WAVES) in Vienna (Austria), see [22], [39]. Part of this work has been published in Wave Motion [17].

7.1.10. Boundary Element Method for 3D Conductive Thin Layer in Eddy Current Problems

Participant: Victor Péron.

Thin conducting sheets are used in many electric and electronic devices. Solving numerically the eddy current problems in presence of these thin conductive sheets requires a very fine mesh which leads to a large system of equations, and becoming more problematic in case of high frequencies. In this work we show the numerical pertinence of asymptotic models for 3D eddy current problems with a conductive thin layer of small thickness based on the replacement of the thin layer by its mid-surface with impedance transmission conditions that satisfy the shielding purpose, and by using an efficient discretization with the Boundary Element Method in order to reduce the computational cost. These results have been obtained in collaboration with M. Issa, R. Perrussel and J-R. Poirier (LAPLACE, CNRS/INPT/UPS, Univ. de Toulouse) and O. Chadebec (G2Elab, CNRS/INPG/UJF, Institut Polytechnique de Grenoble). This work has been published in COMPEL - The international journal for computation and mathematics in electrical and electronic engineering, [16].

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

Participant: Victor Péron.

This work is concerned with the time-harmonic eddy current problem for a medium with a highly conductive thin sheet. We present asymptotic models and impedance conditions up to the second order of approximation for the electromagnetic field. The conditions are derived asymptotically for vanishing sheet thickness where the skin depth is scaled like thickness parameter. The first order condition is the perfect electric conductor boundary condition. The second order condition turns out to be a Poincaré-Steklov map between tangential components of the magnetic field and the electric field. This work has been published in SIAM Journal on Applied Mathematics, [18].

7.2. Understanding the interior of the Earth and the Sun by solving inverse problems

7.2.1. Time-Domain Full Waveform Inversion based on high order discontinuous numerical schemes

Participants: Hélène Barucq, Julien Diaz, Pierre Jacquet.

Full Waveform Inversion (FWI) allows retrieving the physical parameters (e.g. the velocity, the density) from an iterative procedure underlying a global optimization technique. The recovering of the medium corresponds to the minimum of a cost function quantifying the difference between experimental and numerical data. In this study we have considered the adjoint state method to compute the gradient of this cost function.

The adjoint state can be both defined as the adjoint of the continuous equation or the discrete problem. This choice is still under study and complementary results has been presented at WAVES 2019 conference in Vienna [38].

The FWI has been largely developed for time-harmonic wave problems essentially because of computational time which is clearly below the one of corresponding time-dependent problems. However, the memory cost in large 3D domain is overflowing the computer capabilities, which motivates us to develop a FWI algorithm in the time-domain. To fully exploit the information from the seismic traces, while preserving the computational cost, it is important to use an accurate and flexible discretization. For that purpose we study several time schemes such as Runge-Kutta 2/4 or Adam-Bashforth 3 and regarding the space discretization, we employ Discontinuous Galerkin (DG) elements which are well-known not only for their h and p adaptivities but also for their massively parallel computation properties.

In the work-flow of DIP a Reverse Time Migration (RTM) code has been developed in collaboration with Total using their Galerkin Discontinuous acoustic time domain solver. Then this code served as a prototype of the time domain FWI code called utFWI (Unstructured Time-Domain Full Waveform Inversion) (https://bil.inria.fr/fr/software/view/3740/tab). Thanks to this code, 2D acoustic multi-scale reconstructions has been performed. Several optimizers such as gradient descent, non linear conjugate gradient and limited BFGS have also been developed.

This work is a collaboration with Henri Calandra (TOTAL). The time domain FWI results has been presented at Total conference MATHIAS 2019 in Paris [37] and also during the Fall Meeting 2019 AGU in San Francisco [47].

7.2.2. Box-Tomography imaging in the deep mantle

Participant: Yder Masson.

Box Tomography is a seismic imaging method (Masson and Romanowicz, 2017) that allows the imaging of localized geological structures buried at arbitrary depth inside the Earth, where neither seismic sources nor receivers are necessarily present. The big advantage of box-tomography over standard tomographic methods is that the numerical modelling (i.e. the ray tracing in travel time tomography and the wave propagation in waveform tomography or full waveform inversion) is completely confined within the small box-region imaged. Thus, box tomography is a lot more efficient than global tomography (i.e. where we invert for the velocity in the larger volume that encompasses all the sources and receivers), for imaging localized objects in the deep Earth. Following a successful, yet partial, application of box tomography to the imaging of the North American continent (i.e. Clouzet et al, 2018), together with Barbara Romanowicz and Sevan Adourian at the Berkeley Seismological Laboratory, we finished implementing the necessary tools for imaging localized structure in the Earth's lower mantle. The following tasks have been completed:

- Modify the global wave propagation solver Specfem_3D_globe in order to compute Green's functions in our current reference Earth model (SEMUCB).
- Modify the local wave propagation solver RegSEM_globe so that the wavefield can be recorded and stored at the surface of the modeling domain.
- Implement real-time compression for an improved management of computed data.

Preliminary results have been presented at the American geophysical union fall meeting 2019 [24]. In the near Future, the latest implementation of Box-Tomography will be deployed to investigate the deep structure under the Yellowstone hotspot down to 20 seconds period.

7.2.3. Wave-propagation modeling using the distributional finite difference method (DFD).

Participant: Yder Masson.

In the last decade, the Spectral Element Method (SEM) has become a popular alternative to the Finite Difference method (FD) for modeling wave propagation in heterogeneous geological media. Though this can be debated, SEM is often considered to be more accurate and flexible than FD. This is because SEM has exponential convergence, it allows to accurately model material discontinuities, and complex structures can be meshed using multiple elements. In the mean time, FD is often thought to be simpler and more computationally efficient, in particular because it relies on structured meshed that are well adapted to computational architectures. This motivated us to develop a numerical scheme called the Distributional Finite Difference method (DFD), which combines the efficiency and the relative simplicity of the finite difference method together with an accuracy that compares to that of the finite/spectral element method. Similarly to SEM, the DFD method divides the computational domain in multiple elements but their size can be arbitrarily large. Within each element, the computational operations needed to model wave propagation closely resemble that of FD which makes the method very efficient, in particular when large elements are employed. Further, large elements may be combined with smaller ones to accurately mesh certain regions of space having complex geometry and material discontinuities, thus ensuring higher flexibility. An exploratory code allowing to model 2D and 3D wave propagation in complex media has been developed and demonstrated the interest of the proposed scheme. We presented numerical examples showing the accuracy and the interest of the DFD method for modeling wave propagation through the Earth at the EGU general assembly, 2019 and at the AGU fall meeting 2019 [41].

7.2.4. Time-harmonic seismic inverse problem

Participants: Hélène Barucq, Florian Faucher.

We study the inverse problem associated with the propagation of time-harmonic wave. In the seismic context, the available measurements correspond with partial reflection data, obtained from one side illumination (only the Earth surface is available). The inverse problem aims at recovering the subsurface Earth medium parameters and we employ the Full Waveform Inversion (FWI) method, which employs an iterative minimization algorithm of the difference between the measurement and simulation.

In particular, we investigate the use of new misfit functionals, based upon the *reciprocity-gap*. The use of such functional is only possible when specific measurements are available, and relates to Green's identity. The feature of the cost function is to allow a separation between the observational and numerical sources. In fact, the numerical sources do not have to coincide with the observational ones, offering new possibilities to create adapted computational acquisitions, and possibilities to reduce the numerical burden.

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

This work has given rise to a publication [11] and a preprint [53]. It has also been presented in several conferences [32], [35], [33], [34].

7.2.5. Convergence analysis for the seismic inverse problem

Participants: Hélène Barucq, Florian Faucher.

The determination of background velocity by Full Waveform Inversion (FWI) is known to be hampered by the local minima of the data misfit caused by the phase shifts associated with background perturbations. Attraction basins for the underlying optimization problems can be computed around any nominal velocity model and guarantee that the misfit functional has only one (global) minimum. The attraction basins are further associated with tolerable error levels which represent the maximal allowed distance between the (observed) data and the simulations (i.e., the acceptable noise level). The estimates are defined a priori, and only require the computation of (possibly many) first and second order directional derivatives of the (model to synthetic) forward map. The geometry of the search direction and the frequency influence the size of the attraction basins, and complex frequency can be used to enlarge the basins. The size of the attraction basins for the perturbation of background velocities in the classical FWI (global model parametrization) and the data space reflectivity (MBTT) reformulation are compared: the MBTT reformulation increases substantially the size of the attraction basins. Practically, this reformulation compensates for the lack of low frequency data.

This work is a collaboration with Guy Chavent (Inria Rocquencourt) and Henri Calandra (TOTAL). The results have been published in Inverse Problems [12] and in the Research Report [43].

7.2.6. Eigenvector representation for the seismic inverse problem

We study the seismic inverse problem for the recovery of subsurface properties in acoustic media. In order to reduce the ill-posedness of the problem, the heterogeneous wave speed parameter is represented using a limited number of coefficients associated with a basis of eigenvectors of a diffusion equation, following the regularization by discretization approach. We compare several choices for the diffusion coefficient in the partial differential equations, which are extracted from the field of image processing. We first investigate their efficiency for image decomposition (accuracy of the representation with respect to the number of variables and denoising). Next, we implement the method in the quantitative reconstruction procedure for seismic imaging, following the Full Waveform Inversion method.
This work is a collaboration with Otmar Scherzer (University of Vienna) and the results have been documented in [49].

7.2.7. Outgoing solutions for the scalar wave equation in Helioseismology

Participants: Hélène Barucq, Florian Faucher, Ha Pham.

We study the construction and uniqueness of physical solutions for the time-harmonic scalar wave equation arising in helioseismology. The definition of outgoing solutions to the equation in consideration or their construction and uniqueness has not been discussed before in the context of helieoseismology. In our work, we use the Liouville transform to conjugate the original equation to a potential scattering problem for Schrödinger operator, with the new problem containing a Coulomb-type potential. Under assumptions (in terms of density and background sound speed) generalizing ideal atmospheric behavior, we obtain existence and uniqueness of variational solutions.

Solutions obtained in this manner are characterized uniquely by a Sommerfeld-type radiation condition at a new wavenumber. The appearance of this wavenumber is only clear after applying the Liouville transform. Another advantage of the conjugated form is that it makes appear the Whittaker special functions, when ideal atmospheric behavior is extended to the whole domain. This allows for the explicit construction of the outgoing Green kernel and the exact Dirichlet-to-Neumann map and hence reference solutions and radiation boundary condition.

This work has given rise to a report of 135 pages, [45]. Some part have been extracted for a publication which has recently been accepted in ESAIM: M2AN.

Consequently to this work, ongoing research includes the fast construction of the Green's kernel, which is possible thanks to the family of special functions obtained from the previous analytical study, [36]. As part of the ANTS associate team, applications to helioseismology is also ongoing. This work is a collaboration with Damien Fournier and Laurent Gizon (Max Planck Institute at Göttingen). The "Ants workshop on computational helioseismology" has also been organized in this context.

7.2.8. Modeling the propagation of acoustic wave in the Sun

Participants: Hélène Barucq, Juliette Chabassier, Marc Duruflé, Nathan Rouxelin.

We study time-harmonic propagation of acoustic waves in the Sun in the presence of gravity forces.

Galbrun's equation, a Lagrangian description of acoustic wave propagation, is usually used in helioseismology. However, the discretization of this equation with high-order discontinuous Galerkin methods leads to poor numerical results for solar-like background flow and geometries. As better numerical results were obtained by using another model, the Linearized Euler's Equations, we investigate the equivalence between those two models.If compatible boundary conditions are chosen, it should be possible to reconstruct the solution of Galbrun's equation by solving the Linearized Euler's Equations and then a vectorial transport equation. It turns out that finding those boundary conditions is quite difficult and not always possible.

We also have constructed a reduced model for acoustic wave propagation in the presence of gravity. Under some additional assumptions on the background flow and for high frequencies, the Linearized Euler's Equations can be reduced to a scalar equation on the pressure perturbation. This equation is well-posed in a usual variational framework and it will provide a useful reference solution to validate our numerical methods. It also seems that a similar process could be used in more realistic cases.

7.2.9. Equivalent boundary conditions for acoustic media with exponential densities. Application to the atmosphere in helioseismology

Participants: Juliette Chabassier, Marc Duruflé, Victor Péron.

We present equivalent boundary conditions and asymptotic models for the solution of a transmission problem set in a domain which represents the sun and its atmosphere. This problem models the propagation of an acoustic wave in time-harmonic regime. The specific non-standard feature of this problem lies in the presence of a small parameter δ which represents the inverse rate of the exponential decay of the density in the atmosphere. This problem is well suited for the notion of equivalent conditions and the effect of the atmosphere on the sun is as a first approximation local. This approach leads to solve only equations set in the sun. We derive rigorously equivalent conditions up to the fourth order of approximation with respect to δ for the exact solution u. The construction of equivalent conditions is based on a multiscale expansion in power series of δ for u. Numerical simulations illustrate the theoretical results. Finally we measure the boundary layer phenomenon by introducing a characteristic length that turns out to depend on the mean curvature of the interface between the subdomains. This work has been published in Applied Mathematics and Computation [15].

7.3. Hybrid time discretizations of high-order

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

Participants: Juliette Chabassier, Julien Diaz.

A paper was accepted in M2AN [14]. This paper deals with the construction of a fourth order, energy preserving, explicit time discretization for dissipative linear wave equations. This family of schemes is obtained by replacing the inversion of a matrix, that comes naturally after using the technique of the Modified Equation on the second order Leap Frog scheme applied to dissipative linear wave equations, by an explicit approximation of its inverse. The series can be truncated at different orders, which leads to several schemes. The stability of the schemes is studied. Numerical results in 1D illustrate the good behavior regarding space/time convergence and the efficiency of the newly derived scheme compared to more classical time discretizations. A loss of accuracy is observed for non smooth profiles of dissipation, and we propose an extension of the method that fixes this issue. Finally, we assess the good performance of the scheme for a realistic dissipation phenomenon in Lorentz's materials. This work has been done in collaboration with Sébastien Imperiale (Inria Project-Team M3DISIM).

7.3.2. Space-Time Discretization of Elasto-Acoustic Wave Equation in Polynomial Trefftz-DG Bases

Participants: Hélène Barucq, Julien Diaz.

In the context of the strategic action "Depth Imaging Partnership" between Inria and Total we have investigated to the development of an explicit Trefftz-DG formulation for elasto-acoustic problem, solving the global sparse matrix by constructing an approximate inverse obtained from the decomposition of the global matrix into a block-diagonal one. The inversion is then justified under a CFL-type condition. This idea allows for reducing the computational costs but its accuracy is limited to small computational domains. According to the limitations of the method, we have investigated the potential of Tent Pitcher algorithms following the recent works of Gopalakrishnan et al. It consists in constructing a space-time mesh made of patches that can be solved independently under a causality constraint. We have obtained very promising numerical results illustrating the potential of Tent Pitcher in particular when coupled with a Trefftz-DG method involving only surface terms.

In this way, the space-time mesh is composed of elements which are 3D objects at most. It is also worth noting that this framework naturally allows for local time-stepping which is a plus to increase the accuracy while decreasing the computational burden. The results of this work [28] have been presented during the ICIAM conference (Valencia, July 15-19).

7.4. Modeling and design of wind musical instruments

7.4.1. Full Waveform inversion for bore reconstruction of woodwind instruments

Participants: Juliette Chabassier, Augustin Ernoult.

Several techniques can be used to reconstruct the internal geometry of a wind instrument from acoustics measurements. One possibility is to simulate the passive linear acoustic response of an instrument and to use an optimization process to fit the simulation to the measurements. This technique can be seen as a first step toward the design of wind instruments, where the targeted acoustics properties come no more longer from measurements but are imposed by the designer. We applied the FWI methodology, along with 1D spectral finite element discretization in space [19], to the woodwind instruments (with tone holes, losses and radiation). The algorithm have been implemented in Pyhton3 and is know operational to reconstruct the bore of real instrument. This functionality will be available in an upcoming version of the toolbox OpenWind.

7.4.2. Computation of the entry impedance of a wind instrument with toneholes and radiation

Participants: Guillaume Castera, Juliette Chabassier, Augustin Ernoult, Alexis Thibault, Robin Tournemenne.

Modeling the entry impedance of wind instruments pipes is essential for sound synthesis or instrument qualification. Based on a one-dimensional model of acoustic propagation (“telegraphist’s equations”) we find approximate solutions using a high-order Finite Element Method (FEM1D). Contrary to the more standard semi-analytic Transfer Matrix Method (TMM), the FEM1D can take into account arbitrarily complex and variable coefficients [19]. It is therefore well-suited for the realistic cases involving boundary losses, smooth waveguide geometry, and possibly even a temperature gradient along the instrument’s bore. We model toneholes as junctions of one-dimensional waveguides, and an acoustic radiation impedance models the radiation of all open tube ends. A global matrix is assembled to connect all these elements together, and solved for each frequency to compute the impedance curve. Source code is available in our Python3 toolbox OpenWind.

7.4.3. Time-domain simulation of a reed instrument with toneholes

Participants: Juliette Chabassier, Augustin Ernoult, Alexis Thibault, Robin Tournemenne.

As part of the project aiming at providing practical tools for instrument design, we have been developing a sound synthesis module for reed instruments. We model a reed music instrument, such as the oboe or the bassoon, as a coupled system composed of a nonlinear source mechanism (the reed), and a linear resonating part (the air within the instrument’s bore). Acoustic wave propagation inside of the instrument is reduced to a one-dimensional model, on which a variational approximation is performed, yielding high-order finite elements in space. Tone holes on the side of the instrument are taken into account using junctions of one-dimensional waveguides. The acoustic radiation impedance is written as a positive Padé approximation, so that it leads to a stable time domain model even when opening and closing holes during the simulation. For the reed, a one-degree-of-freedom lumped model is used, in which the reed opening follows a second order ODE and acts as a valve, modulating the flow that enters the pipe based on the pressure difference between the player’s mouth and the inside of the instrument. Energy-consistent time discretization schemes have been derived for each component, so that it is possible to simulate instruments with an arbitrary geometry with good numerical accuracy and stability. The simulations have been implemented in Python3 and will be made available in an upcoming version of the toolbox OpenWind.

7.4.4. Time-domain simulation of 1D acoustic wave propagation with boundary layer losses

Participants: Juliette Chabassier, Augustin Ernoult, Alexis Thibault.

Energy dissipation effects are of critical importance in musical acoustics. Boundary layer losses occurring in acoustic waveguides are usually modeled in the frequency domain, leading to slowly-decreasing kernels in the time domain similar to fractional derivatives. We have developed an energy-consistent time-domain model based on positive fraction approximation of the dissipative operators, leading to the use of $2N + 1$ additional variables per degree of freedom. Coefficients for these new variables depend only on N so that they can be tabulated without any prior knowledge of the waveguide geometry. They are found with an optimization procedure. This model can be discretized using 1D finite elements [19], and an energy-consistent time-stepping scheme can be found. The resulting numerical scheme has been implemented numerically, and source code will be made available in an upcoming version of the toolbox OpenWind. An article is being written and will be submitted soon to JASA.

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

Participants: Hélène Barucq, Aurélien Citrain, Julien Diaz.

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

A complementary work has been done on mesh generation. The goal was to encircle an unstructured triangle mesh, obtained by third-party softwares, with a quadrangle mesh layer. At first, we built scripts to generate structured triangle meshes, quadrangle meshes and hybrid meshes (triangles surrounded by quadrangles). In 2018, we have implemented the coupling between Discontinuous Galerkin methods (using the triangles/tetrahedra) and Spectral Element methods (using quadrangles/hexahedra). We have also implemented the PML in the SEM part, and we are now working on the local time-stepping feature.

MNEMOSYNE Project-Team

7. New Results

7.1. Overview

This year we have explored the main cortico-basal loops of the cerebral architecture and their associated memory mechanisms. The limbic loop (*cf.* § 7.2) concerns the taking into account of the emotional and motivational aspects by the respondant and operant conditioning and their relations with the semantic and episodic memories. The associative loop (*cf.* § 7.3) is about mechanisms of working memory and rule manipulation. Concerning the motor loop (*cf.* § 7.4), we have studied mechanisms of song acquisition and production in birds.

We have also worked on the systemic integration of our models (*cf.* § 7.5), raising the question of the conditions of autonomous learning.

Finally, we study the links between our bio-inspired modeling work and other domains like Machine Learning, computer science and educational science (*cf.* § 7.6).

7.2. The limbic loop

Our main contribution this year to advancing our view of the limbic loop is the defense of the PhD of B. T. Nallapu [1] related to the modeling of the orbital and medial loops the two main constituents of the limbic loop, as described in [33]. In short, this work proposes, instead of a global view of the orbital loop as generally proposed, a view with two loops, one corresponding to the lateral part of the orbitofrontal cortex related to respondent conditioning and one to the medial part related to operant conditioning and closely linked to the medial loop. The work shows that such a model replicates a variety of observations in neuroscience and can be used for the autonomous behavior of an agent in the Minecraft video game.

7.3. The associative loop

The prefrontal cortex is known to be involved in many high-level cognitive functions, in particular working memory. Gated working memory is defined as the capacity of holding arbitrary information at any time in order to be used at a later time. Based on electrophysiological recordings, several computational models have tackled the problem using dedicated and explicit mechanisms. We propose instead to consider an implicit mechanism based on a random recurrent neural network. We introduce a robust yet simple reservoir model of gated working memory with instantaneous updates [11]. The model is able to store an arbitrary real value at random time over an extended period of time. The dynamics of the model is a line attractor that learns to exploit reentry and a non-linearity during the training phase using only a few representative values. A deeper study of the model shows that there is actually a large range of hyper parameters for which the results hold (number of neurons, sparsity, global weight scaling, etc.) such that any large enough population, mixing excitatory and inhibitory neurons can quickly learn to realize such gated working memory. This suggests this property could be an implicit property of any random population, that can be acquired through learning. Furthermore, considering working memory to be a physically open but functionally closed system, we give account on some counter-intuitive electrophysiological recordings.

We also developed a model of working memory combining short-term and long-term components [24]. For the long-term component, we used Conceptors in order to store constant temporal patterns. For the short-term component , we used the Gated-Reservoir model [11]. We combined both components in order to obtain a model in which information can go from long-term memory to short-term memory and vice-versa.

The prefrontal cortex is also known to be the place where complex and abstract behavioral rules are implemented. In order to study the mechanisms related to the manipulation of such rules, we have begun the study of networks able to build rules to manipulate such framework as the Wisconsin Card Sorting Test, widely used in the clinical domain [3].

7.4. The motor loop

Sensorimotor learning represents a challenging problem for artificial and natural systems. Several computational models try to explain the neural mechanisms at play in the brain to implement such learning. These models have several common components: a motor control model, a sensory system and a learning architecture. In S. Pagliarini's PhD, our challenge is to build a biologically plausible model for song learning in birds including neuro-anatomical and developmental constraints.

We made a review on a specific type of sensorimotor learning referred to as imitative vocal learning and examplified by song learning in birds or human complex vocalizations[35]. Sensorimotor learning represents a challenging problem for natural and artificial systems. Several computational models have been proposed to explain the neural and cognitive mechanisms at play in the brain [34]. In general, these models can be decomposed in three common components: a sensory system, a motor control device and a learning framework. The latter includes the architecture, the learning rule or optimisation method, and the exploration strategy used to guide learning. In this review, we focus on imitative vocal learning, that is exemplified in song learning in birds and speech acquisition in humans. We aim to synthesise, analyse and compare the various models of vocal learning that have been proposed, highlighting their common points and differences. We first introduce the biological context, including the behavioural and physiological hallmarks of vocal learning and sketch the neural circuits involved. Then, we detail the different components of a vocal learning model and how they are implemented in the reviewed models.

On this topic, X. Hinaut is also collaborating with Catherine del Negro's team (CNRS, NeuroPSI, Orsay) on the representation of syntax in songbird brains. In particular, the project aims at (1) linking the neural activity of a sensori-motor area (HVC) to syntax elements in the songs of domestic canaries ; (2) analysing the audio files and transcripts of canary songs in order to find syntax cues and higher order representations (graph properties of songs, evaluate Markovian forward and backward transition probabilities of various orders). The song transcription and analyses has been done by M1 intern (in Neuroscience) Juliette Giraudon, and preliminary work on song segmentation and classification has been done by L3 intern (ENS) Pierre Marcus. In December, Aurore Cazala (student in the NeuroPSI collaborator team since 2014) defended her PhD at the University of Paris-Saclay on "Codage neuronal de l'ordre des signaux acoustiques dans le chant des oiseaux chanteurs"; X. Hinaut participated in studies done in this PhD.

7.5. Systemic integration

Several global approaches corresponding to the systemic integration of several loops have been studied this year. The PhD work [1] evoked in section § 7.2 for its contributions to modeling the limbic loop, has also been partly devoted to the definition of a global cognitive model and its embodiment in an agent in the Minecraft game, in order to illustrate the performances of autonomous behavior of a system endowed by such a limbic loop [33].

7.6. Association to other scientific domains

Concerning Machine Learning and our work on reservoir computing, X. Hinaut is collaborating with Michael Spranger (Sony Lab, Tokyo, Japan) on grounding of language, adapting Hinaut's previous Reservoir parser (ResPars) with the representational system of Spranger: IRL (Incremental Recruitment Language) [17]. He is also collaborating with Hamburg on the use of reservoir models for robotic tasks (*cf.* § 9.3). In this work, we have shown that the RLM can successfully learn to parse sentences related to home scenarios in fifteen languages [5]. This demonstrates that (1) the learning principle of our model is not limited to a particular language (or particular sentence structures), and (2) it can deal with various kinds of representations (not only predicates), which enable users to adapt it to their own needs. Some people can mix two languages within the same sentence: this is known as intra-sentential code-switching. With M1 intern Pauline Detraz, we collected data from human subjects that were required to mix pairs of given sentences in French and English. The corpus obtained have some very complex mixed sentences: there can be until eleven language switches within the same sentence. Then, we trained our Reservoir-based sentence Parsing model, with the collected corpus.

Surprisingly the model is able to learn and generalize on the mixed corpus with performances nearly as good as the unmixed French-English corpus [16]. A post-doc joined the team in Nov 2019 to work on the project HuRRiCane ("Hierarchical Reservoir Computing for Language Comprehension") project founded by Inria. This project aims at extending the ResPars model to work from speech inputs to sentence comprehension including coherancy checking. In other words, the objective is to experiment how a sentence comprehension model, based on reservoir computing, can learn to understand sentences by exploring which meanings can have the sentences, implying several steps from stream of phonemes to words and from stream of words to sentence comprehension. The model will be implemented on a virtual agent first and then on the Nao humanoid robot.

This project is linked to other projects in the team on the hierarchical organization of the prefrontal cortex (including Broca's area, involved in language). This hierarchy corresponds to an increasingly higher abstraction, which is made by different sub-areas. We will therefore be able to link this post-doc project to existing projects of the team, where different levels of abstractions are necessary for sentence comprehension.

As explained in § 7.6 , song segmentation and classification has been done by L3 intern (ENS) Pierre Marcus.

The on-going work on an original prototype based approach of deep-learning considering not so big data sets, and targeting also interpretability of the result, has been finalized [4], including a fine study on metaparameter adjustment in this context, while both standard learning and meta-learning paradigms have been considered. The capability to easily the "how it works" mechanism to no specialist of the field is an important outcome of the paper.

Co-leaded by Margarida Romero scientific director of the LINE laboratory of the UCA and researchers of our team, a preliminary work regarding artificial intelligence devoted to education (AIDE) was developed to study applications to educational science. This first year has been devoted to study to which extents the existing collaboration between Inria science outreach regarding computational thinking initiation and educational science research in order to be understand the underlying cognitive processes of the former actions and evaluate them, could be enlarged to multi-disciplinary research in both fields. The first outcome of this collaboration has been an analysis of a computational thinking unplugged activity under the perspective of embodied cognition [9] and deep and large review in the field, analyzing how computational thinking in K-12 education could be developed [22], within the scope of studies regarding co-creativity, robotics and maker education. [20], while a qualitative analysis one very large audience (more then 18000 inscriptions) on-line course outcomes has been published [18], with some operational outcomes regarding enlarging comptational thinking training from teachers to all citizens [12].

Software is a fundamental pillar of modern scientific research, across all fields and disciplines. However, there is a lack of adequate means to cite and reference software due to the complexity of the problem in terms of authorship, roles and credits. This complexity is further increased when it is considered over the lifetime of a software that can span up to several decades. Building upon the internal experience of Inria, the French research institute for digital sciences, we provide in this paper a contribution to the ongoing efforts in order to develop proper guidelines and recommendations for software citation and reference. Namely, we recommend: (1) a richer taxonomy for software contributions with a qualitative scale; (2) to put humans at the heart of the evaluation; and (3) to distinguish citation from reference.

MONC Project-Team

7. New Results

7.1. Machine learning and mechanistic modeling for prediction of metastatic relapse in early-stage breast cancer

Authors: *C. Nicolò; C. Périer*; M. Prague; C. Bellera; G. MacGrogan; *O.Saut; S. Benzekry*. Accepted for publication in the Journal of Clinical Oncology: Clinical Cancer Informatics.

Purpose: For patients with early-stage breast cancer, prediction of the risk of metastatic relapse is of crucial importance. Existing predictive models rely on agnostic survival analysis statistical tools (e.g. Cox regression). Here we define and evaluate the predictive ability of a mechanistic model for the time to metastatic relapse.

Methods: The data consisted of 642 patients with 21 clinicopathological variables. A mechanistic model was developed on the basis of two intrinsic mechanisms of metastatic progression: growth (parameter α) and dissemination (parameter μ). Population statistical distributions of the parameters were inferred using mixed-effects modeling. A random survival forest analysis was used to select a minimal set of 5 covariates with best predictive power. These were further considered to individually predict the model parameters, by using a backward selection approach. Predictive performances were compared to classical Cox regression and machine learning algorithms.

Results: The mechanistic model was able to accurately fit the data. Covariate analysis revealed statistically significant association of Ki67 expression with α (p=0.001) and EGFR with μ (p=0.009). Achieving a c-index of 0.65 (0.60-0.71), the model had similar predictive performance as the random survival forest (c-index 0.66-0.69) and Cox regression (c-index 0.62 - 0.67), as well as machine learning classification algorithms.

Conclusion: By providing informative estimates of the invisible metastatic burden at the time of diagnosis and forward simulations of metastatic growth, the proposed model could be used as a personalized prediction tool of help for routine management of breast cancer patients.

7.2. Numerical workflow for clinical electroporation ablation

Authors: *Olivier Gallinato*, *Baudouin Denis de Senneville*, Olivier Seror, *Clair Poignard*. Published in Physics in Medicine and Biology. https://hal.inria.fr/hal-02063020/file/paperIRE_workflow_R3.pdf.

The paper describes a numerical workflow, based on the “real-life” clinical workflow ofirreversible electroporation (IRE) performed for the treatment of deep-seated liver tumors. Thanksto a combination of numerical modeling, image registration algorithm and clinical data, our numerical workflow enables to provide the distribution of the electric field as effectively delivered by theclinical IRE procedure. As a proof of concept, we show on a specific clinical case of IRE ablation ofliver tumor that clinical data could be advantageously combined to numerical simulations in a nearfuture , in order to give to the interventional radiologists information on the effective IRE ablation.We also corroborate the simulated treated region with the post-treatment MRI performed 3 days after treatment.

7.3. A reduced Gompertz model for predicting tumor age using a population approach

Authors: *C. Vaghi*, A. Rodallec, R. Fanciullino, J. Ciccolini, J. Mochel, M. Mastri, *C. Poignard*, J. ML Ebos, *S. Benzekry*. Accepted for publication in PLoS Computational Biology. https://www.biorxiv.org/content/10.1101/670869v2

Tumor growth curves are classically modeled by means of ordinary differential equations. In analyzing the Gompertz model several studies have reported a striking correlation between the two parameters of the model, which could be used to reduce the dimensionality and improve predictive power.

We analyzed tumor growth kinetics within the statistical framework of nonlinear mixed-effects (population approach). This allowed the simultaneous modeling of tumor dynamics and inter-animal variability. Experimental data comprised three animal models of breast and lung cancers, with 833 measurements in 94 animals. Candidate models of tumor growth included the exponential, logistic and Gompertz. The exponential and – more notably – logistic models failed to describe the experimental data whereas the Gompertz model generated very good fits. The previously reported population-level correlation between the Gompertz parameters was further confirmed in our analysis ($R^2 > 0.92$ in all groups). Combining this structural correlation with rigorous population parameter estimation, we propose a reduced Gompertz function consisting of a single individual parameter (and one population parameter). Leveraging the population approach using Bayesian inference, we estimated times of tumor initiation using three late measurement timepoints. The reduced Gompertz model was found to exhibit the best results, with drastic improvements when using Bayesian inference as compared to likelihood maximization alone, for both accuracy and precision. Specifically, mean accuracy was 12.2% versus 78% and mean precision was 15.6 days versus 210 days, for the breast cancer cell line.

These results offer promising clinical perspectives for the personalized prediction of tumor age from limited data at diagnosis. In turn, such predictions could be helpful for assessing the extent of invisible metastasis at the time of diagnosis.

The code and the data used in our analysis are available at https://github.com/cristinavaghi/plumky.

7.4. T2-based MRI Delta-Radiomics Improve Response Prediction in Soft-Tissue Sarcomas Treated by Neoadjuvant Chemotherapy

Authors: *Amandine Crombé*, *Cynthia Perier*, Michèle Kind, *Baudouin Denis de Senneville*, Francois Le Loarer, Antoine Italiano, Xavier Buy, *Olivier Saut*. Published in Journal of Magnetic Reasonance Imaging https://hal.inria.fr/hal-01929807v2

Background: Standard of care for patients with high-grade soft-tissue sarcoma (STS) are being redefined since neoadjuvant chemotherapy (NAC) has demonstrated a positive effect on patients' outcome. Yet, response evaluation in clinical trials still remains on RECIST criteria.

Purpose: To investigate the added value of a Delta-radiomics approach for early response prediction in patients with STS undergoing NAC Study type: Retrospective Population: 65 adult patients with newly-diagnosed, locally-advanced, histologically proven high-grade STS of trunk and extremities. All were treated by anthracycline-based NAC followed by surgery and had available MRI at baseline and after 2 cycles. Field strength/Sequence: Pre- and post-contrast enhanced T1-weighted imaging (T1-WI), turbo spin echo T2-WI at 1.5T.

Assessment: A threshold of <10% viable cells on surgical specimen defined good response (Good-HR). Two senior radiologists performed a semantic analysis of the MRI. After 3D manual segmentation of tumors at baseline and early evaluation, and standardization of voxelsizes and intensities, absolute changes in 33 texture and shape features were calculated. Statistical tests: Classification models based on logistic regression, support vector machine, k-nearest neighbors and random forests were elaborated using cross-validation (training and validation) on 50 patients ('training cohort') and was validated on 15 other patients ('test cohort').

Results: 16 patients were good-HR. Neither RECIST status, nor semantic radiological variables were associated with response except an edema decrease (p=0.003) although 14 shape and texture features were (range of p-values: 0.002-0.037). On the training cohort, the highest diagnostic performances were obtained with random forests built on 3 features, which provided: AUROC=0.86, accuracy=88.1%, sensitivity=94.1%, specificity=66.3%. On the test cohort, this model provided an accuracy of 74.6% but 3/5 good-HR were systematically ill-classified.

Data conclusions: A T2-based Delta-Radiomics approach can improve early response prediction in STS patients with a limited number of features.

7.5. Quantitative mathematical modeling of clinical brain metastasis dynamics in non-small cell lung cancer

Authors: *M. Bilous*, C. Serdjebi, A. Boyer, P. Tomasini, *C. Pouypoudat*, D. Barbolosi, F. Barlesi, F. Chomy, *S. Benzekry*. Published in Scientific Reports. https://hal.inria.fr/hal-01928442v2

Brain metastases (BMs) are associated with poor prognosis in non-small cell lung cancer (NSCLC), but are only visible when large enough. Therapeutic decisions such as whole brain radiation therapy would benefit from patient-specific predictions of radiologically undetectable BMs. Here, we propose a mathematical modeling approach and use it to analyze clinical data of BM from NSCLC. Primary tumor growth was best described by a gompertzian model for the pre-diagnosis history, followed by a tumor growth inhibition model during treatment. Growth parameters were estimated only from the size at diagnosis and histology, but predicted plausible individual estimates of the tumor age (2.1–5.3 years). Multiple metastatic models were further assessed from fitting either literature data of BM probability (n = 183 patients) or longitudinal measurements of visible BMs in two patients. Among the tested models, the one featuring dormancy was best able to describe the data. It predicted latency phases of 4.4–5.7 months and onset of BMs 14–19 months before diagnosis. This quantitative model paves the way for a computational tool of potential help during therapeutic management.

7.6. Optimizing 4D abdominal MRI: Image denoisingusing an iterative back-projection approach

Authors: *B Denis de Senneville*, C R Cardiet, A J Trotier, E J Ribot, L Lafitte, L Facq, S Miraux. Published in Physics in Medicine and Biology. https://hal.archives-ouvertes.fr/hal-02367839/file/2019_4D_reconstruction_revision2_final.pdf

4D-MRI is a promising tool for organ exploration, target delineationand treatmentplanning. Intra-scan motion artifacts may be greatly reduced by increasing the imagingframe rate. However, poor signal-to-noise ratios (SNR) are observed when increasingspatial and/or frame number per physiological cycle, in particular inthe abdomen.In the current work, the proposed 4D-MRI method favored spatial resolution, framenumber, isotropic voxels and large field-of-view (FOV) during MR-acquisition. Theconsequential SNR penalty in the reconstructed data is addressed retrospectively usingan iterative back-projection (IBP) algorithm. Practically, after computing individualspatial 3D deformations present in the images using a deformable image registration(DIR) algorithm, each 3D image is individually enhanced by fusing several successiveframes in its local temporal neighborood, these latter being likely tocover commonindependent informations. A tuning parameter allows one to freely readjust the balancebetween temporal resolution and precision of the 4D-MRI.The benefit of the method was quantitatively evaluated on the thorax of 6 miceunder free breathing using a clinically acceptable duration. Improved 4D cardiacimaging was also shown in the heart of 1 mice. Obtained results are comparedto theoretical expectations and discussed. The proposed implementation is easilyparallelizable and optimized 4D-MRI could thereby be obtained with a clinicallyacceptable duration.

7.7. Optimal Scheduling of Bevacizumab and Pemetrexed/cisplatin Dosing in Non-Small Cell Lung Cancer

Authors: Benjamin Schneider, Arnaud Boyer, Joseph Ciccolini, Fabrice Barlési, Kenneth Wang, *Sébastien Benzekry**, Jonathan Mochel*. * = co-senior authors. Published in CPT: Pharmacometrics and Systems Pharmacology. https://hal.inria.fr/hal-02109335

Bevacizumab-pemetrexed/cisplatin (BEV-PEM/CIS) is a first line therapeutic for advanced non-squamous non-small cell lung cancer (NSCLC). Bevacizumab potentiates PEM/CIS cytotoxicity by inducing transient tumor vasculature normalization. BE V- PEM/CIS has a narrow therapeutic window. Therefore, it is an attractive target for administration schedule optimization. The present study leverages our previous work on BEV-PEM/CIS pharmacodynamic modeling in NSCLC-bearing mice to estimate the optimal gap in the scheduling of sequential BEV-PEM/CIS. We predicted the optimal gap in BEV-PEM/CIS dosing to be 2.0 days

in mice and 1.2 days in humans. Our simulationssuggest that the efficacy loss in scheduling BEV-PEM/CIS at too great of a gap is much less than the efficacy loss in scheduling BEV-PEM/CIS at too short of a gap.

PLEIADE Project-Team

7. New Results

7.1. Genetic Determinisms of Aborted Fermentation in Winemaking

This study contributes to the understanding of the mechanisms leading to stuck fermentation in winemaking, an economic prejudice [7]. A number of factors can trigger stuck or aborted fermentation such as high temperature, high ethanol concentration, low pH. The biodiversity of natural yeast strains used in winemaking starters has as a consequence that some of them are more prone to abort fermentation than others, indicating a genetic determinism. Crosses between strains called "sensitive" or "resistant" to stuck fermentation occurrence, followed by back-crosses with the "sensitive" parent while selecting for the "resistant" phenotype, allowed us to reduce the amount of genetic material inherited from the "resistant" parent in the progeny, ending to 3 small introgression areas after 4 generations. Quantitative Trait Locus (QTL) detection in this progeny (77 strains) involved characterization of SNP inheritence (circa 1200 validated SNPs) from either parent, through micro-array hybridization, mapping of the SNP on the reference genome and phenotypic measurements on the progeny. This analysis made it possible to detect two genes which, when inactivated by naturally occurring mutations, act as major perturbators of several fermentation parameters in winemaking physiological conditions. Consequently, our industrial partner incorporated into its catalogue of winemaking starters, strains carrying the functional forms of these genes.

7.2. Characterization of Molecular Biodiversity

In 2019 PLEIADE developed new methods for characterizing molecular biodiversity (see [4], [5] for applications). This point itself has been developed with two approaches in 2019, each with the beginning of a PhD.

- Building OTUs from a pairwise distance matrix typically is an unsupervised clustering issue. In this new development, the PhD student (PLEIADE) tests whether SBM (Stochastic Block Model) approach yields relevant results for a global characterization of biodiversity beyond a summary by a scalar index. This is done in collaboration with MIAT INRAE research unit in Toulouse and EPC HiePACS. It represents a connection between metabarcoding and statistical modeling, a topic which deserves investigation and is expanding. (Figure 5 from [8])
- A major goal of PLEIADE is to develop a geometric view on biodiversity. The tool selected up to now is to associate a point cloud to a dataset (pairwise distances between sequences) and study its shape. In 2019, PLEIADE has been associated in a collaboration with HiePACS to begin a new topic: comparison between point clouds, each cloud being associated to a data set. Indeed, the development of metabarcoding leads to the new issue of comparison between OTUs built from different dataset. This approach is part of the issues raised in a PhD supervized by HiePACS, in collaboration with PLEIADE.

7.3. Scaling Metabarcoding Programs

Metabarcoding is a series of technical procedures to build molecular based inventories from large datasets of amplicons. We derived new methods and tools to scale metabarcoding programs in collaboration with EPC HiePACS. This has been realized through following participation in research projects:

- Contribution to ADT Gordon project in Inria BSO. The objective of this project (partners: Tadaam (coordinator), STORM, HiePACS, PLEIADE) is to integrate SVD as an available tool in Chameleon, starPU and new Madeleine. The contribution of Pleiade is to bring metabarcoding as a use case, and random projection ([6]) as a method for scaling Multidimensional Scaling (which requires an SVD) in collaboration with HiePACS with a template implemented in Diodon. In 2019, PLEIADE has bought to the project a series of 55 matrices with size about 10^5 rows/column which, assembled, yield a full pairwise distance matrix between one million of sequences. The objective is to reach the million in 2020.

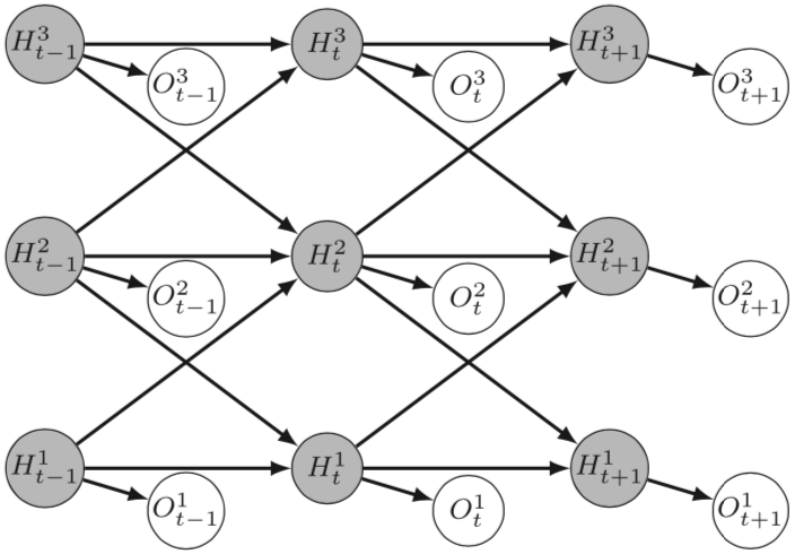


Figure 5. Graphical representation of a coupled HMM with three hidden chains (from [8])

- Contribution to Region Nouvelle Aquitaine project "HPC Scalable Ecosystem." This project is chaired by HiePACS. In collaboration with this EPC, PLEIADE is involved in developing a new approach for comparing OTUs built from different datasets.

7.4. Linking Homology and Function for Algal Desaturases

Polyunsaturated fatty acids (PUFA) such as Omega-3 that are essential for human health cannot be synthesized by the human body and must be acquired through the consumption of certains foods, such as oily fish, that are becoming increasing difficult to produce sustainably. However, fish do not produce them either: their role is to concentrate PUFA through the food chain. There is consequently considerable interest in producing these essential nutriments directly, through the cultivation of domesticated strains of naturally occurring or of engineered strains of green algae.

Ultimately, polyunsaturated fatty acids are produced by molecular machines called **desaturases**. While desaturases are abundant in all branches of life, the link between gene sequence and the precise activity of the corresponding enzyme is poorly understood. The particular challenge is that, while the catalytic active site is well conserved, the features that recognize the substrate and that determine the regiospecificity of the enzyme are not. In order to produce specific PUFA at industrial scale, it is necessary to develop efficient tools for high-throughput identification of candidate genes in algal species, and precise models for designing desaturases through synthetic biology.

In collaboration with the LBM (UMR 5200 CNRS) and with the support of the Inria Project Lab *In silico algae*, we used a core collection of thirteen desaturases from *Osteococcus tauri* to explore the link between homology and function in 23 species ([9]). The study reinforced our understanding of the evolutionary conservation of desaturases and confirmed the identification of substrate and regio- specificity through graph neighborhoods. We were further able to extend the identification of PPR motifs correlated with specificity. This work is ongoing. Since most of the pertinent desaturases are membrane bound, the prediction of protein structure has proved perilous, but we are hopeful that future work will allow us to use structure-inspired prediction to narrow in on the sites responsible for specificity despite their poor sequence conservation.

SISTM Project-Team

7. New Results

7.1. Mechanistic learning

7.1.1. Ebola models

New models have been developed for the response to the Ebola vaccine. The first one has been fitted to Phase 1 trials and has given interesting predictions of the long term duration of the response that are confirmed with the new data coming from phase 2 trials. These results have been published in Journal of Virology. Then, a new model including the B cell memory response has been defined and its mathematical proprieties have been studied. A manuscript has been submitted to Journal of Theoretical Biology. The next step is to estimate model parameters using EBL2001 clinical trial data.

New publication: Pasin C et al. Dynamics of the Humoral Immune Response to a Prime-Boost Ebola Vaccine: Quantification and Sources of Variation. J Virol. 2019 Aug 28;93(18). pii: e00579-19. doi: 10.1128/JVI.00579-19. Print 2019 Sep 15.

7.1.2. Estimation method

A new approach is currently under development by Quentin Clairon to estimate model parameters using a regularization method based on the control theory. The estimation method used an approximation of the original ODE solution for each subject. The expected advantages of this approach are i) to mitigate the effect of model misspecification on estimation accuracy ii) to regularize the estimation problem in presence of poorly identifiable parameters, iii) to avoid estimation of initial conditions. The method is still under development but preliminary results have been presented at the Viral dynamics conference in October 2019.

7.2. High-dimensional and statistical learning

7.2.1. Automatic analysis of cell populations

New publication:

Hejblum BP, Alkhassim C, Gottardo R, Caron F, Thiébaut R, Sequential Dirichlet process mixture of skew t-distributions for model-based clustering of flow cytometry data, Annals of Applied Statistics, 13(1):638-660, 2019. DOI: 10.1214/18-AOAS1209.

7.2.2. High-dimensional compositional data analysis

Perrine Soret (PhD student in the axis "High-dimensional and statistical learning", supervised by M. Avalos) has applied our expertise in high dimensional data analysis to human microbiome field of research:

Soret P, Vandenborght LE, Francis F, Coron N, Enaud R, The Mucofong Investigation Group, Avalos M, Schaeverbeke T, Berger P, Fayon M, Thiébaut R and Delhaes L. Respiratory mycobiome and suggestion of inter-kingdom network during acute pulmonary exacerbation in cystic fibrosis. To appear in *Scientific Reports*.

7.2.3. Missing Value Treatment in Longitudinal High Dimensional Supervised Problems

Poor blood sample quality introduces a large number of missing values in the context of sequencing data production. Furthermore, strong technical biases may force the analyst to remove the considered sequenced samples. Then entire day dependent data are then missing. Hadrien Lorenzo (PhD student in the axis "High-dimensional and statistical learning", supervised by J. Saracco and R. Thiébaut) has developed a multi-block approach: the dd-sPLS method. dd-sPLS has been applied to high dimensional data analysis of different fields of research:

Lorenzo, H., Misbah, R., Odeber, J., Morange, P. E., Saracco, J., Trégouët, D. A., and Thiébaut, R. High-dimensional multi-block analysis of factors associated with thrombin generation potential. In 2019 IEEE 32nd International Symposium on Computer-Based Medical Systems (CBMS) (pp. 453-458). IEEE. https://hal.archives-ouvertes.fr/hal-02429302

Ellies-Oury, M. P., Lorenzo, H., Denoyelle, C., Saracco, J., and Picard, B. An Original Methodology for the Selection of Biomarkers of Tenderness in Five Different Muscles. Foods, 8(6), 206 (2019). https://hal.archives-ouvertes.fr/hal-02164157

7.3. Translational vaccinology

7.3.1. HIV vaccine development

We have finalized the data science analyses of two HIV vaccine clinical trials: 1) ANRS VRI01, a randomized phase I/II trial evaluating for different prime boost vaccine strategies in healthy volunteers; 2) ANRS 149 LIGHT, a randomized phase II trial comparing a prime-boost therapeutic HIV vaccine strategy to placebo in HIV-infected patients undergoing antiretroviral treatment interruption. This included integrative statistical analyses using sPLS methods (as developed by the team) to relate markers from different high-dimensional immunogenicity or gene expression assays or virological assays to each other. In the ANRS VRI01 data set this allowed to disentangle the immune responses induced by the different vaccines used in the prime-boost strategies, showing specific effects of one of the vaccines (MVA HIV-B). We further identified a gene expression signature that correlates with later functional T-cell responses across the three different prime-boost association in which the MVA HIV-B vaccine was used. The corresponding manuscripts are currently in preparation.

Other HIV vaccine trials are currently being set-up by the French VRI (Vaccine Research Institute) and the European consortium EHVA, with strong contributions of SISTM team members to the trial designs.

7.3.2. Ebola vaccine development

The main results of the two randomized phase II Ebola vaccine trials conducted by the IMI-2 EBOVAC2 consortium (coordinated by Rodophe Thiébaut from the SISTM team) were finalized in 2019. and presented at international conferences. The results showed that the tested vaccine strategy (two-dose heterologous Ad26.ZEBOV and MVA-BN®-Filo Ebola vaccine regimen, developed by Janssen) was safe and immunogenic in both European and African volunteers (EBL2001 and EBL2002 trials). Deeper analyses of the induced immune responses are currently ongoing, and systems vaccinology analyses will soon start in the SISTM team.

The SISTM team is also a partner in the related IMI-2 EBOVAC1 and EBOVAC3 consortia (assesing the same vaccine regimen in other trial populations and/or trial phases), in which the main contributions of the team are related to mechanistic modeling of the immune responses (ongoing).

7.3.3. Vaccine development against other pathogens

Two other phase I vaccine trials (one testing a placental malaria vaccine, Primalvac trial; and one testing a nasal Pertussis vaccine, BPZE-1 trial), in which members of the SISTM team were strongly involved, have shown promising results. Results of the Primalvac trial have been accepted for publication in the Lancet Infectious Disease journal, and results of the BPZE-1 trial have been submitted for publication.

7.3.4. Methodological developments for vaccine trials

At the interface between the axis on "Mechanistic learning" and the axis "Translational vaccinology", modelling done within a PhD project (M. Alexandre, supervised by R. Thiébaut and M. Prague) has informed the definition of the primary endpoint and statistical analysis method to be used in two therapeutic HIV vaccine trials with antiretroviral treatment interruption (EHVA T02 and ANRS DALIA-2). The methodological choices and their rationale have been presented to the governance bodies of the research consortia, patient associations and have been submitted for ethics and regulatory approvals. This will be subject to a specific methodology publication.

Edouard Lhomme (PhD student in the axis "Translational Vaccinology, supervised by L. Richert) has developed a statistical method for functional T-cell assay data from vaccine trials (in particular the intracellular cytokine staining assay) that takes into account non-specific immune responses. We propose using a bivariate linear model for the analysis of the cellular immune responses to obtain accurate estimations of the vaccine effect. We benchmarked the performance of the model in terms of both bias and control of type-I and -II errors, and applied it to simulated data as well as real pre- and post-vaccination data from two recent HIV vaccine trials (ANRS VRI01 and ANRS 149 LIGHT in HIV-infected participants). This method has been published in the Journal of Immunological Methods and is now used in the SISTM team as the standard method for analyses of functional cellular data with non-stimulated control conditions, for instance in the currently ongoing analysis of cellular proliferation data from the EBOVAC2 EBL2001 trial. We have also established an online interface based on R Shiny to make this analysis method available for use by immunologists without specific training in statistical modelling (https://shiny-vici.apps.math.cnrs.fr/).

7.3.5. Prediction of the survival of patients based on RNA-seq data

In collaboration with the Inria MONC team, with the Inserm Angiogenesis and Tumor micro-environment team, and with clinicians from Milan and Bergen, the project GLIOMA-PRD aims to improve the prediction of the evolution of the lower grade glioma, a primary brain tumor, based on clinical, imaging and genomic data. In the SISTM team, we first had to determine the sufficient sample size for determining a predictive signature based on RNA-seq data for the survival of the patients. We concluded that 50 patients were a good enough sample size for this aim. We then explored the potential methods to analyze these data, with a particular focus on the methods grouping the genes by pathways, as the pilot data (The Cancer Genome Atlas Research Network, NEJM, 2015) showed a high correlation structure. We particularly compared two methods, Generalized Berk-Jones (GBJ), proposed by [54], and tcgsaseq, proposed by [1]. The first method could be applied to the survival context and thus be appropriate for our data.

Once the RNA-seq data were available, we could observe a high batch effect since the data were sequenced in two different lanes. One of these two batch included only patients that exhibited a particular tumor at the PET-scan, called "COLD", while the second batch included both patient labelled as "COLD", but also the other type of tumor, called "DIFFUSE". As this experiment structure might leads to confusing the difference due to the batch effect with the one due to the biological different, we had to explore methodologies that could remove this batch effect.

Once this batch effect removed, we will then analyze the RNA-seq data in order to identify the genes, or the group of genes, that could be predictive of the survival of the patients.

HIEPACS Project-Team

7. New Results

7.1. High-performance computing on next generation architectures

7.1.1. Memory optimization for the training phase of deep convolutional networks

Training Deep Neural Networks is known to be an expensive operation, both in terms of computational cost and memory load. Indeed, during training, all intermediate layer outputs (called activations) computed during the forward phase must be stored until the corresponding gradient has been computedin the backward phase. These memory requirements sometimes prevent to consider larger batch sizes and deeper networks, so that they can limit both convergence speed and accuracy. Recent works have proposed to offload some of the computed forward activations from the memory of the GPU to the memory of the CPU. This requires to determine which activations should be offloaded and when these transfers from and to the memory of the GPU should take place. In [28], We prove that this problem is NP-hard in the strong sense, and we propose two heuristics based on relaxations of the problem. We perform extensive experimental evaluation on standard Deep Neural Networks. We compare the performance of our heuristics against previous approaches from the literature, showing that they achieve much better performance in a wide variety of situations.

In [23], we also introduce a new activation checkpointing method which allows to significantly decrease memory usage when training Deep Neural Networks with the back-propagation algorithm. Similarly to checkpointing techniques coming from the literature on Automatic Differentiation, it consists in dynamically selecting the forward activations that are saved during the training phase, and then automatically recomputing missing activations from those previously recorded. We propose an original computation model that combines two types of activation savings: either only storing the layer inputs, or recording the complete history of operations that produced the outputs (this uses more memory, but requires fewer recomputations in the backward phase), and we provide an algorithm to compute the optimal computation sequence for this model. This paper also describes a PyTorch implementation that processes the entire chain, dealing with any sequential DNN whose internal layers may be arbitrarily complex and automatically executing it according to the optimal checkpointing strategy computed given a memory limit. Through extensive experiments, we show that our implementation consistently outperforms existing checkpointing approaches for a large class of networks, image sizes and batch sizes.

In [4], [24], we consider the problem of optimally scheduling the backpropagation of Deep Join Networks. Deep Learning training memory needs can prevent the user to consider large models and large batch sizes. In this work, we propose to use techniques from memory-aware scheduling and Automatic Differentiation (AD) to execute a backpropagation graph with a bounded memory requirement at the cost of extra recomputations. The case of a single homogeneous chain, i.e. the case of a network whose all stages are identical and form a chain, is well understood and optimal solutions have been proposed in the AD literature. The networks encountered in practice in the context of Deep Learning are much more diverse, both in terms of shape and heterogeneity. In this work, we define the class of backpropagation graphs, and extend those on which one can compute in polynomial time a solution that minimizes the total number of recomputations. In particular we consider join graphs which correspond to models such as Siamese or Cross Modal Networks.

7.1.2. Sizing and Partitioning Strategies for Burst-Buffers to Reduce IO Contention

Burst-Buffers are high throughput and small size storage which are being used as an intermediate storage between the PFS (Parallel File System) and the computational nodes of modern HPC systems. They can allow to hinder to contention to the PFS, a shared resource whose read and write performance increase slower than processing power in HPC systems. A second usage is to accelerate data transfers and to hide the latency to the PFS. In this work, we concentrate on the first usage. We propose a model for Burst-Buffers and application transfers. We consider the problem of dimensioning and sharing the Burst-Buffers between several

applications. This dimensioning can be done either dynamically or statically. The dynamic allocation considers that any application can use any available portion of the Burst-Buffers. The static allocation considers that when a new application enters the system, it is assigned some portion of the Burst-Buffers, which cannot be used by the other applications until that application leaves the system and its data is purged from it. We show that the general sharing problem to guarantee fair performance for all applications is an NP-Complete problem. We propose a polynomial time algorithms for the special case of finding the optimal buffer size such that no application is slowed down due to PFS contention, both in the static and dynamic cases. Finally, we provide evaluations of our algorithms in realistic settings. We use those to discuss how to minimize the overhead of the static allocation of buffers compared to the dynamic allocation. More information on these results can be found in [9].

7.1.3. Efficient Ordering of Kernel Submission on GPUs

In distributed memory systems, it is paramount to develop strategies to overlap the data transfers between memory nodes with the computations in order to exploit their full potential. In [11], we consider the problem of determining the order of data transfers between two memory nodes for a set of independent tasks with the objective of minimizing the makespan. We prove that, with limited memory capacity, the problem of obtaining the optimal data transfer order is NP-complete. We propose several heuristics to determine this order and discuss the conditions that might be favorable to different heuristics. We analyze our heuristics on traces obtained by running two molecular chemistry kernels, namely, Hartree–Fock (HF) and Coupled Cluster Singles Doubles (CCSD), on 10 nodes of an HPC system. Our results show that some of our heuristics achieve significant overlap for moderate memory capacities and resulting in makespans that are very close to the lower bound.

Concurrent kernel execution is a relatively new feature in modern GPUs, which was designed to improve hardware utilization and the overall system throughput. However, the decision on the simultaneous execution of tasks is performed by the hardware with a leftover policy, that assigns as many resources as possible for one task and then assigns the remaining resources to the next task. This can lead to unreasonable use of resources. In [30], we tackle the problem of co-scheduling for GPUs with and without preemption, with the focus on determining the kernels submission order to reduce the number of preemptions and the kernels makespan, respectively. We propose a graph-based theoretical model to build preemptive and non-preemptive schedules. We show that the optimal preemptive makespan can be computed by solving a Linear Program in polynomial time, and we propose an algorithm based on this solution which minimizes the number of preemptions. We also propose an algorithm that transforms a preemptive solution of optimal makespan into a non-preemptive solution with the smallest possible preemption overhead. We show, however, that finding the minimal amount of preemptions among all preemptive solutions of optimal makespan is a NP-hard problem, and computing the optimal non-preemptive schedule is also NP-hard. In addition, we study the non-preemptive problem, without searching first for a good preemptive solution, and present a Mixed Integer Linear Program solution to this problem. We performed experiments on real-world GPU applications and our approach can achieve optimal makespan by preempting 6 to 9% of the tasks. Our non-preemptive approach, on the other side, obtains makespan within 2.5% of the optimal preemptive schedules, while previous approaches exceed the preemptive makespan by 5 to 12%.

7.1.4. Scheduling Tasks on Two Types of Resources

We consider the problem of scheduling task graphs on two types of unrelated resources, which arises in the context of task-based runtime systems on modern platforms containing CPUs and GPUs. In [10], we focus on an algorithm named HeteroPrio, which was originally introduced as an efficient heuristic for a particular application. HeteroPrio is an adaptation of the well known list scheduling algorithm, in which the tasks are picked by the resources in the order of their acceleration factor. This algorithm is augmented with a spoliation mechanism: a task assigned by the list algorithm can later on be reassigned to a different resource if it allows to finish this task earlier. We propose here the first theoretical analysis of the HeteroPrio algorithm in the presence of dependencies. More specifically, if the platform contains m and n processors of each type, we show that the worst-case approximation ratio of HeteroPrio is between $1 + \max(m/n, n/m)$ and $2 + \max(m/n, n/m)$.

Our proof structure allows to precisely identify the necessary conditions on the spoliation strategy to obtain such a guarantee. We also present an in-depth experimental analysis, comparing several such spoliation strategies, and comparing HeteroPrio with other algorithms from the literature. Although the worst case analysis shows the possibility of pathological behavior, HeteroPrio is able to produce, in very reasonable time, schedules of significantly better quality.

The evolution in the design of modern parallel platforms leads to revisit the scheduling jobs on distributed heterogeneous resources. We contribute to [31], a survey whose goal is to present the main existing algorithms, to classify them based on their underlying principles and to propose unified implementations to enable their fair comparison, both in terms of running time and quality of schedules, on a large set of common benchmarks that we made available for the community [27]. Beyond this comparison, our goal is also to understand the main difficulties that heterogeneity conveys and the shared principles that guide the design of efficient algorithms.

7.1.5. Data-Locality Aware Tasks Scheduling with Replicated Inputs

In [5], we consider the influence on data-locality of the replication of data files, as automatically performed by Distributed File Systems such as HDFS. Replication is known to have a crucial impact on data locality in addition to system fault tolerance. Indeed, intuitively, having more replicas of the same input file gives more opportunities for this task to be processed locally, i.e. without any input file transfer. Given the practical importance of this problem, a vast literature has been proposed to schedule tasks, based on a random placement of replicated input files. Our goal in this paper is to study the performance of these algorithms, both in terms of makespan minimization (minimize the completion time of the last task when non-local processing is forbidden) and communication minimization (minimize the number of non-local tasks when no idle time on resources is allowed). In the case of homogenous tasks, we are able to prove, using models based on "balls into bins" and "power of two choices" problems, that the well known good behavior of classical strategies can be theoretically grounded. Going further, we even establish that it is possible, using semi-matchings theory, to find the optimal solution in very small time. We also use known graph-orientation results to prove that this optimal solution is indeed near-perfect with strong probability. In the more general case of heterogeneous tasks, we propose heuristics solutions both in the clairvoyant and non-clairvoyant cases (i.e. task length is known in advance or not), and we evaluate them through simulations, using actual traces of a Hadoop cluster.

7.2. High performance solvers for large linear algebra problems

7.2.1. Deflation and preconditioning strategies for sequences of sampled stochastic elliptic equations

We are interested in the quantification of uncertainties in discretized elliptic partial differential equations with random coefficients. In sampling-based approaches, this relies on solving large numbers of symmetric positive definite linear systems with different matrices. In this work, we investigate recycling Krylov subspace strategies for the iterative solution of sequences of such systems. The linear systems are solved using deflated conjugate gradient (CG) methods, where the Krylov subspace is augmented with approximate eigenvectors of the previously sampled operator. These operators are sampled by Markov chain Monte Carlo, which leads to sequences of correlated matrices. First, the following aspects of eigenvector approximation, and their effect on deflation, are investigated: (i) projection technique, and (ii) restarting strategy of the eigen-search space. Our numerical experiments show that these aspects only impact convergence behaviors of deflated CG at the early stages of the sampling sequence. Second, unlike sequences with multiple right-hand sides and a constant operator, our experiments with multiple matrices show the necessity to orthogonalize the iterated residual of the linear system with respect to the deflation subspace, throughout the sampling sequence. Finally, we observe a synergistic effect of deflation and block-Jacobi (bJ) preconditioning. While the action of bJ preconditioners leaves a trail of isolated eigenvalues in the spectrum of the preconditioned operator, for 1D problems, the corresponding eigenvectors are well approximated by the recycling strategy. Then, up to a certain number of blocks, deflated CG methods with bJ preconditioners achieve similar convergence behaviors to those observed with CG when using algebraic multigrid (AMG) as a preconditioner.

This work, developed in the framework of the PhD thesis of Nicolas Venkovic in collaboration with P. Mycek (Cerfacs) and O. Le Maitre (CMAP, Ecole Polytechnique), will be presented at the next Copper Mountain conference on iterative methods.

7.2.2. Robust preconditionners via generalized eigenproblems for hybrid sparse linear solvers

The solution of large sparse linear systems is one of the most time consuming kernels in many numerical simulations. The domain decomposition community has developed many efficient and robust methods in the last decades. While many of these solvers fall into the abstract Schwarz (aS) framework, their robustness has originally been demonstrated on a case-by-case basis. In this work, we propose a bound for the condition number of all deflated aS methods provided that the coarse grid consists of the assembly of local components that contain the kernel of some local operators. We show that classical results from the literature on particular instances of aS methods can be retrieved from this bound. We then show that such a coarse grid correction can be explicitly obtained algebraically via generalized eigenproblems, leading to a condition number independent of the number of domains. This result can be readily applied to retrieve or improve the bounds previously obtained via generalized eigenproblems in the particular cases of Neumann-Neumann (NN), Additive Schwarz (AS) and optimized Robin but also generalizes them when applied with approximate local solvers. Interestingly, the proposed methodology turns out to be a comparison of the considered particular aS method with generalized versions of both NN and AS for tackling the lower and upper part of the spectrum, respectively. We furthermore show that the application of the considered grid corrections in an additive fashion is robust in the AS case although it is not robust for aS methods in general. In particular, the proposed framework allows for ensuring the robustness of the AS method applied on the Schur complement (AS/S), either with deflation or additively, and with the freedom of relying on an approximate local Schur complement. Numerical experiments illustrate these statements.

More information on these results can be found in [3]

7.2.3. Rank Revealing QR Methods for Sparse Block Low Rank Solvers

In the context of the ANR Sashimi project and the Phd of Esragul Korkmaz, we have investigated several compression methods of dense blocks appearing inside sparse matrix solvers to reduce the memory consumption, as well as the time to solution.

Solving linear equations of type Ax=b for large sparse systems frequently emerges in science and engineering applications, which creates the main bottleneck. In spite that the direct methods are costly in time and memory consumption, they are still the most robust way to solve these systems. Nowadays, increasing the amount of computational units for the supercomputers became trendy, while the memory available per core is reduced. Therefore, when solving these linear equations, memory reduction becomes as important as time reduction. While looking for the lowest possible compression rank, Singular Value Decomposition (SVD) gives the best result. It is however too costly as the whole factorization is computed to find the resulting rank. In this respect, rank revealing QR decomposition variants are less costly, but can introduce larger ranks. Among these variants, column pivoting or matrix rotation can be applied on the matrix A, such that the most important information in the matrix is gathered to the leftmost columns and the remaining unnecessary information can be omitted. For reducing the communication cost of the classical QR decomposition with column pivoting, blocking versions with randomization are suggested as an alternative solution to find the pivots. In these randomized variants, the matrix A is projected on a much lower dimensional matrix by using an independent and identically distributed Gaussian matrix so that the pivoting/rotational matrix can be computed on the lower dimensional matrix. In addition, to avoid unnecessary updates of the trailing matrix at each iteration, a truncated randomized method is suggested and shown to be more efficient for larger matrix sizes. Thanks to these methods, closer results to SVD can be obtained and the cost of compression can be reduced.

A comparison of all these methods in terms of complexity, numerical stability and performance have been presented at the national conference COMPAS'2019 [18], and at the international workshop SparseDay'2019 [19].

7.2.4. Accelerating Krylov linear solvers with agnostic lossy data compression

In the context of the Inria International Lab JLESC we have an ongoing collaboration with Argonne National Laboratory on the use of agnostic compression techniques to reduce the memory footprint of iterative linear solvers. Krylov methods are among the most efficient and widely used algorithms for the solution of large linear systems. Some of these methods can, however, have large memory requirements. Despite the fact that modern high-performance computing systems have more and more memory available, the memory used by applications remains a major concern when solving large scale problems. This is one of the reasons why interest in lossy data compression techniques has grown tremendously in the last two decades: it can reduce the amount of information that needs to be stored and communicated. Recently, it has also been shown that Krylov methods allow for some inexactness in the matrix-vector product that is typically required in each iteration. We showed that the loss of accuracy caused by compressing and decompressing the solution of the preconditioning step in the flexible generalized minimal residual method can be interpreted as an inexact matrix-vector product. This allowed us to find a bound on the maximum compression error in each iteration based on the theory of inexact Krylov methods. We performed a series of numerical experiment in order to validate our results. A number of "relaxed compression strategies" was also considered in order to achieve higher compression ratios.

The results of this joint effort will be presented to the next SIAM conférence on Parallel processing SIAM-PP'20.

7.2.5. Energy Analysis of a Solver Stack for Frequency-Domain Electromagnetics

High-performance computing aims at developing models and simulations for applications in numerous scientific fields. Yet, the energy consumption of these HPC facilities currently limits their size and performance, and consequently the size of the tackled problems. The complexity of the HPC software stacks and their various optimizations makes it difficult to finely understand the energy consumption of scientific applications. To highlight this difficulty on a concrete use-case, we perform in [8] an energy and power analysis of a software stack for the simulation of frequency-domain electromagnetic wave propagation. This solver stack combines a high order finite element discretization framework of the system of three-dimensional frequency-domain Maxwell equations with an algebraic hybrid iterative-direct sparse linear solver. This analysis is conducted on the KNL-based PRACE-PCP system. Our results illustrate the difficulty in predicting how to trade energy and runtime.

7.2.6. Exploiting Parameterized Task-graph in Sparse Direct Solvers

Task-based programming models have been widely studied in the context of dense linear algebra, but remains less studied for the more complex sparse solvers. In this talk [17], we have presented the use of two different programming models: Sequential Task Flow from StarPU, and Parameterized Task Graph from PaRSEC to parallelize the factorization step of the `PaStiX` sparse direct solver. We have presented how those programming models have been used to integrate more complex and finer parallelism to take into account new architectures with many computational units. Efficiency of such solutions on homogeneous and heterogeneous architectures with a spectrum of matrices from different applications have been shown. We also have presented how such solutions enable, without extra cost to the programmer, better performance on irregular computations such as in the block low-rank implementation of the solver.

7.2.7. Block Low-rank Algebraic Clustering for Sparse Direct Solvers

In these talks [20], [21], we adressed the Block Low-Rank (BLR) clustering problem, to cluster unknowns within separators appearing during the factorization of sparse matrices. We have shown that methods considering only intra-separators connectivity (i.e., k-way or recursive bissection) as well as methods managing only interaction between separators have some limitations. The new strategy we proposed consider interactions between a separator and its children to pre-select some interactions while reducing the number of off-diagonal blocks. We demonstrated how this method enhance the BLR strategies in the sparse direct supernodal solver PaStiX, and discuss how it can be extended to low-rank formats with more than one level of hierarchy.

7.2.8. Leveraging Task-Based Polar Decomposition Using PARSEC on Massively Parallel Systems

In paper [13], we describe how to leverage a task-based implementation of the polar decomposition on massively parallel systems using the PARSEC dynamic runtime system. Based on a formulation of the iterative QR Dynamically-Weighted Halley (QDWH) algorithm, our novel implementation reduces data traffic while exploiting high concurrency from the underlying hardware architecture. First, we replace the most time-consuming classical QR factorization phase with a new hierarchical variant, customized for the specific structure of the matrix during the QDWH iterations. The newly developed hierarchical QR for QDWH exploits not only the matrix structure, but also shortens the length of the critical path to maximize hardware occupancy. We then deploy PARSEC to seamlessly orchestrate, pipeline, and track the data dependencies of the various linear algebra building blocks involved during the iterative QDWH algorithm. PARSEC enables to overlap communications with computations thanks to its asynchronous scheduling of fine-grained computational tasks. It employs look-ahead techniques to further expose parallelism, while actively pursuing the critical path. In addition, we identify synergistic opportunities between the task-based QDWH algorithm and the PARSEC framework. We exploit them during the hierarchical QR factorization to enforce a locality-aware task execution. The latter feature permits to minimize the expensive inter-node communication, which represents one of the main bottlenecks for scaling up applications on challenging distributed-memory systems. We report numerical accuracy and performance results using well and ill-conditioned matrices. The benchmarking campaign reveals up to 2X performance speedup against the existing state-of-the-art implementation for the polar decomposition on 36,864 cores.

7.3. Parallel Low-Rank Linear System and Eigenvalue Solvers Using Tensor Decompositions

It is common to accelerate the boundary element method by compression techniques (FMM, $\mathcal{H}$-matrix/ACA) that enable a more accurate solution or a solution in higher frequency. In this article, we present a compression method based on a transformation of the linear system into the tensor-train format by the quantization technique. The method is applied to a scattering problem on a canonical object with a regular mesh and improves the performance obtained from existing methods. This method has been presented at the 22nd International Conference on the Computation of Electromagnetic Field, (COMPUMAG 2019) [12] and an extented version is accepted in IEEE TRANSACTIONS ON MAGNETICS.

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

7.4.1. StarPart Redesign

In the context of the french ICARUS project (FUI), which focuses the development of high-fidelity calculation tools for the design of hot engine parts (aeronautics & automotive), we are looking to develop new load-balancing algorithms to optimize the complex numerical simulations of our industrial and academic partners (Turbomeca, Siemens, Cerfacs, Onera, ...). Indeed, the efficient execution of large-scale coupled simulations on powerful computers is a real challenge, which requires revisiting traditional load-balancing algorithms based on graph partitioning. A thesis on this subject has already been conducted in the Inria HiePACS team in 2016, which has successfully developed a co-partitioning algorithm that balances the load of two coupled codes by taking into account the coupling interactions between these codes. This work was initially integrated into the StarPart platform. The necessary extension of our algorithms to parallel & distributed (increasingly dynamic) versions has led to a complete redesign of StarPart, which has been the focus of our efforts this year (as in the previous year). The StarPart framework provides the necessary building blocks to develop new graph algorithms in the context of HPC, such as those we are targeting. The strength of StarPart lies in the fact that it is a light runtime system applied to the issue of "graph computing". It provides a unified data model and a uniform programming interface that allows easy access to a dozen partitioning libraries, including Metis,

Scotch, Zoltan, etc. Thus, it is possible, for example, to load a mesh from an industrial test case provided by our partners (or an academic graph collection as DIMACS'10) and to easily compare the results for the different partitioners integrated in StarPart.

Alongside this work, we are beginning to work on the application of learning techniques to the problem of graph partitioning. Recent work on GCNs (Graph Convolutional Networks) is an interesting approach that we will explore.

7.5. Application Domains

7.5.1. Material physics

7.5.1.1. EigenSolver

The adaptive vibrational configuration interaction algorithm has been introduced as a new eigennvalues method for large dimension problem. It is based on the construction of nested bases for the discretization of the Hamiltonian operator according to a theoretical criterion that ensures the convergence of the method. It efficiently reduce the dimension of the set of basis functions used and then we are able solve vibrationnal eigenvalue problem up to the dimension 15 (7 atoms). Beyond this molecule size, two major issues appear. First, the size of the approximation domain increases exponentially with the number of atoms and the density of eigenvalues in the target area.
This year we have worked on two main areas. First of all, not all the eigenvalues that are calculated are determined by spectroscopy and therefore do not interest chemists. Only eigenvalues with an intensity are relevant. Also, we have set up a selection of interesting eigenvalues using the intensity operator. This requires calculating the scalar product between the smallest eigenvalues and the dipole moment applied to an eigenvector to evaluate its intensity. In addition, to get closer to the experimental values, we introduced the Coriolis operator into the Hamiltonian. A document is being written on these last two points showing that we can reach for a molecule 10 atoms the area of interest (i.e. more than 2 400 eigenvalues). Moreover, we continue to extend our shared memory parallelization to distributed memory using the message exchange paradigm to speedup the eigensolver time.

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

7.5.2.1. Numerical and parallel scalable hybrid solvers in large scale calculations

We have been working with the NACHOS team on the treatment of the system of three-dimensional frequency-domain (or time-harmonic) Maxwell equations using a high order hybridizable discontinuous Galerkin (HDG) approximation method combined to domain decomposition (DD) based hybrid iterative-direct parallel solution strategies. The proposed HDG method preserves the advantages of classical DG methods previously introduced for the time-domain Maxwell equations, in particular in terms of accuracy and flexibility with regards to the discretization of complex geometrical features, while keeping the computational efficiency at the level of the reference edge element based finite element formulation widely adopted for the considered PDE system. We study in details the computational performances of the resulting DD solvers in particular in terms of scalability metrics by considering both a model test problem and more realistic large-scale simulations performed on high performance computing systems consisting of networked multicore nodes. More information on these results can be found in [2].

In the context of a parallel plasma physics simulation code, we perform a qualitative performance study between two natural candidates for the parallel solution of 3D Poisson problems that are multigrid and domain decomposition. We selected one representative of each of these numerical techniques implemented in state of the art parallel packages and show that depending on the regime used in terms of number of unknowns per computing cores the best alternative in terms of time to solution varies. Those results show the interest of having both types of numerical solvers integrated in a simulation code that can be used in very different configurations in terms of problem sizes and parallel computing platforms. More information on these results will be shortly available in an Inria scientific report.

7.5.2.2. Efficient Parallel Solution of the 3D Stationary Boltzmann Transport Equation for Diffusive Problems

In the context of a collaboration with EDF-Lab with the Phd of Salli Moustafa, we present an efficient parallel method for the deterministic solution of the 3D stationary Boltzmann transport equation applied to diffusive problems such as nuclear core criticality computations. Based on standard MultiGroup-Sn-DD discretization schemes, our approach combines a highly efficient nested parallelization strategy with the PDSA parallel acceleration technique applied for the first time to 3D transport problems. These two key ingredients enable us to solve extremely large neutronic problems involving up to 10^{12} degrees of freedom in less than an hour using 64 super-computer nodes.

These contributions have been published in Journal of Computational Physics (JCP) [7].

7.5.2.3. Bridging the Gap Between $\mathcal{H}$-Matrices and Sparse Direct Methods for the Solution of Large Linear Systems

For the sake of numerical robustness in aeroacoustics simulations, the solution techniques based on the factorization of the matrix associated with the linear system are the methods of choice when affordable. In that respect, hierarchical methods based on low-rank compression have allowed a drastic reduction of the computational requirements for the solution of dense linear systems over the last two decades. For sparse linear systems, their application remains a challenge which has been studied by both the community of hierarchical matrices and the community of sparse matrices. On the one hand, the first step taken by the community of hierarchical matrices most often takes advantage of the sparsity of the problem through the use of nested dissection. While this approach benefits from the hierarchical structure, it is not, however, as efficient as sparse solvers regarding the exploitation of zeros and the structural separation of zeros from non-zeros. On the other hand, sparse factorization is organized so as to lead to a sequence of smaller dense operations, enticing sparse solvers to use this property and exploit compression techniques from hierarchical methods in order to reduce the computational cost of these elementary operations. Nonetheless, the globally hierarchical structure may be lost if the compression of hierarchical methods is used only locally on dense submatrices. In [1], we have reviewed the main techniques that have been employed by both those communities, trying to highlight their common properties and their respective limits with a special emphasis on studies that have aimed to bridge the gap between them. With these observations in mind, we have proposed a class of hierarchical algorithms based on the symbolic analysis of the structure of the factors of a sparse matrix. These algorithms rely on a symbolic information to cluster and construct a hierarchical structure coherent with the non-zero pattern of the matrix. Moreover, the resulting hierarchical matrix relies on low-rank compression for the reduction of the memory consumption of large submatrices as well as the time to solution of the solver. We have also compared multiple ordering techniques based on geometrical or topological properties. Finally, we have opened the discussion to a coupling between the Finite Element Method and the Boundary Element Method in a unified computational framework.

7.5.2.4. Design of a coupled MUMPS - $\mathcal{H}$-Matrix solver for FEM-BEM applications

In that approach, the FEM matrix is eliminated by computing a Schur complement using MUMPS. Given the size of the BEM matrix, this can not be done in one operation, so it is done block by block, and added in the $\mathcal{H}$-matrix, which is then factorized to complete the process. The overall process yields an interesting boost in performance when compared to the previously existing approach that coupled MUMPS with a classical dense solver. However, a full comparison with all the other existing methods must still be performed (full $\mathcal{H}$-matrix solver with [22] or without nested dissection, iterative approaches, etc.).

7.5.2.5. Metabarcoding

Distance Geometry Problem (DGP) and Nonlinear Mapping (NLM) are two well established questions: DGP is about finding a Euclidean realization of an incomplete set of distances in a Euclidean space, whereas Nonlinear Mapping is a weighted Least Square Scaling (LSS) method. We show how all these methods (LSS, NLM, DGP) can be assembled in a common framework, being each identified as an instance of an optimization problem with a choice of a weight matrix. In [6], we studied the continuity between the solutions (which are point clouds) when the weight matrix varies, and the compactness of the set of solutions (after centering). We finally studied a numerical example, showing that solving the optimization problem is far from being simple and that the numerical solution for a given procedure may be trapped in a local minimum.

We are involved in the ADT Gordon ((partners: TADAAM (coordinator), STORM, HIEPACS, PLEIADE). The objectives of this ADT is to scale our slolver stack on a PLEIADE dimensioning metabarcoding application (multidimensional scaling method). Our goal is to be able to handle a problem leading to a distance matrix around 100 million individuals. Our contribution concerns the the scalability of the multidimensional scaling method and more particulary the random projection methods to speed up the SVD solver. Experiments son PlaFRIM and MCIA CURTA plateforms have allowed us to show that the solver stack was able to solve efficiently a large problem up to 300,000 individuals in less than 10 minutes on 25 nodes. This has highlighted that for these problem sizes the management of I/O, inputs and outputs with the disks, becomes critical and dominates calculation times.

STORM Project-Team

7. New Results

7.1. Multi-Valued Expression Analysis for Collective Checking

Determining if a parallel program behaves as expected on any execution is challenging due to non-deterministic executions. Static analysis helps to detect all execution paths that can be executed concurrently by identifying multi-valued expressions, i.e. expressions evaluated differently among processes. This can be used to find collective errors in parallel programs. The PARallel COntrol flow Anomaly CHecker (PARCOACH) framework has been extended with a multi-valued expressions detection to find such errors [9]. The new analysis corrects the previous one and analyzes parallel applications using MPI, OpenMP, UPC and CUDA.

7.2. Hiding the latency of MPI communication

As developers spend significant effort performing manual latency optimizations, the goal of Van Man Nguyen Ph.D Thesis is to automatically provide maximal communication overlap for MPI communication (collective, point-to-point and RMA Put/Get operations). A method that moves operations and their completion (e.g. Isend/Wait) as far apart as possible in the program while preserving memory consistency is under development.

7.3. Performance monitoring and Steering Framework

Two frameworks were developed within the context of the project H2020 EXA2PRO to offer performance monitoring and steering APIs into the StarPU runtime system, to be targeted by external tools.

The performance monitoring framework enables StarPU to export performance counters in a generic, safe, extensible way, to give external tools access to internal metrics and statistics, such as the peak number of tasks in the dependence waiting queue, the cumulated execution time by worker thread, and the number of ready tasks of a given kind waiting for an execution slot.

The performance steering framework defines a set of runtime-actionable knobs that can be used to steer the execution of an application on top of StarPU from an external tool, with similar properties of genericity, safety and extensibility as the performance monitoring framework.

7.4. Heterogeneous task scheduling

Taking advantage of heterogeneous systems requires to carefully choose which tasks should be accelerated. Simple heuristics allow to get fairly good performance, but do not have approximation ratio that would provide performance guarantees. The ROMA Inria team designed advanced heuristics which do have approximation ratios. We have implemented one within StarPU and indeed improved the performance over existing heuristics. The implementation required to revamp part of the StarPU toolkit dedicated to writing scheduling heuristics.

7.5. Task scheduling with memory constraints

When dealing with larger and larger datasets processed by task-based applications, the amount of system memory may become too small to fit the working set, depending on the task scheduling order. The ROMA Inria team proposed a heuristic to introduce additional dependencies to the task graph enough so that any scheduling order will meet the memory constraint, while avoiding to extend the critical path length. On the other hand, banker algorithms allow to achieve this online, within the task scheduler, but do not have an overall view of the task graph, and may thus severely increase the critical path. We have thus started to design a collaboration between visionary heuristics which take a global but coarse view of the task graph, and online heuristics which have a local but precise view of the task graph.

7.6. Leveraging compiler analysis for task scheduling

Polyhedral analysis of task graph submission loops allow to get at compilation time a representation of the task graph, and perform insightful analyses thanks to the obtained overview of the whole task graph. Task scheduling heuristics, on the other hand, usually keep only a limited view over the task graph, to avoid prohibitive algorithmic costs. We have started to collaborate with the CASH Inria team to transfer some of the insights of the compiler to the compiler. We have notably made the compiler automatically compute a cut of the task graph below which the availability parallelism is lower than the capacities of the target hardware. The scheduler can then at that point switch between a heuristic which privileges task acceleration, and a heuristic which privileges the critical path. Only preliminary results have been obtained so far.

7.7. Failure Tolerance for StarPU

Since supercomputers keep growing in terms of core numbers, the reliability decreases the same way. The project H2020 EXA2PRO aims to propose solutions for the failure tolerance problem, including StarPU. While exploring decades of research about the resilience techniques, we have identified properties in our runtime's paradigm that can be exploited in order to propose a solution with lower overhead than the generic existing ones. An implementation of a solution is currently being developed for evaluation, with an interface that can be easily plugged into StarPU.

7.8. Static and Dynamic Adaptation of Task parallelism

This work is the result of Pierre Huchant PhD thesis, and the objectives are to adapt statically and dynamically OpenCL tasks. The adaptation consists in splitting tasks automatically into multiple sub-tasks, taking into account the heterogeneity of the architecture (sub-tasks are specifically created for one processing unit), the load imbalance within a parallel OpenCL, between the different iterations in space, and if the task graph is repeatedly executed, between the iterations in time, and it takes into account the time of the communications generated by splitting the tasks [2]. The method is able to cope with sequential task graphs (tasks are parallel themselves but scheduled sequentially) and deals with tasks manipulating complex data structures as shown on an N-body particle simulation mini-app.

7.9. AFF3CT

The AFF3CT library, developed jointly between IMS and the STORM team, which aims to model error correcting codes for numerical communications has been further improved in different ways. Additional new algorithms have been designed and evaluated within the AFF3CT library [6], [7], and a new approach for generating and exploring automatically high performance error correction codes from matrix description is an on-going work. Besides, an on-going work is on the automatic parallelization of the tasks describing the simulation of a whole chain of signal transmission. In order to be able to make accessible the simulation results obtained with AFF3CT and to be able to replay easily the same simulations, a web interface allowing users to browse through the results, and the simulation setup for a large range of inputs has been designed. This graphical interface is currently in use at IMS by other researchers. Finally, a call for the creation of a consortium on AFF3CT is available on the web page of AFF3CT https://aff3ct.github.io/.

7.10. Matlab API for AFF3CT

As part of the AFF3CT development action, an API compatible with the Matlab mathematical software was designed on top of the AFF3CT fast forward error correction toolbox to allow the library to be used in a high-level manner, directly from the Matlab environment. Due to the relatively large number of classes exported by AFF3CT, an automatized process was designed to let AFF3CT classes be wrapped adequately for Matlab's MEX interface to external libraries.

7.11. InKS framework

The InKS framework was developed by Ksander Ejjaaouani as part of his Ph.D Thesis co-supervised by the university of Strasbourg, the Maison de la Simulation and the STORM team. It enables separating algorithms of time loop based scientific simulations into platform independent algorithms and platform specific optimisation files, thus enforcing separation of concerns between the algorithmic design process on one side, and the optimization process of specifying platform-dependent aspects such as operation ordering and memory placement on the other side.

7.12. HPC Big Data Convergence

A Java interface for StarPU has been implemented and allows to execute Map Reduce applications on top of StarPU. We have made some preliminary experiments on Cornac, a big data application for visualising huge graphs.

7.13. Hierarchical Tasks

We are continuing our work, on the partitioning of the data and the prioritization of task graphs to optimize the use of resources of a machine. Hierarchical tasks allow a better control over the submission of an application's task graph by allowing to dynamically adapt the granularity of the calculations. In the ANR project Solharis, hierarchical tasks are proposed as a solution for a better management of dynamic task graphs. We have continued to explore new solutions for maximizing the performance of hierarchical tasks.

7.14. ADT Gordon

In collaboration with the HIEPACS and TADAAM Inria teams, we are strengthening the relations between the Chameleon linear algebra library from HIEPACS, our StarPU runtime scheduler, and the NewMadeleine high-performance communication library from TADAAM. More precisely, we have improved the interoperation between StarPU and NewMadeleine, to more carefully decide when NewMadeleine should proceed with communications. We have then introduced the notion of dynamic collective operations, which opportunistically introduce communication trees to balance the communication load.

7.15. StarPU in Julia

Julia is a modern language for parallelism and simulation that aims to ease the effort for developing high performance codes. In this context, we carry on the development of a StarPU binding inside Julia. It is possible to launch StarPU tasks inside Julia, either given as libraries, or described in Julia directly. The tasks described in Julia are compiled into either source OpenMP code or CUDA code. We improved further the support of StarPU in Julia, but this is still a work in progress.

7.16. Simulation of OpenMP task based programs

A simulator for OpenMP task based programs is being designed as part of Inria's IPL HAC-Specis project, and the Ph.D Thesis of Idriss Daoudi. The goal is to extend the SimGrid HPC simulation framework with the ability to simulate OpenMP applications.

7.17. OpenMP enabled version of Chameleon

An OpenMP enabled version of the Chameleon linear algebra library was designed within the context of European Project PRACE-5IP. This enables the Chameleon library to be available on any platform for which an OpenMP compiler is installed, without any requirement for third party task-based runtime systems. A preliminary support of the OpenMP port for heterogeneous, accelerated platform was also designed as part of this work.

TADAAM Project-Team

7. New Results

7.1. Management of heterogeneous and non-volatile memories in HPC

The emergence of non-volatile memory that may be used either as fast storage or slow high-capacity memory brings many opportunities for application developers.

We studied the impact of those new technologies on the allocation of resources in HPC platforms. We showed that co-scheduling HPC applications will possibly different needs in term of storage and memories brings constraints of the way non-volatile memory should be exposed by the hardware and operating system to bring both flexibility and performance. [21]

We also worked with Lawrence Livermore National Lab to propose an API to help application choose between the different kinds of available memory (high-bandwidth (HBM), normal (DDR), slow (non-volatile)). We exposed several useful criteria for selecting target memories as well as ways to rank them. [22]

7.2. Modeling and Visualizing Many-core HPC Platforms

As the number of cores keeps increasing inside processors, new kinds of hierarchy are added to organize and interconnect them. We worked with Intel to leverage new groups of cores such as *Dies* in newest Xeon Advanced Performance models. We also designed ways to clarify the modeling and visualisation of those many cores by factorizing identical parts of the platforms.

7.3. Co-scheduling HPC workloads on cache-partitioned CMP platforms

Co-scheduling techniques are used to improve the throughput of applications on chip multiprocessors (CMP), but sharing resources often generates critical interferences.

In collaboration with ENS Lyon and Georgia Tech, we looked at the interferences in the last level of cache (LLC) and use the *Cache Allocation Technology* (CAT) recently provided by Intel to partition the LLC and give each co-scheduled application their own cache area.

We considered m iterative HPC applications running concurrently and answer the following questions: (i) how to precisely model the behavior of these applications on the cache partitioned platform? and (ii) how many cores and cache fractions should be assigned to each application to maximize the platform efficiency? Here, platform efficiency is defined as maximizing the performance either globally, or as guaranteeing a fixed ratio of iterations per second for each application. Through extensive experiments using CAT, we demonstrated the impact of cache partitioning when multiple HPC application are co-scheduled onto CMP platforms. [2]

7.4. Modeling High-throughput Applications for in situ Analytics

In this work [3], we proposed to model HPC applications in the framework of in situ analytics. Typically, an HPC application is composed of a simulation tasks (data and compute intensive), and a set of analysis tasks that post-process the data. Currently, the performance of the I/O system in HPC platform prohibits the storage of all simulation data to process analysis post-mortem. Hence, in situ framework proposes to treat the data "on the fly", directly where it is produced. Hence, it leverages the amount of data to store as we only keep the result of analytics phase. However, simulation and analysis have to be scheduled in parallel and compete for shared resources. It generates resource conflicts and can lead to severe performance degradation for the simulation.

Hence, we proposed to model both platform (number of nodes and cores, memory, etc) and application (profile of each tasks) in order to optimize the execution of such applications. We propose a resource partitioning model that affects computational resources to the different tasks, as so as a scheduling of those tasks in order to maximize resource usage and minimize total application makespan. Tasks are assumed to be fully parallel to solve the partitioning problem.

We evaluated different scheduling heuristics combined to the resource partitioning model and show important features that influence in situ analytics performance.

This work is done in collaboration with Bruno RAFFIN from Inria team DATAMOVE of Inria Grenoble.

7.5. Modeling Non-Uniform Memory Access and Heterogeneous Memories on Large Compute Nodes with the Cache-Aware Roofline Model

The trend of increasing the number of cores on-chip is enlarging the gap between compute power and memory performance. This issue leads to design systems with heterogeneous memories, creating new challenges for data locality. Before the release of those memory architectures, the Cache-Aware Roofline Model [43] (CARM) offered an insightful model and methodology to improve application performance with knowledge of the cache memory subsystem.

With the help of the HWLOC library, we are able to leverage the machine topology to extend the CARM for modeling NUMA and heterogeneous memory systems, by evaluating the memory bandwidths between all combinations of cores and NUMA nodes. The new Locality Aware Roofline Model [6] (LARM) scopes most contemporary types of large compute nodes and characterizes three bottlenecks typical of those systems, namely contention, congestion and remote access. We also designed a hybrid memory bandwidth model to better estimate the roof when heterogeneous memories are involved or when read and write bandwidths differ.

We also developed an hybrid bandwidth model that combines the performance of different memories and their respective read/write bandwidth with the application memory access pattern to predict the performance of these accesses on heterogeneous memory platforms.

This work has been achieved in collaboration with the authors of the CARM from University of Lisbon.

7.6. Statistical Learning for Task and Data Placement in NUMA Architecture

Achieving high performance for multi-threaded application requires both a careful placement of threads on computing units and a thorough allocation of data in memory. Finding such a placement is a hard problem to solve, because performance depends on complex interactions in several layers of the memory hierarchy.

We proposed a black-box approach to decide if an application execution time can be impacted by the placement of its threads and data, and in such a case, to choose the best placement strategy to adopt [18]. We show that it is possible to reach near-optimal placement policy selection by looking at hardware performance counters, and at counters obtained from application instrumentation. Furthermore, solutions work across several recent processor architectures (from Haswell to Skylake), across several applications, and decisions can be taken with a single run of low overhead profiling.

This work has been achieved in collaboration with Thomas ROPARS from University of Grenoble.

7.7. On-the-fly scheduling vs. reservation-based scheduling for unpredictable workflows

Scientific insights in the coming decade will clearly depend on the effective processing of large datasets generated by dynamic heterogeneous applications typical of workflows in large data centers or of emerging fields like neuroscience. In this work [8], we show how these big data workflows have a unique set of characteristics that pose challenges for leveraging HPC methodologies, particularly in scheduling. Our findings indicate that execution times for these workflows are highly unpredictable and are not correlated with the size of the dataset involved or the precise functions used in the analysis. We characterize this inherent variability and sketch the need for new scheduling approaches by quantifying significant gaps in achievable performance. Through simulations, we show how on-the-fly scheduling approaches can deliver benefits in both system-level and user-level performance measures. On average, we find improvements of up to 35% in system utilization and up to 45% in average stretch of the applications, illustrating the potential of increasing performance through new scheduling approaches.

7.8. Scheduling strategies for stochastic jobs

Following the observations of made in 7.7 , we studied stochastic jobs (coming from neuroscience applications) which we want to schedule on a reservation-based platform (e.g. cloud, HPC).

The execution time of jobs is modeled using a (known) probability distribution. The platform to run the job is reservation-based, meaning that the user has to request fixed-length time slots for its job to be executed. The aim of this project is to study efficient strategies of reservation for an user given the cost associated to the machine. These reservations are all paid until a job is finally executed.

As a first step we derived efficient strategies without any additional asumptions [15]. This allowed us to set up properly the problem. These strategies were general enough that they could take as input any probability distributions, and performed better than any more natural strategies. Then we extended our strategies by including checkpoint/restart to well-chosen reservations in order to avoid wasting the benefits of work during underestimated reservations [35]. We were able to develop a fully polynomial-time approximation for continuous distribution of job execution time whose performance we then experimentally studied.

The final works of this project focused on the case without checkpointing: we studied experimentally how the strategies developed in [15] would perform in a parallel setup and showed that they improve both system utilization and job response time. Finally we started to study the robustness of such solutions when the job distributions were not perfectly known [19] and observed that the performance were still correct even with a very low quantity of information.

7.9. Online Prediction of Network Utilization

Stealing network bandwidth helps a variety of HPC runtimes and services to run additional operations in the background without negatively affecting the applications. A key ingredient to make this possible is an accurate prediction of the future network utilization, enabling the runtime to plan the background op- erations in advance, such as to avoid competing with the application for network bandwidth. In this work [23], we have proposed a portable deep learning predictor that only uses the information available through MPI introspection to construct a recurrent sequence-to-sequence neural network capable of forecasting network utilization. We leverage the fact that most HPC applications exhibit periodic behaviors to enable predictions far into the future (at least the length of a period). Our online approach does not have an initial training phase, it continuously improves itself during application execution without incurring significant computational overhead. Experimental results show better accuracy and lower computational overhead compared with the state-of-the-art on two representative applications.

7.10. An Introspection Monitoring Library

In this work [36] we have described how to improve communication time of MPI parallel applications with the use of a library that enables to monitor MPI applications and allows for introspection (the program itself can query the state of the monitoring system). Based on previous work, this library is able to see how collective communications are decomposed into point-to-point messages. It also features monitoring sessions that allow suspending and restarting the monitoring, limiting it to specific portions of the code. Experiments show that the monitoring overhead is very small and that the proposed features allow for dynamic and efficient rank reordering enabling up to 2-time reduction of communication parts of some program.

7.11. Tag matching in constant time

Tag matching is the operation, inside an MPI library, of pairing a packet arriving from the network, with its corresponding receive request posted by the user. This operation is not straightforward given that matching criterions are the communicator, the source of the message, a user-supplied tag, and since there are wildcards for tag and source. State of the art algorithms are linear with the number of pending packets and requests, or don't support wildcards.

We proposed [17] an algorithm that is able perform the matching operation in constant time, in all cases, even with wildcard requests. We implemented the algorithm in our NEWMADELEINE communication library, and demonstrated it actually improves performance of Cholesky factorization with CHAMELEON running on top of STARPU.

7.12. Dynamic broadcasts in StarPU/NewMadeleine

We worked on the improvement of broadcast performance in STARPU runtime with NEWMADELEINE. Although STARPU supports MPI, its distributed and asynchronous model to schedule tasks makes it impossible to use MPI optimized routines, such as `MPI_Bcast`. Indeed these functions need that all nodes participating in the collective are synchronized and know each others, which makes it unusable in practice for STARPU.

We proposed [42], a dynamic broadcast algorithm that runs without synchronization among participants, and where only the root node needs to know the others. Recipient don't even have to know whether the message will arrive as a plain send/receive or through a dynamic broadcast, which allows for a seamless integration in STARPU. We implemented the algorithm in our NEWMADELEINE communication library, leveraging its event-based paradigm and background progression of communications. Preliminary experiments using Cholesky factorization from the CHAMELEON library show a sensible performance improvement.

7.13. Task based asynchronous MPI collectives optimisation

Asynchronous collectives are more complex than plain non-blocking point-to-point communications. They need specific mechanisms for progression. Task based progression is a good way to improve the performance of applications with overlap.

We worked on a benchmarking tool [41] measuring specific collective overlapping, taking into account time shift between different nodes. Using this tool, we were able to experiment with different task execution policies in the NEWMADELEINE communication library.

We propose a progression policy consisting of a dedicated a core for progression tasks; modern processors have more and more cores, so it is profitable on that kind of processors. The only function of this core is to progress communications, so we use a particularly aggressive algorithm for this progression.

7.14. Dynamic placement of progress thread for overlapping MPI non-blocking collectives on manycore processor

To amortize the cost of MPI collective operations, non-blocking collectives have been proposed so as to allow communications to be overlapped with computation. Unfortunately, collective communications are more CPU-hungry than point-to-point communications and running them in a communication thread on a single dedicated CPU core makes them slow. On the other hand, running collective communications on the application cores leads to no overlap. To address these issues, we proposed [5] an algorithm for tree-based collective operations that splits the tree between communication cores and application cores. To get the best of both worlds, the algorithm runs the short but heavy part of the tree on application cores, and the long but narrow part of the tree on one or several communication cores, so as to get a trade-off between overlap and absolute performance. We provided a model to study and predict its behavior and to tune its parameters. We implemented it in the MPC framework, which is a thread-based MPI implementation. We have run benchmarks on manycore processors such as the KNL and Skylake and got good results both in terms of performance and overlap.

7.15. Dynamic placement of Hybrid MPI +X coupled applications

We continued our collaboration with CERFACS in order to propose the HIPPO software that addresses the issue of dynamic placement of computing kernels that feature each their own placement/mapping/binding policy of MPI processes and OpenMP threads. In such a case, enforcing a global placement policy for the whole application composed of several such kernels may be detrimental to the overall performance. HIPPO (based on our HSPLIT library and the HWLOC software) is able to make the selection of the relevent resource

on which some master MPI processes are going to execute and spawn OpenMP parallel sections while the remaining MPI processes are put in a "quiescence" state. HIPPO is currently at the prototype stage and the interface and the set of provided functionnalities need some refinement, however, preliminary results are very encouraging, especially on climate modelling applications from Météo France.

7.16. Scheduling on Two Unbounded Resources with Communication Costs

Heterogeneous computing systems are popular and powerful platforms, containing several heterogeneous computing elements (e.g. CPU+GPU). In [13], we consider a platform with two types of machines , each containing an unbounded number of elements. We want to execute an application represented as a Directed Acyclic Graph (DAG) on this platform. Each task of the application has two possible execution times, depending on the type of machine it is executed on. In addition we consider a cost to transfer data from one platform to the other between successive tasks. We aim at minimizing the execution time of the DAG (also called makespan). We show that the problem is NP-complete for graphs of depth at least three but polynomial for graphs of depth at most two. In addition, we provide polynomial-time algorithms for some usual classes of graphs (trees, series-parallel graphs).

7.17. H-Revolve: A Framework for Adjoint Computation on Synchrone Hierarchical Platforms

In this work [38], we study the problem of checkpointing strategies for adjoint computation on synchrone hierarchical platforms. Specifically we consider computational platforms with several levels of storage with different writing and reading costs. When reversing a large adjoint chain, choosing which data to checkpoint and where is a critical decision for the overall performance of the computation. We introduce H-Revolve, an optimal algorithm for this problem. We make it available in a public Python library along with the implementation of several state-of-the-art algorithms for the variant of the problem with two levels of storage. We provide a detailed description of how one can use this library in an adjoint computation software in the field of automatic differentiation or backpropagation. Finally, we evaluate the performance of H-Revolve and other checkpointing heuristics though an extensive campaign of simulation.

7.18. Sizing and Partitioning Strategies for Burst-Buffers to Reduce IO Contention

Burst-Buffers are high throughput and small size storage which are being used as an intermediate storage between the PFS (Parallel File System) and the computational nodes of modern HPC systems. They can allow to hinder to contention to the PFS, a shared resource whose read and write performance increase slower than processing power in HPC systems. A second usage is to accelerate data transfers and to hide the latency to the PFS. In this work [14], we concentrate on the first usage. We propose a model for Burst-Buffers and application transfers. We consider the problem of dimensioning and sharing the Burst-Buffers between several applications. This dimensioning can be done either dynamically or statically. The dynamic allocation considers that any application can use any available portion of the Burst-Buffers. The static allocation considers that when a new application enters the system, it is assigned some portion of the Burst-Buffers, which cannot be used by the other applications until that application leaves the system and its data is purged from it. We show that the general sharing problem to guarantee fair performance for all applications is an NP-Complete problem. We propose a polynomial time algorithms for the special case of finding the optimal buffer size such that no application is slowed down due to PFS contention, both in the static and dynamic cases. Finally, we provide evaluations of our algorithms in realistic settings. We use those to discuss how to minimize the overhead of the static allocation of buffers compared to the dynamic allocation.

7.19. Optimal Memory-aware Backpropagation of Deep Join Networks

Deep Learning training memory needs can prevent the user to consider large models and large batch sizes. In our work [4] (extended version [34]), we propose to use techniques from memory-aware scheduling and Automatic Differentiation (AD) to execute a backpropagation graph with a bounded memory requirement at the cost of extra recomputations. The case of a single homogeneous chain, i.e. the case of a network whose all stages are identical and form a chain, is well understood and optimal solutions have been proposed in the AD literature. The networks encountered in practice in the context of Deep Learning are much more diverse, both in terms of shape and heterogeneity. In this work, we define the class of backpropagation graphs, and extend those on which one can compute in polynomial time a solution that minimizes the total number of recomputations. In particular we consider join graphs which correspond to models such as Siamese or Cross Modal Networks.

7.20. Optimal checkpointing for heterogeneous chains: how to train deep neural networks with limited memory

This work [33] introduces a new activation checkpointing method which allows to significantly decrease memory usage when training Deep Neural Networks with the back-propagation algorithm. Similarly to checkpoint-ing techniques coming from the literature on Automatic Differentiation, it consists in dynamically selecting the forward activations that are saved during the training phase, and then automatically recomputing missing activations from those previously recorded. We propose an original computation model that combines two types of activation savings: either only storing the layer inputs, or recording the complete history of operations that produced the outputs (this uses more memory, but requires fewer recomputations in the backward phase), and we provide an algorithm to compute the optimal computation sequence for this model. This paper also describes a PyTorch implementation that processes the entire chain, dealing with any sequential DNN whose internal layers may be arbitrarily complex and automatically executing it according to the optimal checkpointing strategy computed given a memory limit. Through extensive experiments, we show that our implementation consistently outperforms existing checkpoint-ing approaches for a large class of networks, image sizes and batch sizes.

7.21. I/O scheduling strategy for HPC applications

With the ever-growing need of data in HPC applications, the congestion at the I/O level becomes critical in supercomputers. Architectural enhancement such as burst buffers and pre-fetching are added to machines, but are not sufficient to prevent congestion. Recent online I/O scheduling strategies have been put in place, but they add an additional congestion point and overheads in the computation of applications.

In this project, we studied application pattern (such as periodicity), in order to develop efficient scheduling strategies [7], [32] for their I/O transfers.

7.22. A New Framework for Evaluating Straggler Detection Mechanisms in MapReduce

In this work [10] we present a new framework for evaluating straggler detection mechanisms in MapReduce. We then show how to use it efficiently.

7.23. Clarification of the MPI semantics

In the framework of the MPI Forum, we have been involved in several active working groups, in particular the “Terms and Conventions” Working Group. The work carried out in this group has lead to a timely study and proposed clarifications, revisions, and enhancements to the Message Passing Interface’s (MPI’s) Semantic Terms and Conventions. To enhance MPI, a clearer understanding of the meaning of the key terminology has proven essential, and, surprisingly, important concepts remain underspecified, ambiguous and, in some cases,

inconsistent and/or conflicting despite 26 years of standardization. This work [16] addresses these concerns comprehensively and usefully informs MPI developers, implementors, those teaching and learning MPI, and power users alike about key aspects of existing conventions, syntax, and semantics. This work will also be a useful driver for great clarity in current and future standardization and implementation efforts for MPI.

7.24. Adaptive Request Scheduling for the I/O Forwarding Layer using Reinforcement Learning

I/O optimization techniques such as request scheduling can improve performance mainly for the access patterns they target, or they depend on the precise tune of parameters. In this work [40], we propose an approach to adapt the I/O forwarding layer of HPC systems to the application access patterns by tuning a request scheduler. Our case study is the TWINS scheduling algorithm, where performance improvements depend on the time window parameter, which depends on the current workload. Our approach uses a reinforcement learning technique — contextual bandits — to make the system capable of learning the best parameter value to each access pattern during its execution, without a previous training phase. We evaluate our proposal and demonstrate it can achieve a precision of 88% on the parameter selection in the first hundreds of observations of an access pattern. After having observed an access pattern for a few minutes (not necessarily contiguously), we demonstrate that the system will be able to optimize its performance for the rest of the life of the system (years).

Auctus Team

7. New Results

7.1. Posture and motion capture by smart textile

The objective of Postex is to design an intelligent textile jacket, without the use of additional sensors, to determine an operator's posture.

Since 2017, we offer an innovative solution based on the electrical properties of a conductive stretch fabric that is used in the manufacture of an intelligent garment. We use Electrical Impedance Tomography (EIT) to identify fabric deformations. A neural network is used to correlate the different postures and movements measured using a reference device with the electric field measured in the intelligent textile.

By 2018, we had successfully identified the movements of an elbow and then filed a patent under the number FR1860192. In 2019, we identified the shoulder and worked on the design of the fabric parts. In order to valorize the technology, the Touch Sensity startup was created at the end of 2019.

7.2. Set-based evaluation of robot capabilities

Set-based approaches allow to model serial mechanisms with varying levels of geometric uncertainties. The Kinematic Chain Appropriate Design Library (KCADL) has been created for the purpose of modelling imprecise serial kinematic chains and provides numerous certified methods, implemented using the IBEX interval analysis library, for analyzing the capabilities of these modelled mechanisms. The KCADL software provides a set of public routines to build arbitrary serial mechanisms by incrementally adding rigid-body segments with associated parent-child uncertainties. Efficient Forward Kinematic (FK) and Inverse Kinematic (IK) solvers have been formulated and integrated into the software. These solvers are fully compatible with set-based inputs and are capable of handling sets of poses or sets of joint configurations. In addition to the FK and IK solvers, analysis routines which are applicable to imprecise kinematic chains with set-based inputs are also implemented in the software (e.g., evaluating the mechanism's: force/velocity/acceleration capabilities, precision). These routines provide offline analysis and design tools as well as online real-time capable tools for reliably evaluating current and future capabilities.

7.3. Redundancy tube

A set-based approach for modelling the human upper-limb and its complex articular constraints as a 7-degree-of-freedom (dof) constrained imprecise kinematic chain is formulated and implemented in the AUCTUS-RT software. This software allows to easily customize the geometric parameters and articular constraints. This permits to adapt the upper-limb model to each unique human subject and may also be used to model sets of human subjects. Various visualization tools are available for Python and Matlab to aid in the selection of appropriate articular constraints. When given a task with m redundant dofs and the desired workspace resolution, the software is capable of computing certified inside and outside regions of the (m+1)-dimensional redundant workspace associated with the task and upper-limb model. When a temporal dimension is added to the task description, the redundant workspace varies over time and produces a tube of redundant motions, which we refer to as the redundancy tube. The software accepts spatial-temporal task descriptions and allows to compute the full redundancy tube or slices of the tube. Furthermore, the software allows to model individual and/or sets of trajectories to describe tasks exactly or with varying level of uncertainties. Much effort has been put into improving the efficiency of the redundancy tube evaluations and parallel computing, both locally and non-locally via PlaFRIM, using OpenMP is supported. The AUCTUS-RT software is currently being used for the ongoing study of human motor-variabilities for the AUCTUS Mover project.

Related publications: [16]

7.4. Mover project

The Mover project is an experiment-based project to evaluate and study the links between human motor-variability, expertise, and fatigue associated with repetitive tasks. A wireloop game serves as the experimental task where a human subject is tasked with moving a metal ring along a fixed conductive wire while trying to maximize a score which is a function of the task time and the number of ring-wire contacts. To more easily study motor-variabilities, the wireloop ring is actively constrained by a collaborative robot (the Frank Emika Panda), allowing to isolate desired task variabilities and easily modify the task (e.g., through applied disturbances, changes in robot stiffness, task guidance). A preliminary version of the experiment is currently being developed to test all aspects of the project (i.e., experimental protocol, robot control, motion capture, human modelling, and redundancy tube evaluation). All experimental aspects of the project will be finalized before March 2020.

7.5. Interactions with a chatbot

In the context of the CIFRE Orange PhD work by Nicolas Simonazzi under the supervision of Jean-Marc Salotti and with the objective of analyzing and identifying emotions during interactions with a chatbot, a first experiment was conducted. It involved a user, the use of a smartphone, viewing videos and asking questions about the content of the video and the feeling of the user just after the answer. The collected data were numerous: the accuracy of the answers to the questions, the emotional feeling (choice of emoticons by the user) as well as the real-time measurements of the accelerometer of the smartphone. An analysis of the data was carried out with Russell's relatively simple emotional model as an explanatory framework based on two variables, the positive or negative valence of the emotion, and the degree of excitement. The experimental results showed that there was a slight correlation between the valence indicated by the user, the accuracy of the answers to the questions and the accelerations of the smartphone. However, it was hoped that the videos would have an impact on the valence, because their content had an intrinsic valence, but it proved impossible to find a correlation with the valence indicated by the user, probably due to a lack of the user engagement and also because of the focus on the questions that followed and the accuracy of the answers. A new experimental protocol is currently being studied with a priori more impactful videos (likely to produce an emotion with a greater degree of excitement).

Related publications:[10], [19]

7.6. Prediction of human error in robotics

The INRS provides a database of accidents at work, from which we can extract those concerning robotics. Many accidents are due to a deterioration of situational awareness. However, as there are many different causes and human factors are not well understood, it is very difficult for experts to provide probabilistic risk assessments. We proposed to simplify the problem by classifying the accidents according to the main demons that degrade the consciousness of the situation (Endsley model) and to use a Bayesian approach with the Noisy-Or nodes. We had already tried such an interpretation in the field of aeronautics. We propose to extend it to the field of robotics. Even if the approach remains empirical and approximate, it is possible to infer general probabilities of risk of human error leading to accidents and to deduce actions to reduce risks.

Related publications: [18]

7.7. Securing industrial operators with collaborative robots: simulation and experimental validation for a carpentry task

In this work, a robotic assistance strategy is developed to improve the safety in an artisanal task that involves a strong interaction between a machine-tool and an operator. Wood milling is chosen as a pilot task due to its importance in carpentry and its accidentogenic aspect.

In order to analyze the wood milling task, a wood shaping training was conducted in collaboration with a carpentry learning institute which allowed to collect information related to the task (perceived effort, position of the operator, accident circumstances).

To analyze the human-machine interaction, a formalization of the problem as a dynamic exchange of spatial forces inspired by the grasping theory has been performed. This theory presents structural similarities with the studied task. Based on this formalization, a behavior simulator of the system "wood + human + tool" has been developed.

To propose a credible and a realistic assistance solution, accidentogenic situations are simulated (see Woobotsim). Based on the observation made with these simulations, the use of a collaborative robot to secure wood instability cases has been explored and validated by an experiment. An operational space damping behaviour appears to be the most appropriate solution to improve safety in the studied cases.

The experiment was designed to reproduce two cases of instability during a carpentry milling task based on the entry and exit of the tool into and out of a wood node. For safety reasons, the experiment is performed on a safe but tangible simulation of the task. We then show how a robot ((Franka Emika's Panda, 7-DOF)) controlled in torque can instantly stabilize the wood to avoid an accident without modifying the carpenter's sensations.
Related publications: [22]

7.8. A software architecture for the control of a 7 dof robot for the conduct of several experiment

The Franka Emika Panda robot is a 7 dof robot. Using the Robot Operating System (ROS), an experimental setup has been built to exploit this robot. The experimental setup consists in:

- the Panda robot;
- several RGBD sensors (Kinects);
- a safety laser scanner.

Dedicated algorithms have been developed to exploit the capacities of the Kinects to visualize the environment surrounding the robot and compute the distance to the closest obstacle. Several Kinects can be used simultaneously. Specific drivers have been developed to exploit the data given by the laser scanner to also determine the closest distance between a human and a robot.

Within the framework of Arcol (see 6.6) A software architecture has been developed to ease the development of different controllers. The robot can be controlled in joint position, velocity and torque using standard state-of-the-art control technique or constrained convex optimization methods. Trajectories can be easily defined, played, and modified at run time. The robot can be simulated using the GAZEBO dynamic simulator or run on the real robot with similar behaviours. All this software architecture works on a real-time patched computer.

7.9. Modulation of the robot velocity capabilities according to the distance to a human

Using the setup described in 7.8 , a controller has been defined to constrain the robot maximum velocity according to the distance to a human. The aim of this work is to be able to use the robot optimally at all time. When a human comes near the robot workspace, the robot must stop to avoid any dangerous interaction. Several strategies exist to reduce the robot velocity as a function of the distance to the human. In this work, the goal is to determine the maximum deceleration capabilities of the robot in real time and determine if the robot has the capacity to stop before a contact. If not, the maximum allowable joint velocities of the robot are reduced. When the human reaches the robot workspace, these joint velocities must be null to ensure safety. Simply reducing the joint velocity of the robot without modulating the trajectory would induce a bad tracking of the robot task. Hence, the trajectory is updated in real time to take into account the capacities of the robot. This work has been submitted for publication at the ICRA 2020 conference.

Related publications: [23]

7.10. Human motion analysis in ecological environment

The estimation of human motion from sensors that can be used in an ecological environment is an important issue being it for home assistance for frail people or for human/robot interaction in industrial contexts. We are continuing our work on data fusion from RGB-D sensors using extended Kalman filters. The original approach uses a biomechanical model of the person to obtain anthropomorphically constrained joint angles to make their estimation physically coherent. In addition, we propose a method for the optimal adjustment of the covariance matrices of the extended Kalman filter. The proposed approach was tested with six healthy subjects performing 4 rehabilitation tasks. The accuracy of the joint estimates was evaluated with a reference stereophotogrammetric system. Our results show that an affordable RGB-D sensor can be used for simple home rehabilitation when using a constrained biomechanical model. This work has led to the writing of an article now in submission to the MBEC (Medical & Biological Engineering & Computing).

In a second step, we compared the joint centre estimates obtained with the new Kinect 3 (Azure Kinect) sensor, the Kinect 2 (Kinect for Windows) and a reference stereophotogrammetric system. Regardless of the system used, we have shown that our algorithm improves the body tracker data. This study also shows the importance of defining good heuristics to merge the data according to the body tracking operation. This study is submitted for publication at ICRA 2020.

7.11. Human motion decomposition

The aim of the work is to find ways of representing human movement in order to extract meaningful physical and cognitive information.

After the realization of a state of the art on human movement, several methods are compared: principal component analysis (PCA), Fourier series decomposition and inverse optimal control.

These methods are used on a signal comprising all the angles of a walking human being. PCA makes it possible to understand the correlations between the different angles during the trajectory. Fourier series decomposition methods are used for a harmonic analysis of the signal. Finally, inverse optimal control sets up a modeling of the engine control to highlight qualitative properties characteristic of the whole motion. These three methods are tested, combined and compared on data from the EWalk database (http://gamma.cs.unc.edu/GAIT/#EWalk) in order to test emotion recognition based on these decompositions and simple classifiers.

7.12. New method for cobotic task evaluation

Two industrial studies allowed us improving our methodology for cobotic task evaluation.

- Thanks to the partnership with Suez and the work of ENSC student Nina Docteur under Auctus supervision, there are several interesting results: first, the methodological approach has been reinforced. There was a detailed analysis of an accident-prone gesture (the pipe cover raising), meetings with field agents, supervisory teams, discussions with SUEZ ergonomic expert and field observations. Second, there was a theoretical framework - a model - for the general evaluation of a cobotic task, as well as the exploration of rules for evaluating the components of this model. Five main components have been proposed for the evaluation: bio-mechanics, cognitics, usability, hedonism and social impact. An important difficulty was to mix every component and to unify the evaluation. In order to mature the model, it has been decided to carry on the partnership with a PhD work.
- The PORTAGE project (Plateforme de RoboTisation et d'Automatisation GénériquE de bâtis industriels) involves AKKA Technologies, Ez-Wheel, IIDRE and IMS laboratory. It aims at developing semi-autonomous solutions for moving heavy structures within industrial environments (e.g. aircraft industry). Our contributions is concerned with human-robot interactions, and especially accounting for real-life constraints of operators' job within their industrial environment, and translate them into design choices and requirements for the to-be-developed robotic solution. In order to identify relevant elements from the work situation, three Human Factors models have been used: Reason's model

[54], Situation Awareness model [27], and Skill Rule Knowledge (SRK) model [53]. The Reason's model details the different layers to explain accidents, notably in the aircraft industry. These layers gather equipment, procedure, training, management, policies and even psychological precursors of the operator. Therefore, this model allows investigations on potential latent causes of accident in complement with "obvious" patent causes usually more easily identified. The Situation Awareness model of Endsley describes cognitive mechanisms involved in a given situation for a person when performing actions, based on 1) the perception of elements of the current situation, 2) the understanding of the current situation, and 3) the projection of the future status of the situation. This model leads to identifying 8 daemons where situation awareness can be deteriorated, potentially resulting in accidents. The SRK model describes the decision process of an operator (or any person) performing a given task, based on his/her familiarity with this task. This model, coupled with the Situation Awareness model, can be leveraged to identify elements to be accounted when developing collaborative robots in industrial environments.

7.13. Situation awareness analysis

Baptiste Prébot, PhD student under the direction of Jean-Marc Salotti developed a methodology to analyze and assess representation sharing and situation awareness in groups of humans, possibly involving robots or artificial intelligence systems. An experiment has been carried out with two persons and a vehicle. The first person was assigned the role of mission control and the second person the role of an astronaut driving a vehicle using a real driving interface but in a simulated environment (surface of Mars in virtual reality). The two persons were located in different rooms and could communicate only by voice. Mission control had to guide the driver to a specific location. Every minute, the experiment was stopped and the two persons were asked to make a cross on a map corresponding to what they believe was the location of the vehicle. Comparisons of crosses on the maps, including ground truth locations, enabled us to determine the exactness of the localization and the degree of correct sharing of the situation. This experiment helped us better understanding communication and sharing issues, which are particularly relevant for the design of tasks and procedures for robotic operations.

Related publications:[12], [13], [9]

FLOWERS Project-Team

7. New Results

7.1. Curiosity-Driven Learning in Humans

7.1.1. Computational Models Of Information-Seeking and Curiosity-Driven Learning in Human Adults

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

This project involves a collaboration between the Flowers team and the Cognitive Neuroscience Lab of J. Gottlieb at Columbia Univ. (NY, US), on the understanding and computational modeling of mechanisms of curiosity, attention and active intrinsically motivated exploration in humans.

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

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

7.1.1.1. Context

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

Because a curiosity-based system acts in conditions of extreme uncertainty (when the distributions of events may be entirely unknown) there is in general no optimal solution to the question of which exploratory action to take [108], [130], [141]. Therefore we hypothesize that, rather than using a single optimization process as it has been the case in most previous theoretical work [86], curiosity is comprised of a family of mechanisms that include simple heuristics related to novelty/surprise and measures of learning progress over longer time scales [128] [56], [118]. These different components are related to the subject's epistemic state (knowledge and beliefs) and may be integrated with fluctuating weights that vary according to the task context. Our aim is to quantitatively characterize this dynamic, multi-dimensional system in a computational framework based on models of intrinsically motivated exploration and learning.

Because of its reliance on epistemic currencies, curiosity is also very likely to be sensitive to individual differences in personality and cognitive functions. Humans show well-documented individual differences in curiosity and exploratory drives [106], [140], and rats show individual variation in learning styles and novelty seeking behaviors [80], but the basis of these differences is not understood. We postulate that an important component of this variation is related to differences in working memory capacity and executive control which, by affecting the encoding and retention of information, will impact the individual's assessment of learning, novelty and surprise and ultimately, the value they place on these factors [133], [149], [50], [155]. To start understanding these relationships, about which nothing is known, we will search for correlations between curiosity and measures of working memory and executive control in the population of children we test in our tasks, analyzed from the point of view of a computational models of the underlying mechanisms.

A final premise guiding our research is that essential elements of curiosity are shared by humans and non-human primates. Human beings have a superior capacity for abstract reasoning and building causal models, which is a prerequisite for sophisticated forms of curiosity such as scientific research. However, if the task is adequately simplified, essential elements of curiosity are also found in monkeys [106], [98] and, with adequate characterization, this species can become a useful model system for understanding the neurophysiological mechanisms.

7.1.1.2. Objectives

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

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

7.1.1.3. Current results: experiments in Active Categorization

In 2018, we have been occupied by analyzing data of the human adult experiment conducted in 2017. In this experiment we asked whether humans possess, and use, metacognitive abilities to guide task choices in two contexts motivational contexts, in which they could freely choose to learn about 4 competing tasks. Participants (n = 505, recruited via Amazon Mechanical Turk) were asked to play a categorization game with four distinct difficulty levels. Some participants had been explicitly prescribed a goal of maximizing their learning across the difficulty levels (across tasks), while others did not receive any specific instructions regarding the goal of the game. The experiment yielded a rich but complex set of data. The data includes records of participants' classification responses, task choices, reaction times, and post-task self-reports about various subjective evaluations of the competing tasks (e.g. subjective interest, progress, learning potential, etc.). We are now finalizing the results and a computational model of the underlying cognitive and motivational mechanisms in order to prepare them for public dissemination.

The central question going into the study was, how do active learners become interested in specific learning exercises: how do they decide which task to engage with, when none of the tasks provide external rewards. Last year, we identified some of the key behavioral observations that merited further attention. First, we saw a clear effect of an external goal prescription on the participants' overall task selection strategy. People who were explicitly instructed to try to maximize their learning across the 4 tasks challenged themselves more by giving preference towards harder tasks. In contrast, those who were simply familiarized with the rules of the game

and not given any explicit suggestions from the experiments did not show this overchallenge bias and had a slight preference for easier tasks (see figure 5). Second, we observed that although strategies varied between the two instruction groups, there was some considerable within-group variability in learning. We found that in both groups, people had varying success in learning the classification task for each task family resulting in four distinct performance based groups: learners of 0, 1, 2, or 3 tasks (task 4 was unlearnable), as shown in figure 6 . Importantly, successful 3-task learners in both instruction groups exhibited similar task preferences, suggesting that (1) even in the absence of external instruction, people can be motivated to explore the task space and (2) intrinsically motivated exploration is similar to strategies employed when a learner is trying to maximize her learning.

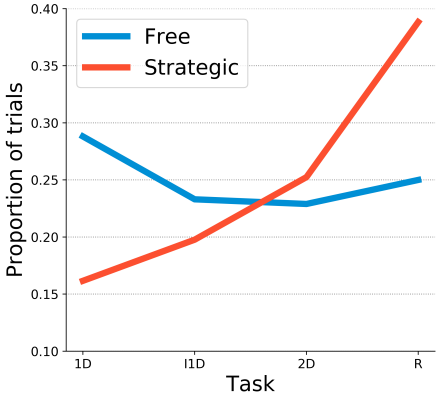


Figure 5. Proportion of trials on each task (1D, I1D, 2D, and R). The group with an externally prescribed learning maximization goal is referred to as "Strategic", and the unconstrained group is referred to as "Free". 1D was the task were categorization was determined by a single variable dimension. In I1D (ignore 1D), the stimuli varied across 2 dimensions, but only one determined the stimulus category. In 2D, there were 2 variable dimensions and both jointly determined the category. Finally in R, there were 2 variable dimensions, but none of them could reliably predict the stimulus class. The top plot shows data aggregated across experimental groups, shown separately in the bottom plot. In hte figure, task difficulty increases from left to right.

Assuming that task choice decisions are based on a subjective evaluative process that assigns value to choice candidates, we considered a simple choice model of task selection. In a classic conditional logit model [116], choices are made probabilistically and the choice probabilities are proportional to choice utilities (the inherent subjective value of a choice; also see ref [159]). We elaborated on the utility component of the basic choice model to consider two utility aspects of interest: a relative measure of learning progress (LP) and an absolute measure of proportion correct (PC). Although both measures are based on empirical feedback (correct / incorrect), the LP measure is considered relative, because it captures how performance changes over time by comparing performance estimates across different time scales, while PC is absolute in a sense that it only characterizes performance at a given instance. While PC alone does not differentiate between an unfamiliar (but potentially easy) task on which the performance might be low and a familiar but very hard task, the former can have markedly LP (due to the gradual improvement on that task) than the latter. Only the tasks characterized by high LP are then worthy of time and effort if the goal is to master tasks. The utility component in our model thus includes two principal quantities:

$$u_{i,t} = \alpha LP_{i,t} + \beta PC_{i,t}$$

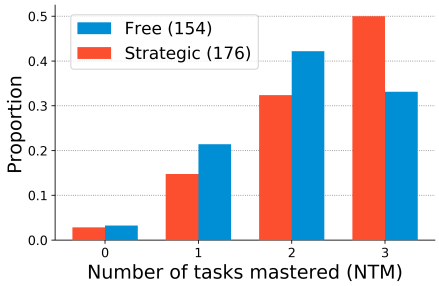


Figure 6. Learning outcomes by group. The group with an externally prescribed learning maximization goal is referred to as "Strategic", and the unconstrained group is referred to as "Free". Number of people in the groups are given in parentheses. The figure shows that some people seem to set and pursue learning maximization goals in the absence of external rewards and instructions. Additionally, learning goals have to be matched with appropriate behavioral strategies in order to be reached, which explains the failure to learn all 3 tasks by some people in the Strategic group.

where α and β are free parameters indexing the model's sensitivity to LP and PC, respectively. Index i designates the task, while t indexes time. Thus, in our model, utility is seen as a dual-component linear computation of both relative and absolute competence quantities. Task utilities enter the decision-making process that assigns relative preference to each task:

$$p(task_i) = \frac{e^{u_i/\tau}}{\sum_j e^{u_j/\tau}}$$

where τ is another free parameter (known as temperature) that controls the stochasticity of utility-based decision. The sum over j elements, $\sum_j$ constitutes the total exponentiated utility of each task in the task space, thus normalizing each individual exponentiated task utility u_i.

The computation of the utility components is important for the model, because it ultimately determines how well the model can fit to choice data. We started exploring the model with a simple definition for both LP and PC. Both components are based on averaging binary feedback over multiple trials in the past. Since the familiarization stage of our experiment was 15-trials-long, the first free choice was made based on feedback data from 15 trials on each task. Accordingly, we defined PC to be the proportion of correct guesses in 15 trials. LP was defined as the absolute difference between the first 9 and the last 6 trials of the recent most 15 trials in the past. While a participant was engaged with one of the tasks, LP and PC for that task changed according to her dynamic performance, while LPs and PCs for other tasks remained unchanged. We acknowledge that there are probably multiple other components at play when it comes to utility computation, some of which may have little to do with task competence. We also submit that there are multiple ways of defining the PC and LP components that are more biologically rooted and plausible given what we know of memory and metacognition. Finally, we do not rule out the possibility of dynamic changes of free parameters themselves, corresponding to changes in motivation during the learning process. All of these considerations are worthy directions of future research, but in this study we focused on finding some necessary evidence for the sensitivity to learning progress.

We fitted the model to each individual's choice data using maximum likelihood estimation. Assuming that choice probabilities on each trial come from a categorical distribution (also called a generalized Bernoulli distribution), where the probability of choosing item i is given by:

$$P(\mathbf{x} \mid \mathbf{p}) = \prod_{i=1}^{k} p_i^{x_i}$$

where $\mathbf{p}$ is a vector of probabilities associated with k events, and $\mathbf{p}$ is a one-hot encoded vector representing discrete items x_i. We add a time index to indicate the dynamic quality of choice probabilities, so that:

$$P_t(\mathbf{x} \mid \mathbf{p}_t) = \prod_{i=1}^{k} p_{i,t}^{x_i}$$

Then, the likelihood of the choice model ($p(task_i \mid \alpha, \beta, \tau)$) at time t is equal to the product of choice probabilities given by that model for that time step:

$$L_t(\alpha, \beta, \tau \mid \mathbf{x}) = \prod_{i}^{k} p_{i,t}^{x_i}$$

and since the empirical choice data can be represented in a one-hot format, the likelihood of the model for a given time point boils down to the predicted probability of the actualy choice:

$$L_t(\alpha, \beta, \tau \mid choice = i) = p_{i,t} = \frac{e^{u_{i,t}/\tau}}{\sum_j e^{u_{j,t}/\tau}}$$

The likelihood of the model across all m trials is obtained by applying the product rule of probability:

$$L_{overall}(\alpha, \beta, \tau \mid choice = i) = \prod_{t}^{m} p_{i,t} = \frac{e^{u_{i,t}/\tau}}{\sum_j e^{u_{j,t}/\tau}}$$

For convenience, we use the negative log transformation to avoid computational precision problems and convert a likelihood maximization objective into negative likelihood minimization problem solvable by publicly accessible optimization tools:

$$-\log L_{overall}(\alpha, \beta, \tau \mid choice = i) = -\sum_{t}^{m} \log p_{i,t}$$

Having formulated the likelihood function, we optimized the free parameters to obtain a model that fits the individual data best. We thus fit an individual-level model to each participant's choice data. The fitted parameters can be interpreted as relative sensitivity to the competence quantities of interest (LP and PC), since these quantities share the same range of values (0 to 1). Finally, we performed some group-level analyses on these individual-level parameter estimates to evaluate certain group-level effects that might influence them (e.g. effect of instruction or learning proficiency).

The group with a learning maximization goal devalued tasks with higher positive feedback expectation. Qualitatively, this matched our prior observations showing their strong preference for harder tasks. However, the best learners (3-task learners) across both instruction groups showed a slight preference for learning progress and a relatively strong aversion to positive feedback. It appears that what separated better and worse learners among the learning maximizers was whether they followed learning progress, and not just the feedback heuristic. On the other hand, while less successful learners in the unconstrained group seemed to choose tasks according to learning progress, they valued positive feedback over it, which prevented them from exploring more challenging learnable tasks. This is reflected in group mean values of the fitted parameters summarized in table 1

Table 1. Fitted parameter values averaged within number of tasks learned and instruction groups. The group with an externally prescribed learning maximization goal is referred to as "Strategic", and the unconstrained group is referred to as "Free". NTM stands for the number of tasks mastered

Group	NTM	Learning progress	Proportion correct	Temperature
	1	0.27	0.56	6.92
	2	0.07	0.32	5.90
	3	0.12	-0.15	6.14
	1	-0.01	-0.56	7.14
	2	-0.15	-0.41	6.42
	3	0.11	-0.44	6.59

We also looked at the relative importance of the utility model parameters by performing model comparisons. We compared 4 models based on combinations of 2 factors: PC and LP. According to the AIC scores (see Table 2), the best model was the one which included both LP and PC factors, followed by the PC-only model, and then by the LP-only model. The random-choice model came in last with the highest AIC score. The results of this model comparison show that both learning progress and positive feedback expectation factors provide substantive improvement to model likelihood compared to when these factors are included alone, or when neither of them is present. This is potentially important, as it provides some evidence for the role of relative competence kind of measure in autonomous exploration. We are planning to submit the work described about to a high impact peer-reviewed journal focusing on computational modeling of human behavior.

Table 2. Model comparisons.

Model	AIC	$AIC - AIC_{min}$
LP + PC	568.99	-
PC	593.51	24.52
LP	658.46	89.47
Random	693.60	124.61

7.1.2. Experimental study of the role of intrinsic motivation in developmental psychology experiments and in the development of tool use

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

Children are so curious to explore their environment that it can be hard to focus their attention on one given activity. Many experiments in developmental psychology evaluate particular skills of children by setting up a task that the child is encouraged to solve. However, children may sometimes be following their own motivation to explore the experimental setup or other things in the environment. We suggest that considering the intrinsic motivations of children in those experiments could help understanding their role in the learning of related skills and on long-term child development. To illustrate this idea, we reanalyzed and reinterpretd a typical experiment aiming to evaluate particular skills in infants. In this experiment run by Lauriane Rat-Fischer et al, 32 21-month old infants have to retrieve a toy stuck inside a tube, by inserting several blocks in sequence into the tube. In order to understand the mechanisms of the motivations of babies, we studied in detail their behaviors, goals and strategies in this experiment. We showed that their motivations are diverse and do not always coincide with the target goal expected and made salient by the experimenter. Intrinsically motivated exploration seems to play an important role in the observed behaviors and to interfere with the measured success rates. This new interpretation provides a motivation for studying curiosity and intrinsic motivations in robotic models.

7.2. Intrinsically Motivated Learning in Artificial Intelligence

7.2.1. Intrinsically Motivated Goal Exploration and Goal-Parameterized Reinforcement Learning

Participants: Sébastien Forestier, Pierre-Yves Oudeyer [correspondant], Olivier Sigaud, Cédric Colas, Adrien Laversanne-Finot, Rémy Portelas, Grgur Kovac.

7.2.1.1. Intrinsically Motivated Exploration of Modular Goal Spaces and the Emergence of Tool use

A major challenge in robotics is to learn goal-parametrized policies to solve multi-task reinforcement learning problems in high-dimensional continuous action and effect spaces. Of particular interest is the acquisition of inverse models which map a space of sensorimotor goals to a space of motor programs that solve them. For example, this could be a robot learning which movements of the arm and hand can push or throw an object in each of several target locations, or which arm movements allow to produce which displacements of several objects potentially interacting with each other, e.g. in the case of tool use. Specifically, acquiring such repertoires of skills through incremental exploration of the environment has been argued to be a key target for life-long developmental learning [54].

Recently, we developed a formal framework called "Intrinsically motivated goal exploration processes" (IMGEPs), enabling study agents that generate and sequence their own goals to learn world models and skill repertoirs, that is both more compact and more general than our previous models [82]. We experimented several implementations of these processes in a complex robotic setup with multiple objects (see Fig. 7), associated to multiple spaces of parameterized reinforcement learning problems, and where the robot can learn how to use certain objects as tools to manipulate other objects. We analyzed how curriculum learning is automated in this unsupervised multi-goal exploration process, and compared the trajectory of exploration and learning of these spaces of problems with the one generated by other mechanisms such as hand-designed learning curriculum, or exploration targeting a single space of problems, and random motor exploration. We showed that learning several spaces of diverse problems can be more efficient for learning complex skills than only trying to directly learn these complex skills. We illustrated the computational efficiency of IMGEPs as these robotic experiments use a simple memory-based low-level policy representations and search algorithm, enabling the whole system to learn online and incrementally on a Raspberry Pi 3.

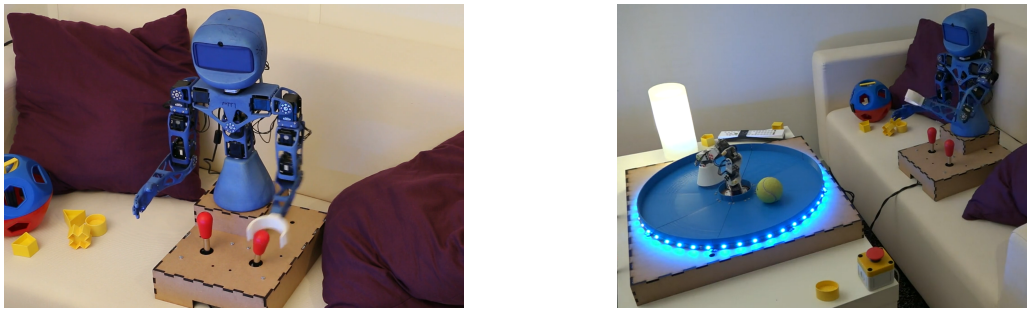


Figure 7. Robotic setup. Left: a Poppy Torso robot (the learning agent) is mounted in front of two joysticks. Right: full setup: a Poppy Ergo robot (seen as a robotic toy) is controlled by the right joystick and can hit a tennis ball in the arena which changes some lights and sounds.

In order to run many systematic scientific experiments in a shorter time, we scaled up this experimental setup to a platform of 6 identical Poppy Torso robots, each of them having the same environment to interact with. Every robot can run a different task with a specific algorithm and parameters each (see Fig. 8). Moreover, each Poppy Torso can also perceives the motion of a second Poppy Ergo robot, than can be used, this time, as a distractor performing random motions to complicate the learning problem. 12 top cameras and 6 head

cameras can dump video streams during experiments, in order to record video datasets. This setup is now used to perform more experiments to compare different variants of curiosity-driven learning algorithms.

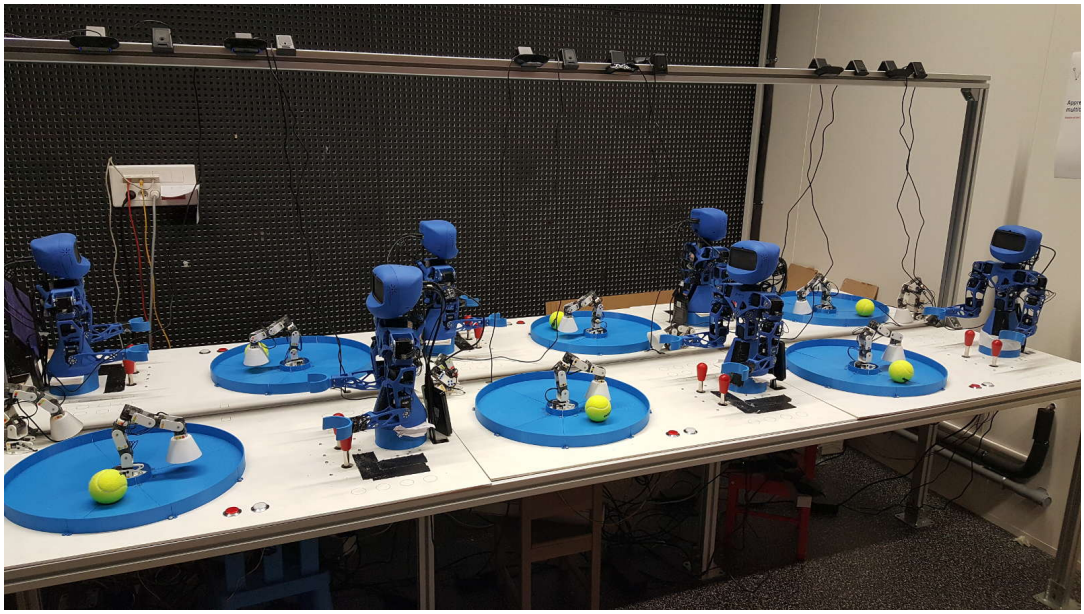


Figure 8. Platform of 6 robots with identical environment: joysticks, Poppy Ergo, ball in an arena, and a distractor. The central bar supports the 12 top cameras.

7.2.1.2. Leveraging the Malmo Minecraft platform to study IMGEP in rich simulations

We continued to leverage the Malmo platform to study curiosity-driven learning applied to multi-goal reinforcement learning tasks (https://github.com/Microsoft/malmo). The first step was to implement an environment called Malmo Mountain Cart (MMC), designed to be well suited to study multi-goal reinforcement learning (see figure [9]). We then showed that IMGEP methods could efficiently explore the MMC environment without any extrinsic rewards. We further showed that, even in the presence of distractors in the goal space, IMGEP methods still managed to discover complex behaviors such as reaching and swinging the cart, especially Active Model Babbling which ignored distractors by monitoring learning progress.

7.2.1.3. Unsupervised Learning of Modular Goal Spaces for Intrinsically Motivated Goal Exploration

Intrinsically motivated goal exploration algorithms enable machines to discover repertoires of policies that produce a diversity of effects in complex environments. These exploration algorithms have been shown to allow real world robots to acquire skills such as tool use in high-dimensional continuous state and action spaces, as shown in previous sections. However, they have so far assumed that self-generated goals are sampled in a specifically engineered feature space, limiting their autonomy. We have proposed an approach using deep representation learning algorithms to learn an adequate goal space. This is a developmental 2-stage approach: first, in a perceptual learning stage, deep learning algorithms use passive raw sensor observations of world changes to learn a corresponding latent space; then goal exploration happens in a second stage by sampling goals in this latent space. We made experiments with a simulated robot arm interacting with an object, and we show that exploration algorithms using such learned representations can closely match, and even sometimes improve, the performance obtained using engineered representations. This work was presented at ICLR 2018 [136].

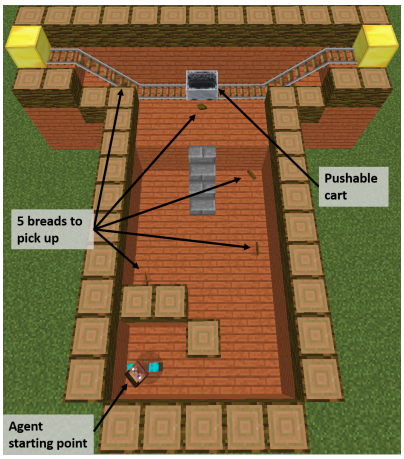


Figure 9. Malmo Mountain Cart. In this episodic environment the agent starts at the bottom left corner of the arena and is able to act on the environment using 2 continuous action commands: move and strafe. If the agent manages to get out of its starting area it may be able to collect items dispatched within the arena. If the agent manages to climb the stairs it may get close enough to the cart to move it along its railroad.

However, in the case of more complex environments containing multiple objects or distractors, an efficient exploration requires that the structure of the goal space reflects the one of the environment. We studied how the structure of the learned goal space using a representation learning algorithm impacts the exploration phase. In particular, we studied how disentangled representations compare to their entangled counterparts in a paper published at CoRL 2019 [101], associated with a blog post available at: https://openlab-flowers.inria.fr/t/discovery-of-independently-controllable-features-through-autonomous-goal-setting/494.

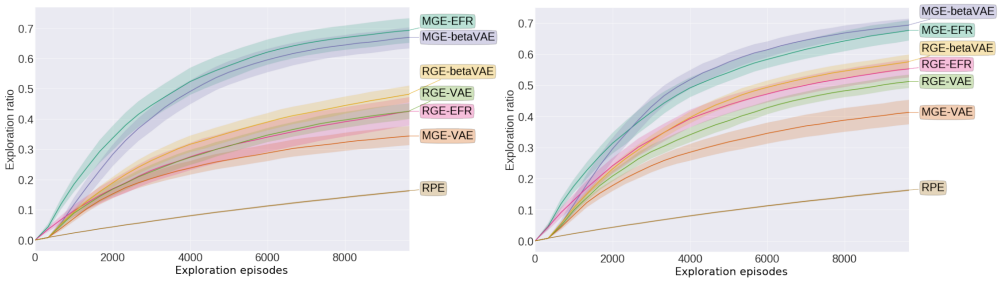
Those ideas were evaluated on a simple benchmark where a seven joints robotic arm evolves in an environment containing two balls. One of the ball can be grasped by the arm and moved around whereas the second one acts as a distractor: it cannot be grasped by the robotic arm and moves randomly across the environment.

Our results showed that using a disentangled goal space leads to better exploration performances than an entangled goal space: the goal exploration algorithm discovers a wider variety of outcomes in less exploration steps (see Figure 10). We further showed that when the representation is disentangled, one can leverage it by sampling goals that maximize learning progress in a modular manner. Lastly, we have shown that the measure of learning progress, used to drive curiosity-driven exploration, can be used simultaneously to discover abstract independently controllable features of the environment.

Finally, we experimented the applicability of those principles on a real-world robotic setup, where a 6-joint robotic arm learns to manipulate a ball inside an arena, by choosing goals in a space learned from its past experience, presented in this technical report: https://arxiv.org/abs/1906.03967.

7.2.1.4. Monolithic Intrinsically Motivated Modular Multi-Goal Reinforcement Learning

In this project we merged two families of algorithms. The first family is the population-based Intrinsically Motivated Goal Exploration Processes (IMGEP) developped in the team (see [83] for a presentation). In this family, autonomous learning agents sets their own goals and learn to reach them. Intrinsic motivation under the form of absolute learning progress is used to guide the selection of goals to target. In some variations of this framework, goals can be represented as coming from different *modules* or *tasks*. Intrinsic motivations are then used to guide the choice of the next task to target.



(a) Small exploration noise ($\sigma = 0.05$) (b) Large exploration noise ($\sigma = 0.1$)

Figure 10. Exploration ratio during exploration for different exploration noises. Architectures with disentangled representations as a goal space (βVAE) explore more than those with entangled representations (VAE). Similarly modular architectures (MGE) explore more than flat architecture (RGE).

The second family encompasses goal-parameterized reinforcement learning algorithms. The first algorithm of this category used an architecture called Universal Value Function Approximators (UVFA), and enabled to train a single policy on an infinite number of goals (continuous goal spaces) [144] by appending the current goal to the input of the neural network used to approximate the value function and the policy. Using a single network allows to share weights among the different goals, which results in faster learning (shared representations). Later, HER [51] introduced a goal replay policy: the actual goal aimed at, could be replaced by a fictive goal when learning. This could be thought of as if the agent were pretending it wanted to reach a goal that it actually reached later on in the trajectory, in place of the true goal. This enables cross-goal learning and speeds up training. Finally, UNICORN [111] proposed to use UVFA to achieve multi-task learning with a discrete task-set.

In this project, we developed CURIOUS [33] (ICML 2019), an intrinsically motivated reinforcement learning algorithm able to achieve both multiple tasks and multiple goals with a single neural policy. It was tested on a custom multi-task, multi-goal environment adapted from the OpenAI Gym Fetch environments [61], see Figure 11 . CURIOUS is inspired from the second family as it proposes an extension of the UVFA architecture. Here, the current task is encoded by a one-hot code corresponding to the task id. The goal is of size $\sum_{i=1}^{N} dim(\mathcal{G}_i)$ where $\mathcal{G}_i$ is the goal space corresponding to task T_i. All components are zeroed except the ones corresponding to the current goal g_i of the current task T_i, see Figure 12 .

CURIOUS is also inspired from the first family, as it self-generates its own tasks and goals and uses a measure of learning progress to decide which task to target at any given moment. The learning progress is computed as the absolute value of the difference of non-overlapping window average of the successes or failures

$$LP_i(t) = \frac{|\sum_{\tau=t-2l}^{t-l} S_\tau - \sum_{\tau=t-l}^{t} S_\tau|}{2l},$$

where S_τ describes a success (1) or a failure (0) and l is a time window length. The learning progress is then used in two ways: it guides the selection of the next task to attempt, and it guides the selection of the task to replay. Cross-goal and cross-task learning are achieved by replacing the goal and/or task in the transition by another. When training on one combination of task and goal, the agent can therefore use this sample to learn about other tasks and goals. Here, we decide to replay and learn more on tasks for which the absolute learning progress is high. This helps for several reasons: 1) the agent does not focus on already learned tasks, as the corresponding learning progress is null, 2) the agent does not focus on impossible tasks for the same reason.

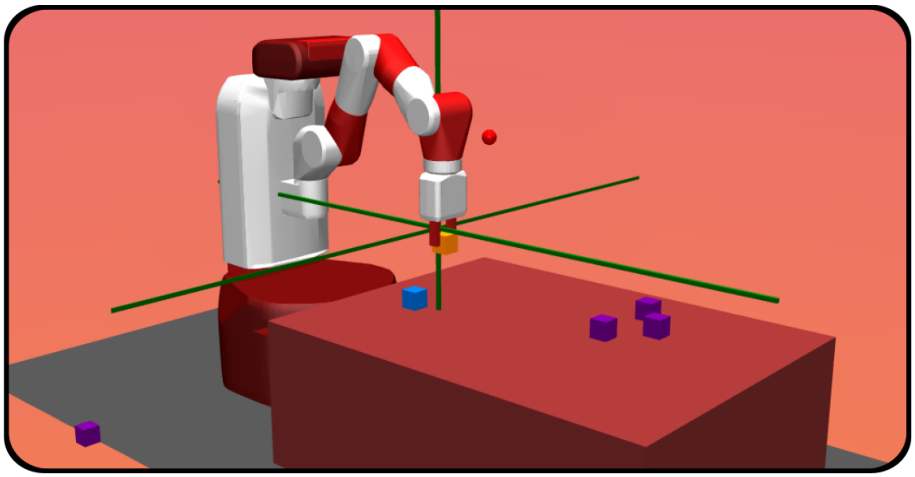


Figure 11. Custom multi-task and multi-goal environment to test the CURIOUS algorithm.

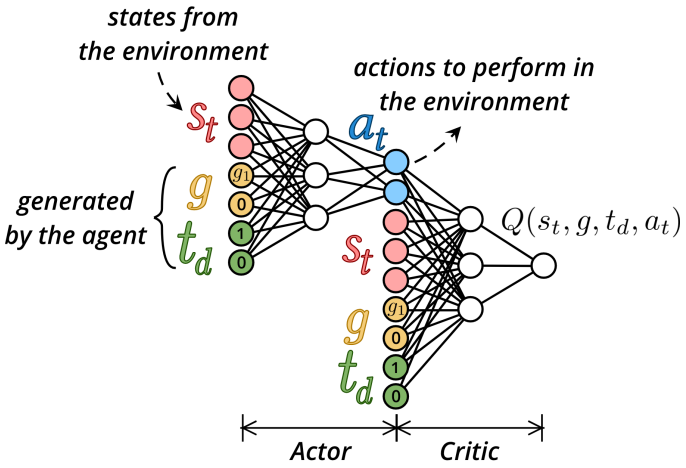


Figure 12. Architecture extended from Universal Value Function Approximators. In this example, the agent is targeting task T_1 among two tasks, each corresponding to a 1 dimension goal.

The agent focuses more on tasks that are being learned (therefore maximizing learning progress), and on tasks that are being forgotten (therefore fighting the problem of forgetting). Indeed, when many tasks are learned in a same network, chances are tasks that are not being attempted often will be forgotten after a while.

In this project, we compare CURIOUS to two baselines: 1) a flat representation algorithm where goals are set from a multi dimensional space including all tasks (equivalent to HER); 2) a task-expert algorithm where a multi-goal UVFA expert policy is trained for each task. The results are shown in Figure 13 .

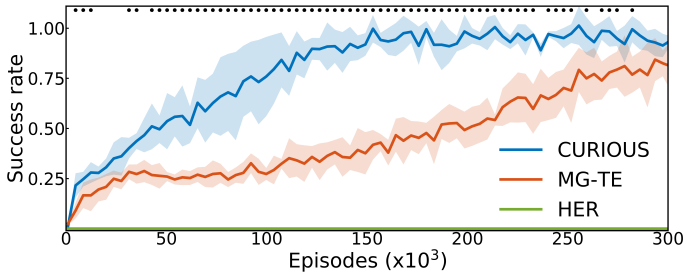


Figure 13. Comparison of CURIOUS to alternative algorithms.

7.2.1.5. Autonomous Multi-Goal Reinforcement Learning with Natural Language

This project follows the CURIOUS project on intrinsically motivated modular multi-goal reinforcement learning [33]. In the CURIOUS algorithm, we presented an agent able to tackle multiple goals of multiple types using a single controller. However, the agent needed to have access to the description of each of the goal types, and their associated reward functions. This represents a considerable amount of prior knowledge for the engineer to encode into the agent. In our new project, the agent builds its own representations of goals, can tackle a growing set of goals, and learns its own reward function, all this through interactions in natural language with a social partner.

The agent does not know any potential goal at first, and act randomly. As it reaches outcomes that are meaningful for the social partner, the social partner provides descriptions of the scene in natural language. The agent stores descriptions and corresponding states for two purposes. First it builds a list of potential goals, reaching back these outcomes that the social partner described. Second, it uses the combination of state and state description to learn a reward function, mapping current state and language descriptions to a binary feedback: 1 if the description is satisfied by the current state, 0 if not.

The agent sets goals to itself from the set of previously discovered descriptions, and is able to learn how to reach them thanks to its learned internal reward function. Concretely, the agent learns a set of 50+ goals from these interactions. We showed co-learning of the reward function and the policy did not produce consequent overhead compared to using an oracle reward function (see Figure 14). This project led to an article accepted at the NeurIPS workshop Visually Grounded Interaction and Language [35]. Current work aims at learning the language model mapping the description into a continuous goal space used as input to the policy and reward function using recurrent networks.

7.2.1.6. Intrinsically Motivated Exploration and Multi-Goal RL with First-Person Images

The aim of this project is to create an exploration process in first-person 3D environments. Following the work presented in [121], [135], [157] the current algorithm is setup in a similar manner. The agent observes the environment in a first person manner and is given a goal to reach. Furthermore, the goal policy samples goals that encourage the agent to explore the environment.

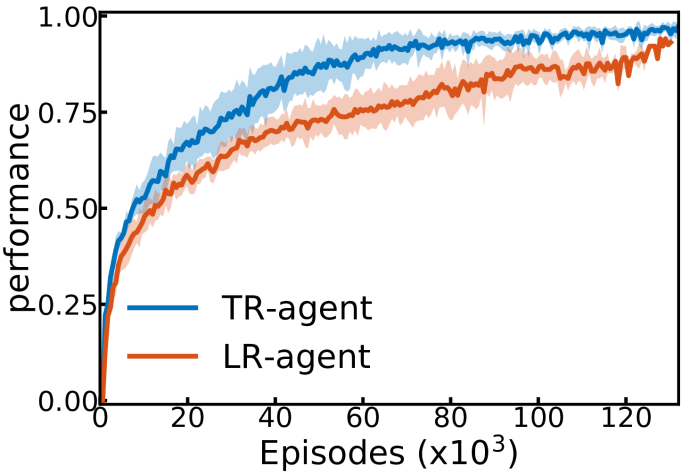


Figure 14. Comparison of performance of agents having access to the true oracle reward function (TR-agent), and agents co-learning policy and reward function (LR-agent).

Currently, there are two ways of representing and sampling goals. Following [157] the goal policy has a buffer of states previously visited by the agent and samples from this buffer the next goals as first person observations. In this setting a goal conditioned reward function is also learned in the form of a reachability network introduced in [142]. On the other hand, following [136], [121] and [135] one can learn a latent representation of a goal using an autoencoder (VAE) and then sample from this generative model or in the latent space. Then, we can use the L2 distance in the latent space as a reward function. The experiments are conducted on a set of Unity environments created in the team.

7.2.2. Teacher algorithms for curriculum learning of Deep RL in continuously parameterized environments

Participants: Remy Portelas [correspondant], Katja Hoffman, Pierre-Yves Oudeyer.

In this work we considered the problem of how a teacher algorithm can enable an unknown Deep Reinforcement Learning (DRL) student to become good at a skill over a wide range of diverse environments. To do so, we studied how a teacher algorithm can learn to generate a learning curriculum, whereby it sequentially samples parameters controlling a stochastic procedural generation of environments. Because it does not initially know the capacities of its student, a key challenge for the teacher is to discover which environments are easy, difficult or unlearnable, and in what order to propose them to maximize the efficiency of learning over the learnable ones. To achieve this, this problem is transformed into a surrogate continuous bandit problem where the teacher samples environments in order to maximize absolute learning progress of its student. We presented ALP-GMM (see figure 15), a new algorithm modeling absolute learning progress with Gaussian mixture models. We also adapted existing algorithms and provided a complete study in the context of DRL. Using parameterized variants of the BipedalWalker environment, we studied their efficiency to personalize a learning curriculum for different learners (embodiments), their robustness to the ratio of learnable/unlearnable environments, and their scalability to non-linear and high-dimensional parameter spaces. Videos and code are available at https://github.com/flowersteam/teachDeepRL.

Overall, this work demonstrated that LP-based teacher algorithms could successfully guide DRL agents to learn in difficult continuously parameterized environments with irrelevant dimensions and large proportions of unfeasible tasks. With no prior knowledge of its student's abilities and only loose boundaries on the task

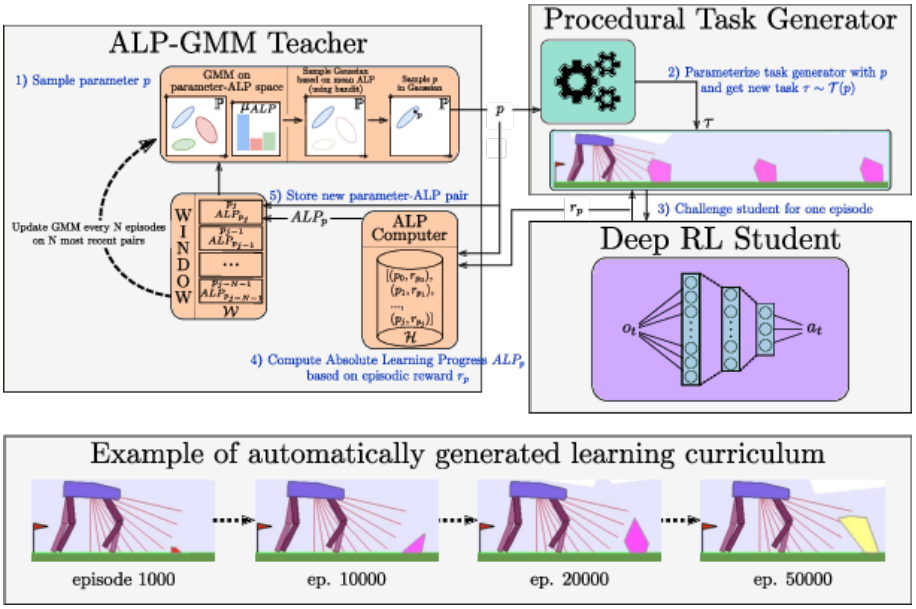


Figure 15. Schematic view of an ALP-GMM teacher's workflow

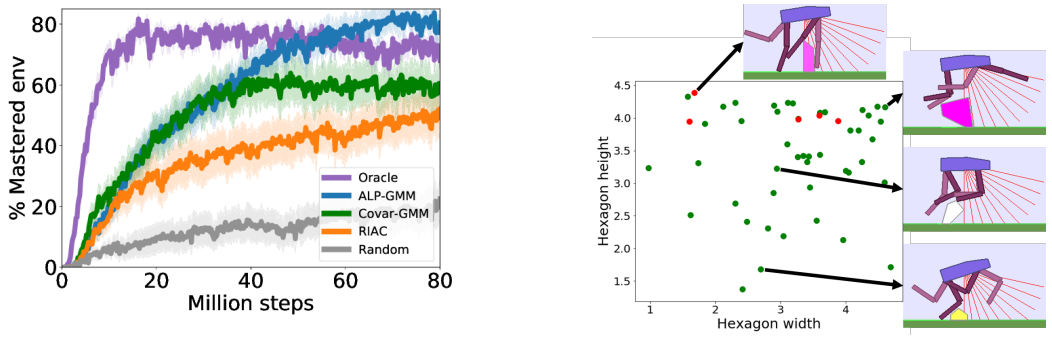


Figure 16. ***Teacher-Student approaches in Hexagon Tracks.Left:*** *Evolution of mastered tracks for Teacher-Student approaches in Hexagon Tracks. 32 seeded runs (25 for Random) of 80 Millions steps where performed for each condition. The mean performance is plotted with shaded areas representing the standard error of the mean.* ***Right:*** *A visualization of which track distributions of the test-set are mastered (i.e* $r_t > 230$*, shown by green dots) by an ALP-GMM run after 80 million steps.*

space, ALP-GMM, our proposed teacher, consistently outperformed random heuristics and occasionally even expert-designed curricula (see figure 16). This work was presented at CoRL 2019 [38].

ALP-GMM, which is conceptually simple and has very few crucial hyperparameters, opens-up exciting perspectives inside and outside DRL for curriculum learning problems. Within DRL, it could be applied to previous work on autonomous goal exploration through incremental building of goal spaces [101]. In this case several ALP-GMM instances could scaffold the learning agent in each of its autonomously discovered goal spaces. Another domain of applicability is assisted education, for which current state of the art relies heavily on expert knowledge [68] and is mostly applied to discrete task sets.

7.3. Automated Discovery in Self-Organizing Systems

7.3.1. Curiosity-driven Learning for Automated Discovery of Physico-Chemical Structures

Participants: Chris Reinke [correspondant], Mayalen Etcheverry, Pierre-Yves Oudeyer.

7.3.1.1. Introduction

Intrinsically motivated goal exploration algorithms (IMGEPs) enable machines to discover repertoires of action policies that produce a diversity of effects in complex environments. In robotics, these exploration algorithms have been shown to allow real world robots to acquire skills such as tool use [81] [55]. In other domains such as chemistry and physics, they open the possibility to automate the discovery of novel chemical or physical structures produced by complex dynamical systems [134]. However, they have so far assumed that self-generated goals are sampled in a specifically engineered feature space, limiting their autonomy. Recent work has shown how unsupervised deep learning approaches could be used to learn goal space representations [136] but they have used precollected data to learn the representations. This project studies how IMGEPs can be extended and used for automated discovery of behaviours of dynamical systems in physics or chemistry without using assumptions or knowledge about such systems.

As a first step towards this goal we choose Lenia [66], a simulated high-dimensional complex dynamical system, as a target system. Lenia is a continuous cellular automaton where diverse visual structures can self-organize (Fig.17 , c). It consists of a two-dimensional grid of cells $A \in [0,1]^{256\times256}$ where the state of each cell is a real-valued scalar activity $A^t(x) \in [0,1]$. The state of cells evolves over discrete time steps t. The activity change is computed by integrating the activity of neighbouring cells. Lenia's behavior is controlled by its initial pattern $A^{t=1}$ and several settings that control the dynamics of the activity change. Lenia can produce diverse patterns with different dynamics. Most interesting, spatially localized coherent patterns that resemble in their shapes microscopic *animals* can emerge. Our goal was to develop methods that allow to explore a high diversity of such animal patterns.

We could successfully accomplish this goal [30] based on two key contributions of our research: 1) the usage of compositional pattern producing networks (CPPNs) for the generation of initial states for Lenia, and 2) the development of a novel IMGEP algorithm that learns goal representations online during the exploration of the system.

7.3.1.2. 1) CPPNs for the generation of initial states

A key role in the generation of patterns in dynamical systems is their initial state $A^{t=1}$. IMGEPs sample these initial states and apply random perturbations to them during the exploration. For Lenia this state is a two-dimensional grid with 256×256 cells. Performing directly a random sampling of the 256×256 grid cells results in initial patterns that resemble white noise. Such random states result mainly in the emergence of global patterns that spread over the whole state space, complicating the search for spatially localized patterns. We solved the sampling problem for the initial states by using compositional pattern producing networks (CPPNs) [148]. CPPNs are recurrent neural networks that allow the generation of structured initial states (Fig.17 , a). The CPPNs are used as part of the system parameters which are explored by the algorithms. They are defined by their network structure (number of neurons, connections between neurons) and their connection weights. They include a mechanism for random mutation of the weights and structure.

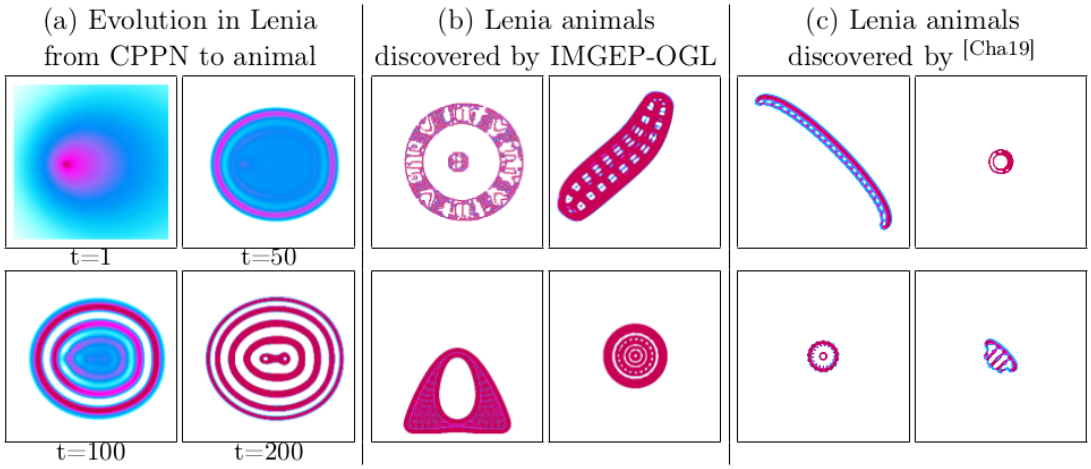


Figure 17. Example patterns produced by the Lenia system. Illustration of the dynamical morphing from an initial CPPN image to an animal (a). The automated discovery (b) is able to find similar complex animals as a human-expert manual search (c) by [66].

7.3.1.3. 2) IMGEP for Online Learning of Goal Space Representations

We proposed an online goal space learning IMGEP (IMGEP-OGL), which learns the goal space incrementally during the exploration process. A variational autoencoder (VAE) is used to encode Lenia patterns into a 8-dimensional latent representation used as goal space. The training procedure of the VAE is integrated in the goal sampling exploration process by first initializing the VAE with random weights. The VAE network is then trained every K explorations for E epochs on the previously idetnfied patterns during the exploration.

7.3.1.4. Experiments

We evaluated the performance of the novel IMGEP-OGL to other exploration algorithms by comparing the diversity of their identified patterns. Diversity is measured by the spread of the exploration in an *analytic behavior space*. This space is defined by a latent representation space that was build through the training of a VAE to learn the important features over a very large dataset of Lenia patterns identified during the many experiments over all evaluated algorithms. We then augmented that space by concatenating hand-defined features. Each identified Lenia pattern is represented by a specific point in this space. The space was then discretized in a fixed number of areas/bins of equal size. The final diversity measure of each algorithm is the number of areas/bins in which at least one explored pattern exists.

We compared different exploration algorithms to the novel IMGEP-OGL: 1) Random exploration of system parameters, 2) IMGEP-HGS: IMGEP with a hand-defined goal space, 3) IMGEP-PGL: IMGEP with a learned goal space via an VAE by a precollected dataset of Lenia patterns, and 4) IMGEP-RGS: IMGEP with a VAE with random weights that defines the goal space.

The system parameters θ consisted of a CPPN that generates the initial state $A^{t=1}$ for Lenia and 6 further settings defining Lenia's dynamics: $\theta = [\text{CPPN} \rightarrow A^{t=1}, R, T, \mu, \sigma, \beta_1, \beta_2, \beta_3]$. The CPPNs were initialized and mutated by a random process that defines their structure and connection weights as done. The random initialization of the other Lenia settings was done by an uniform distribution and their mutation by a Gaussian distribution around the original values.

7.3.1.5. Results

The diversity of identified patterns in the analytic behavior space show that IMGEP approaches with learned goal spaces via VAEs (PGL, OGL) could identify the highest diversity of patterns overall (Fig. 18 , a). They were followed by the IMGEP with a hand-defined goal space (HGS). The lowest performance had the random exploration and the IMGEP with a random goal space (RGS). The advantage of learned goals space approaches (PGL, OGL) over all other approaches was even stronger for the diversity of animal patterns, i.e. the main goal of our exploration (Fig. 18 , b).

(a) Diversity of all identified Patterns (b) Diversity of Animal Patterns

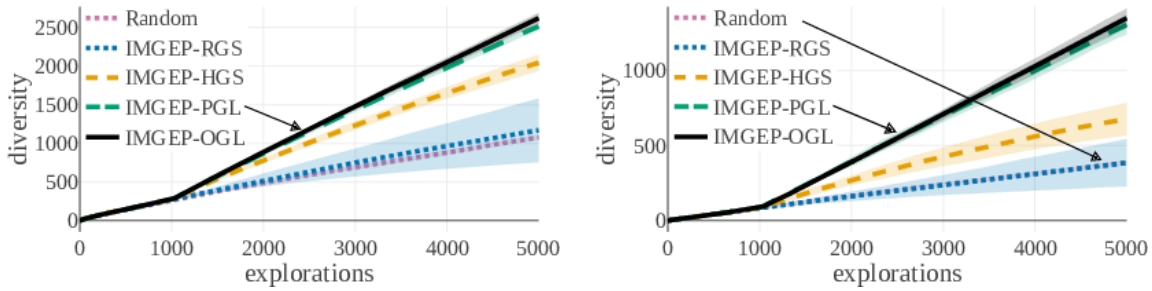


Figure 18. (a) All IMGEPs reach a higher diversity in the analytic behavior space over all patterns than random search. (b) IMGEPs with a learned goal space are especially successful in identifying a diversity of animal patterns. Depicted is the average diversity ($n = 10$) with the standard deviation as shaded area (for some not visible because it is too small).

7.3.1.6. Conclusion

Our goal was to investigate new techniques based on intrinsically motivated goal exploration for the automated discovery of patterns and behaviors in complex dynamical systems. We introduced a new algorithm (IMGEP-OGL) which is capable of learning unsupervised goal space representations during the exploration of an unknown system. Our results for Lenia, a high-dimensional complex dynamical system, show its superior performance over hand-defined goal spaces or random exploration. It shows the same performance as a learned goal space based on precollected data, showing that such a precollection of data is not necessary. We furthermore introduced the usage of CPPNs for the successful initialization of the intial states of the dynamical systems. Both advances allowed us to explore an unknown and high-dimensional dynamical system which shares many similarities with different physical or chemical systems.

7.4. Representation Learning

7.4.1. State Representation Learning in the Context of Robotics

Participants: David Filliat [correspondant], Natalia Diaz Rodriguez, Timothee Lesort, Antonin Raffin, René Traoré, Ashley Hill, Te Sun, Lu Lin, Guanghang Cai, Bunthet Say.

During the DREAM project, we participated in the development of a conceptual framework of open-ended lifelong learning [77] based on the idea of representational re-description that can discover and adapt the states, actions and skills across unbounded sequences of tasks.

In this context, State Representation Learning (SRL) is the process of learning without explicit supervision a representation that is sufficient to support policy learning for a robot. We have finalized and published a large state-of-the-art survey analyzing the existing strategies in robotics control [103], and we developed unsupervised methods to build representations with the objective to be minimal, sufficient, and that encode the relevant information to solve the task. More concretely, we used the developed and open sourced[0] the S-RL

[0] https://github.com/araffin/robotics-rl-srl

toolbox [137] containing baseline algorithms, data generating environments, metrics and visualization tools for assessing SRL methods. Part of this study is the [105] where we present a robustness analysis on Deep unsupervised state representation learning with robotic priors loss functions.

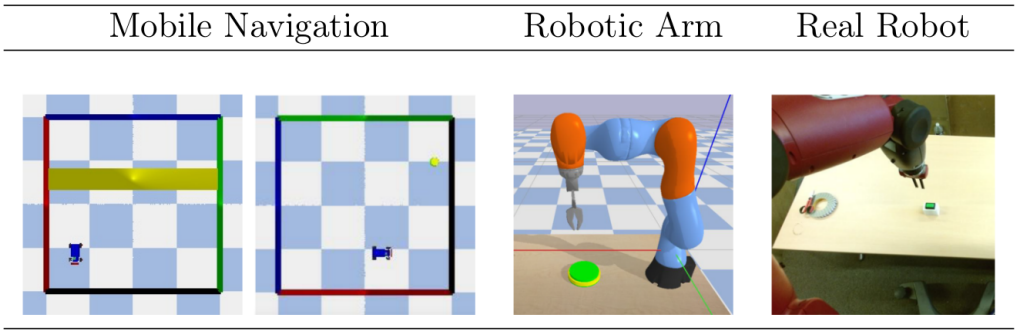


Figure 19. Environments and datasets for state representation learning.

The environments proposed in Fig. 19 are variations of two environments: a 2D environment with a mobile robot and a 3D environment with a robotic arm. In all settings, there is a controlled robot and one or more targets (that can be static, randomly initialized or moving). Each environment can either have a continuous or discrete action space, and the reward can be sparse or shaped, allowing us to cover many different situations.

The evaluation and visualization tools are presented in Fig. 20 and make it possible to qualitatively verify the learned state space behavior (e.g., the state representation of the robotic arm dataset is expected to have a continuous and correlated change with respect to the arm tip position).

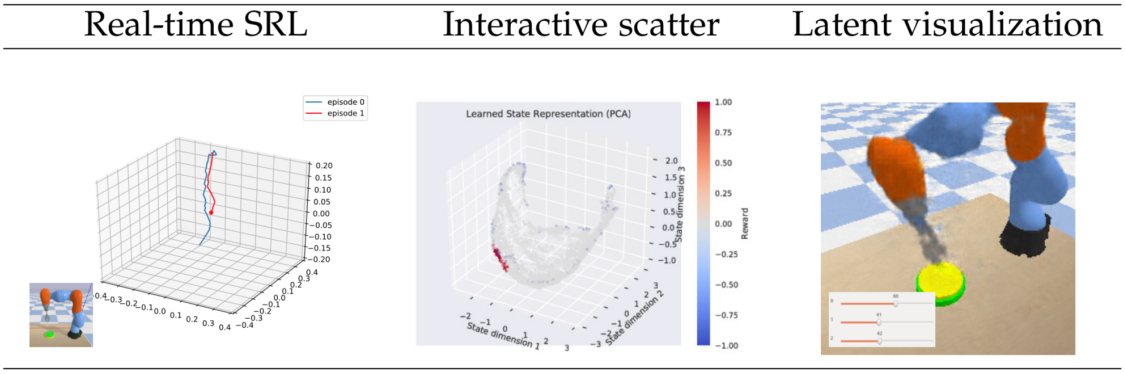


Figure 20. Visual tools for analysing SRL; Left: Live trajectory of the robot in the state space. Center: 3D scatter plot of a state space; clicking on any point displays the corresponding observation. Right: reconstruction of the point in the state space defined by the sliders.

We also proposed a new approach that consists of learning a state representation that is split into several parts where each optimizes a fraction of the objectives. In order to encode both target and robot positions, auto-encoders, reward and inverse model losses are used.

The latest work on decoupling feature extraction from policy learning, was presented at the SPIRL workshop at ICLR2019 in New Orleans, LA [138]. We assessed the benefits of state representation learning in goal based robotic tasks, using different self-supervised objectives.

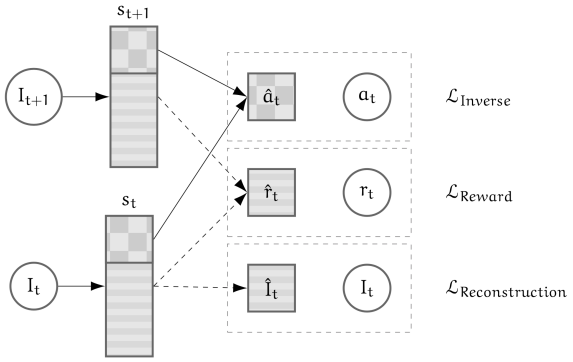


Figure 21. SRL Splits model: combines a reconstruction of an image I, a reward (r) prediction and an inverse dynamic models losses, using two splits of the state representation s. Arrows represent model learning and inference, dashed frames represent losses computation, rectangles are state representations, circles are real observed data, and squares are model predictions.

Because combining objectives into a single embedding is not the only option to have features that are *sufficient* to solve the tasks, by stacking representations, we favor *disentanglement* of the representation and prevent objectives that can be opposed from cancelling out. This allows a more stable optimization. Fig. 21 shows the split model where each loss is only applied to part of the state representation.

As using the learned state representations in a Reinforcement Learning setting is the most relevant approach to evaluate the SRL methods, we use the developed S-RL framework integrated algorithms (A2C, ACKTR, ACER, DQN, DDPG, PPO1, PPO2, TRPO) from Stable-Baselines [92], Augmented Random Search (ARS), Covariance Matrix Adaptation Evolutionary Strategy (CMA-ES) and Soft Actor Critic (SAC). Due to its stability, we perform extensive experiments on the proposed datasets using PPO and states learned with the approaches described in [137] along with ground truth (GT).

Table 22 illustrates the qualitative evaluation of a state space learned by combining forward and inverse models on the mobile robot environment. It also shows the performance of PPO algorithm based on the states learned by several baseline approaches.

We verified that our new approach (described in Task 2.1) makes it possible for reinforcement learning to converge faster towards the optimal performance in both environments with the same amount of budget timesteps. Learning curve in Fig. 23 shows that our unsupervised state representation learned with the split model even improves on the supervised case.

7.4.2. Continual learning

Participants: David Filliat [correspondant], Natalia Díaz Rodríguez, Timothee Lesort, Hugo Caselles-Dupré.

Continual Learning (CL) algorithms learn from a stream of data/tasks continuously and adaptively through time to better enable the incremental development of ever more complex knowledge and skills. The main problem that CL aims at tackling is catastrophic forgetting [115], i.e., the well-known phenomenon of a neural network experiencing a rapid overriding of previously learned knowledge when trained sequentially on new data. This is an important objective quantified for assessing the quality of CL approaches, however, the almost exclusive focus on catastrophic forgetting by continual learning strategies, lead us to propose a

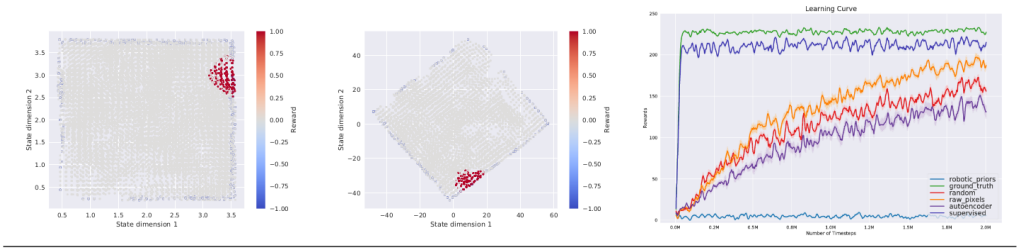


Figure 22. Ground truth states (left), states learned (Inverse and Forward) (center), and RL performance evaluation (PPO) (right) for different baselines in the mobile robot environment. Colour denotes the reward, red for positive, blue for negative and grey for null reward (left and center).

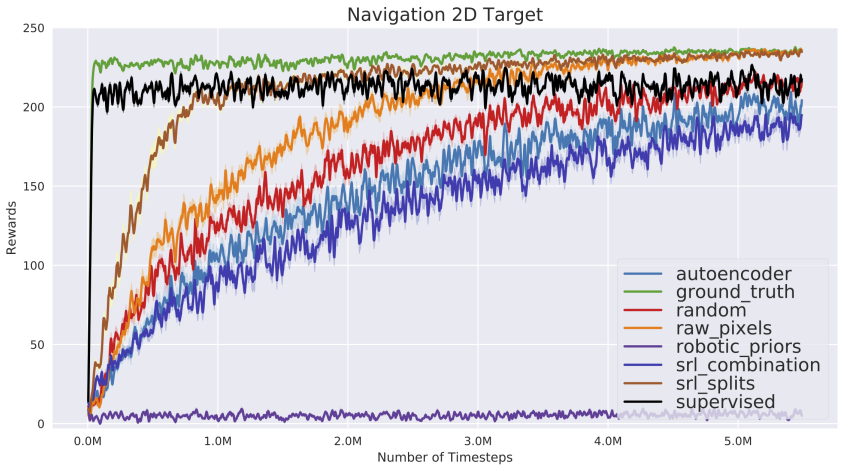


Figure 23. Performance (mean and standard error for 10 runs) for PPO algorithm for different state representations learned in Navigation 2D random target environment.

set of comprehensive, implementation independent metrics accounting for factors we believe have practical implications worth considering with respect to the deployment of real AI systems that learn continually, and in "Non-static" machine learning settings. In this context we developed a framework and a set of comprehensive metrics [78] to tame the lack of consensus in evaluating CL algorithms. They measure Accuracy (A), Forward and Backward (/*remembering*) knowledge transfer (FWT, BWT, REM), Memory Size (MS) efficiency, Samples Storage Size (SSS), and Computational Efficiency (CE). Results on iCIFAR-100 classification sequential class learning is in Table 24 .

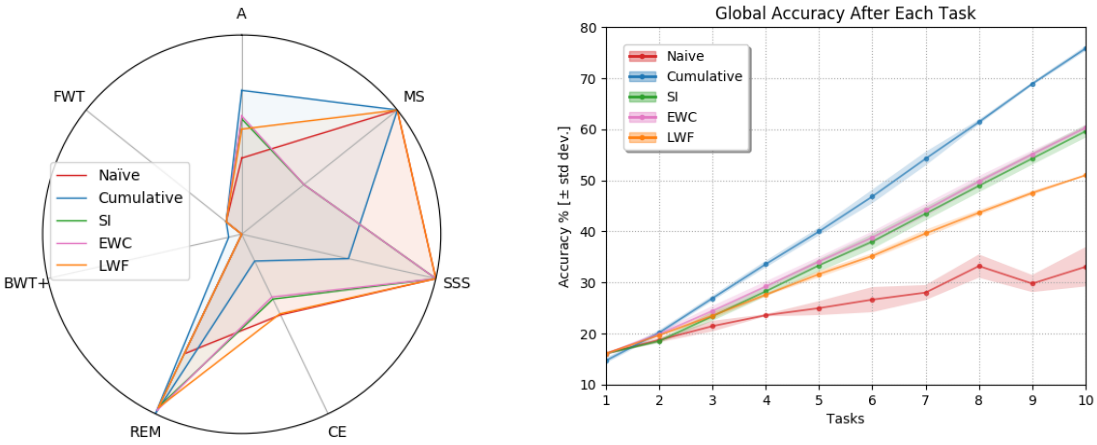


Figure 24. (left) Spider chart: CL metrics per strategy (larger area is better) and (right) Accuracy per CL strategy computed over the fixed test set.

Generative models can also be evaluated from the perspective of Continual learning which we investigated in our work [102]. This work aims at evaluating and comparing generative models on disjoint sequential image generation tasks. We study the ability of Generative Adversarial Networks (GANS) and Variational Auto-Encoders (VAEs) and many of their variants to learn sequentially in continual learning tasks. We investigate how these models learn and forget, considering various strategies: rehearsal, regularization, generative replay and fine-tuning. We used two quantitative metrics to estimate the generation quality and memory ability. We experiment with sequential tasks on three commonly used benchmarks for Continual Learning (MNIST, Fashion MNIST and CIFAR10). We found (see Figure 26) that among all models, the original GAN performs best and among Continual Learning strategies, generative replay outperforms all other methods. Even if we found satisfactory combinations on MNIST and Fashion MNIST, training generative models sequentially on CIFAR10 is particularly instable, and remains a challenge. This work has been published at the NIPS workshop on Continual Learning 2018.

Another extension of previous section on state representation learning (SRL) to the continual learning setting is in our paper [65]. This work proposes a method to avoid catastrophic forgetting when the environment changes using generative replay, i.e., using generated samples to maintain past knowledge. State representations are learned with variational autoencoders and automatic environment change is detected through VAE reconstruction error. Results show that using a state representation model learned continually for RL experiments is beneficial in terms of sample efficiency and final performance, as seen in Figure 26 . This work has been published at the NIPS workshop on Continual Learning 2018 and is currently being extended.

The experiments were conducted in an environment built in the lab, called Flatland [64]. This is a lightweight first-person 2-D environment for Reinforcement Learning (RL), designed especially to be convenient for Continual Learning experiments. Agents perceive the world through 1D images, act with 3 discrete actions,

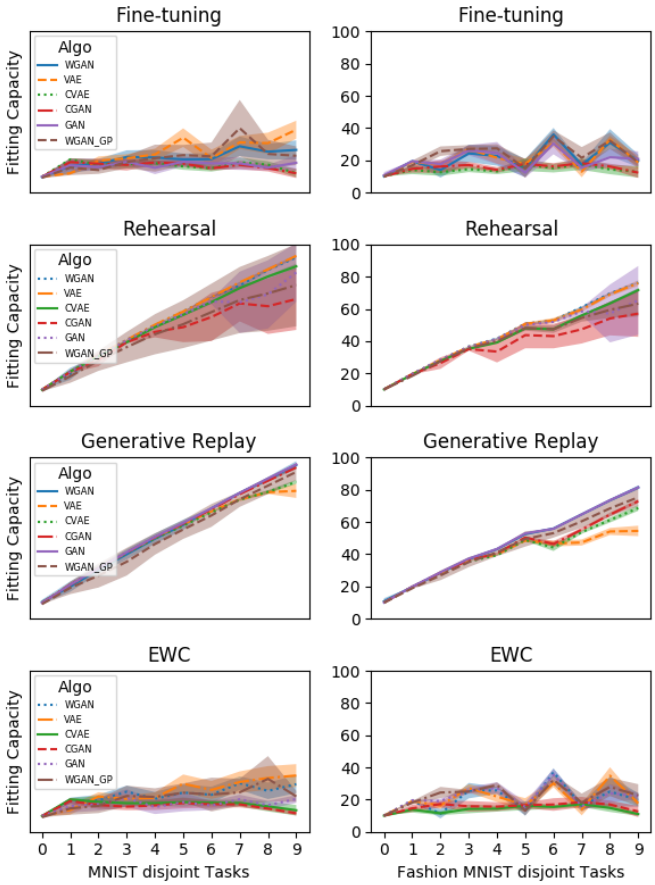


Figure 25. Means and standard deviations over 8 seeds of Fitting Capacity metric evaluation of VAE, CVAE, GAN, CGAN and WGAN. The four considered CL strategies are: Fine Tuning, Generative Replay, Rehearsal and EWC. The setting is 10 disjoint tasks on MNIST and Fashion MNIST.

and the goal is to learn to collect edible items with RL. This work has been published at the ICDL-Epirob workshop on Continual Unsupervised Sensorimotor Learning 2018, and was accepted as oral presentation.

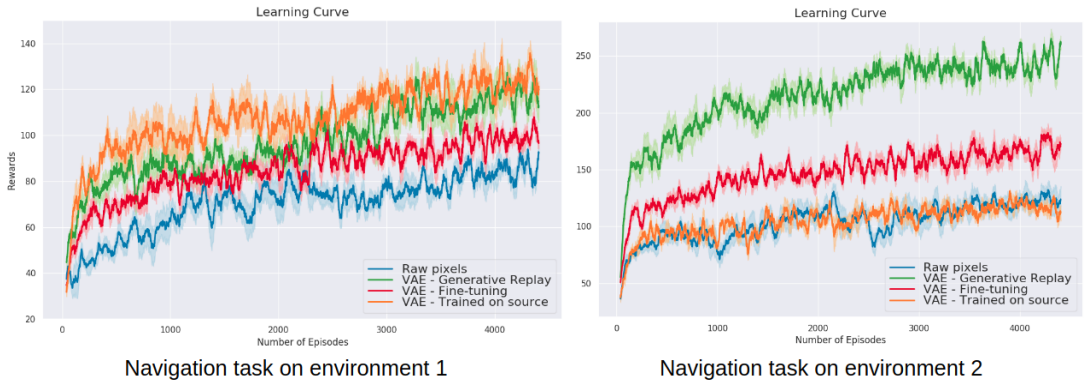


Figure 26. Mean reward and standard error over 5 runs of RL evaluation using PPO with different types of inputs. Fine-tuning and Generative Replay models are trained sequentially on the first and second environment, and then used to train a policy for both tasks. Generative Replay outperforms all other methods. It shows the need for continually learning features in State Representation Learning in settings where the environment changes.

In the last year, we published a survey on continual learning models, metrics and contributed a CL framework to categorize the approaches on this area [104]. Figure 27 shows the different approaches cited and the strategies proposed and a small subset of examples analyzed.

We also worked on validating a distillation approach for multitask learning in a continual learning reinforcement learning setting [152], [153].

Applying State Representation Learning (SRL) into a continual learning setting of reinforcement learning was possible by learning a compact and efficient representation of data that facilitates learning a policy. The proposed a CL algorithm based on distillation does not manually need to be given a task indicator at test time, but learns to infer the task from observations only. This allows to successfully apply the learned policy on a real robot.

We present 3 different 2D navigation tasks to a 3 wheel omni-directional robot to be learned to be solved sequentially. The robot has first access to task 1 only, and then to task 2 only, and so on. It should learn a single policy that solves all tasks and be applicable in a real life scenario. The robot can perform 4 high level discrete actions (move left/right, move up/down). The tasks where the method was validated are in Fig. 28 :

Task 1: Target Reaching (TR): Reaching a red target randomly positioned.

Task 2: Target Circling (TC): Circling around a fixed blue target.

Task 3: Target Escaping (TE): Escaping a moving robot.

DisCoRL (Distillation for Continual Reinforcement learning) is a modular, effective and scalable pipeline for continual RL. This pipeline uses policy distillation for learning without forgetting, without access to previous environments, and without task labels in order to transfer policies into real life scenarios [152]. It was presented as an approach for continual reinforcement learning that sequentially summarizes different learned policies into a dataset to distill them into a student model. Some loss in performance may occur while transferring knowledge from teacher to student, or while transferring a policy from simulation to real life. Nevertheless, the experiments show promising results when learning tasks sequentially, in simulated environments and real life settings.

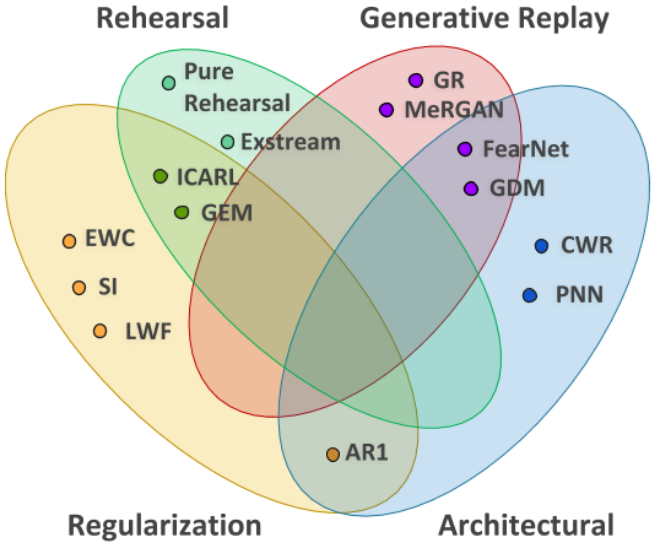


Figure 27. Venn diagram of some of the most popular CL strategies w.r.t the main approaches in the literature (CWR,PNN EWC, SI, LWF, ICARL, GEM, FearNet, GDM, ExStream, GR, MeRGAN, and AR1. Rehearsal and Generative Replay upper categories can be seen as a subset of replay strategies. Better viewed in color [104].

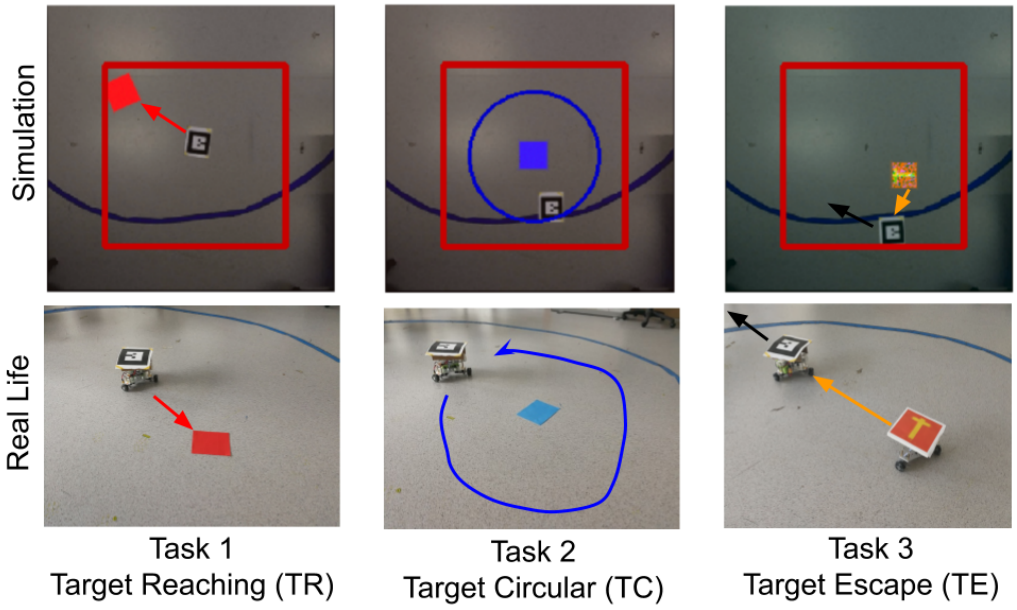


Figure 28. The three tasks, in simulation (top) and in real life (bottom), sequentially experienced. Learning is performed in simulation, the real life setting is only used at test time.

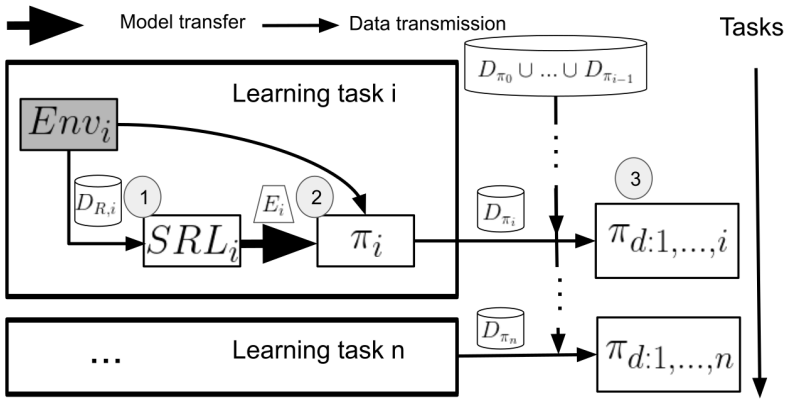


Figure 29. White cylinders are for datasets, gray squares for environments, and white squares for learning algorithms, whose name correspond to the model trained. Each task i is learned sequentially and independently by first generating a dataset $D_{R,i}$ with a random policy to learn a state representation with an encoder E_i with an SRL method (1), then we use E_i and the environment to learn a policy π_i in the state space (2). Once trained, π_i is used to create a distillation dataset D_{π_i} that acts as a memory of the learned behaviour. All policies are finally compressed into a single policy $\pi_{d:1,..,i}$ by merging the current dataset D_{π_i} with datasets from previous tasks $D_{\pi_1} \cup ... \cup D_{\pi_{i-1}}$ and using distillation (3).

The overview of DisCoRL full pipeline for Continual Reinforcement Learning is in Fig. 29 .

7.4.3. Disentangled Representation Learning for agents

Participants: Hugo Caselles-Dupré [correspondant], David Filliat.

Finding a generally accepted formal definition of a disentangled representation in the context of an agent behaving in an environment is an important challenge towards the construction of data-efficient autonomous agents. Higgins et al. (2018) recently proposed Symmetry-Based Disentangled Representation Learning, a definition based on a characterization of symmetries in the environment using group theory. We build on their work and make observations, theoretical and empirical, that lead us to argue that Symmetry-Based Disentangled Representation Learning cannot only be based on static observations: agents should interact with the environment to discover its symmetries.

Our research was published in NeuRIPS 2019 [32] at Vancouver, Canada.

7.5. Tools for Understanding Deep Learning Systems

7.5.1. Explainable Deep Learning

Participants: Natalia Díaz Rodríguez [correspondant], Adrien Bennetot.

Together with Segula Technologies and Sorbonne Université, ENSTA Paris has been working on eXplainable Artificial Intelligence (XAI) in order to make machine learning more interpretable. While opaque decision systems such as Deep Neural Networks have great generalization and prediction skills, their functioning does not allow obtaining detailed explanations of their behaviour. The objective is to fight the trade-off between performance and explainability by combining connectionist and symbolic paradigms [47].

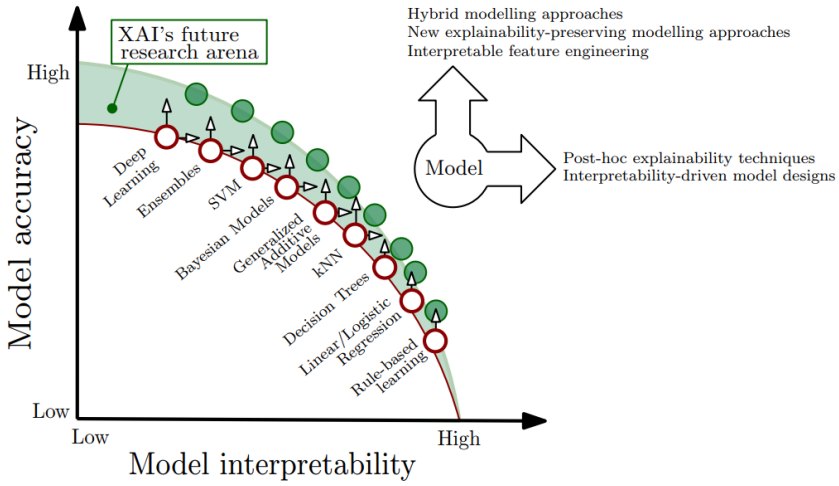


Figure 30. Trade-off between model interpretability and performance, and a representation of the area of improvement where the potential of XAI techniques and tools resides [46].

Broad consensus exists on the importance of interpretability for AI models. However, since the domain has only recently become popular, there is no collective agreement on the different definitions and challenges that constitute XAI. The first step is therefore to summarize previous efforts made in this field. We presented a taxonomy of XAI techniques in [46] and we are currently working on a prediction model that generates itself an explanation of its rationale in natural language while keeping performance as close as possible to the the state of the art [47].

7.5.2. Methods for Statistical Comparison of RL Algorithms

Participants: Cédric Colas [correspondant], Pierre-Yves Oudeyer, Olivier Sigaud.

Following a first article in 2018 [71], we pursued the objective of providing key tools to robustly compare reinforcement learning (RL) algorithms to practioners and researchers. In this year's extension, we compiled a hitchhicker's guide for statistical comparisons of RL algorithms. In particular, we provide a list of statistical tests adapted to compare RL algorithms and compare them in terms of false positive rate and statistial power. In particular, we study the robustness of these tests when their assumptions are violated (non-normal distributions of performances, different distributions, unknown variance, unequal variances etc). We provided an extended study using data from synthetic performance distributions, as well as empirical distributions obtained from running state-of-the-art RL algorithms (TD3 [84] and SAC [88]). From these results we draw a selection of advice for researchers. This study led to an article accepted at the ICLR conference workshop on Reproducibility in Machine Learning [34], to be sumitted to the Neural Networks journal.

7.5.3. Knowledge engineering tools for neural-symbolic learning

Participants: Natalia Díaz Rodríguez [correspondant], Adrien Bennetot.

Symbolic artificial intelligence methods are experiencing a come-back in order to provide deep representation methods the explainability they lack. In this area, a survey on RDF stores to handle ontology-based triple databases has been contributed [97], as well as the use of neural-symbolic tools that aim at integrating both neural and symbolic representations [58].

7.6. Applications in Educational Technologies

7.6.1. Machine Learning for Adaptive Personalization in Intelligent Tutoring Systems

Participants: Pierre-Yves Oudeyer [correspondant], Benjamin Clément, Didier Roy, Helene Sauzeon.

7.6.1.1. The Kidlearn project

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

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

The system is based on the combination of three approaches. First, it leverages recent models of intrinsically motivated learning by transposing them to active teaching, relying on empirical estimation of learning progress provided by specific activities to particular students. Second, it uses state-of-the-art Multi-Arm Bandit (MAB) techniques to efficiently manage the exploration/exploitation challenge of this optimization process. Third, it leverages expert knowledge to constrain and bootstrap initial exploration of the MAB, while requiring only coarse guidance information of the expert and allowing the system to deal with didactic gaps in its knowledge. The system was evaluated in several large-scale experiments relying on a scenario where 7-8 year old schoolchildren learn how to decompose numbers while manipulating money [68]. Systematic experiments were also presented with simulated students.

7.6.1.2. Kidlearn Experiments 2018-2019: Evaluating the impact of ZPDES and choice on learning efficiency and motivation

An experiment was held between mars 2018 and July 2019 in order to test the Kidlearn framework in classrooms in Bordeaux Metropole. 600 students from Bordeaux Metropole participated in the experiment. This study had several goals. The first goal was to evaluate the impact of the Kidlearn framework on motivation and learning compared to an Expert Sequence without machine learning. The second goal was to observe the impact of using learning progress to select exercise types within the ZPDES algorithm compared to a random policy. The third goal was to observe the impact of combining ZPDES with the ability to let children make different kinds of choices during the use of the ITS. The last goal was to use the psychological and contextual data measures to see if correlation can be observed between the students psychological state evolution, their profile, their motivation and their learning. The different observations showed that generally, algorithms based on ZPDES provided a better learning experience than an expert sequence. In particular, they provide a better motivating and enriching experience to self-determined students. Details of these new results, as well as the overall results of this project, are presented in Benjamin Clément PhD thesis [69] and are currently being processed to be published.

7.6.1.3. Kidlearn and Adaptiv'Math

The algorithms developed during the Kidlearn project and Benjamin Clement thesis [69] are being used in an innovation partnership for the development of a pedagogical assistant based on artificial intelligence intended for teachers and students of cycle 2. The algorithms are being written in typescript for the need of the project. The expertise of the team in creating the pedagogical graph and defining the graph parameters used for the algorithms is also a crucial part of the role of the team for the project. One of the main goal of the team here is to transfer technologies developed in the team in a project with the perspective of industrial scaling and see the impact and the feasibility of such scaling.

7.6.1.4. Kidlearn for numeracy skills with individuals with autism spectrum disorders

Few digital interventions targeting numeracy skills have been evaluated with individuals with autism spectrum disorder (ASD) [114]. Yet, some children and adolescents with ASD have learning difficulties and/or a significant academic delay in mathematics. While ITS are successfully developed for typically developed students to personnalize learning curriculum and then to foster the motivation-learning coupling, they are not or fewly proposed today to student with specific needs. The objective of this pilot study is to test the feasibility of a digital intervention using an STI with high school students with ASD and/or intellectual disability. This application (KidLearn) provides calculation training through currency exchange activities, with a dynamic exercise sequence selection algorithm (ZPDES). 24 students with ASD and/or DI enrolled in specialized classrooms were recruited and divided into two groups: 14 students used the KidLearn application, and 10 students received a control application. Pre-post evaluations show that students using KidLearn improved their calaculation performance, and had a higher level of motivation at the end of the intervention than the control group. These results encourage the use of an STI with students with specific needs to teach numeracy skills, , but need to be replicated on a larger scale. Suggestions for adjusting the interface and teaching method are suggested to improve the impact of the application on students with autism. (Paper is in progress).

7.6.2. Curiosity-driven interaction systems for education

Participants: Pierre-Yves Oudeyer, Hélène Sauzéon [correspondant], Mehdi Alami, Didier Roy, Edith Law.

Three studies have been developed and conducted to newly design curiosity-driven interaction systems aiming to foster learning performance across lifespan : the first two studies include children and the last one includes the older adults.

The first study regards a new interactive robotic system to foster curiosity-driven learning. This led to an article in CHI 2019 [29]. In this work, we explored whether a social peer robot's verbal expression of curiosity can be perceived by participants, produce emotional or behavioural contagion effects, and impact learning. In a between-subject experiment involving 30 participants, a peer robot was manipulated to verbally express: curiosity, curiosity plus rationale, or no curiosity (neutral), within the context of LinkedIt!, a cooperative game we designed for teaching students how to classify rocks. Results show that participants were able to reliably recognize curiosity in the robot and curious robots can be used to elicit significantly more curiosity-driven behaviours among participants.

The second study regards a new interactive educational application to foster curiosity-driven question-asking in children. This study has been performed during the Master 2 internship of Mehdi Alaimi co-supervised by H. Sauzéon, E. Law and PY Oudeyer. The paper submission to CHI'20 is just accepted in december 2019 (« Pedagogical Agents for Fostering Question-Asking Skills in Children »). It addresses a key challenge for 21st-century schools, i.e., teaching diverse students with varied abilities and motivations for learning, such as curiosity within educational settings. Among variables eliciting curiosity state, one is known as « knowledge gap », which is a motor for curiosity-driven exploration and learning. It leads to question-asking which is an important factor in the curiosity process and the construction of academic knowledge. However, children questions in classroom are not really frequent and don't really necessitate deep reasoning. Determined to improve children's curiosity, we developed a digital application aiming to foster curiosity-related question-asking from texts and their perception of curiosity. To assess its efficiency, we conducted a study with 95 fifth grade students of Bordeaux elementary schools. Two types of interventions were designed, one trying to focus children on the construction of low-level question (i.e. convergent) and one focusing them on high-level questions (i.e. divergent) with the help of prompts or questions starters models. We observed that both interventions increased the number of divergent questions, the question fluency performance, while they did not significantly improve the curiosity perception despite high intrinsic motivation scores they have elicited in children. The curiosity-trait score positively impacted the divergent question score under divergent condition, but not under convergent condition. The overall results supported the efficiency and usefulness of digital applications for fostering children's curiosity that we need to explore further.

Finaly, the third study investigates the role of intrinsic motivation in spatial learning in late adulthood [25]. We investigated age differences in memory for spatial routes that were either actively (i.e., intinsic

motivation condition) or passively (i.e., control condition) encoded. A series of virtual environments were created and presented to 20 younger (Mean age = 19.71) and 20 older (Mean age = 74.55) adults, through a cardboard viewer. During encoding, participants explored routes presented within city, park, and mall virtual environments, and were later asked to re-trace their travelled routes. Critically, participants encoded half the virtual environments by passively viewing a guided tour along a pre-selected route, and half through active exploration with volitional control of their movements by using a button press on the viewer. During retrieval, participants were placed in the same starting location and asked to retrace the previously traveled route. We calculated the percentage overlap in the paths travelled at encoding and retrieval, as an indicator of spatial memory accuracy, and examined various measures indexing individual differences in their cognitive approach and visuo-spatial processing abilities. Results showed that active navigation, compared to passive viewing during encoding, resulted in a higher accuracy in spatial memory, with the magnitude of this memory enhancement being significantly larger in older than in younger adults. Results suggest that age-related deficits in spatial memory can be reduced by active encoding. In other words, this means that conditions where intrinsic motivation is involved, reduce negative effects of aging on spatial learning.

7.6.3. Poppy Education: Designing and Evaluating Educational Robotics Kits

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

The Poppy Education project aims to create, evaluate and disseminate all-inclusive pedagogical kits, open-source and low cost, for teaching computer science and robotics in secondary education and higher education, scientific literacy centers and Fablabs.

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

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

- web interface connection (see figure 31)

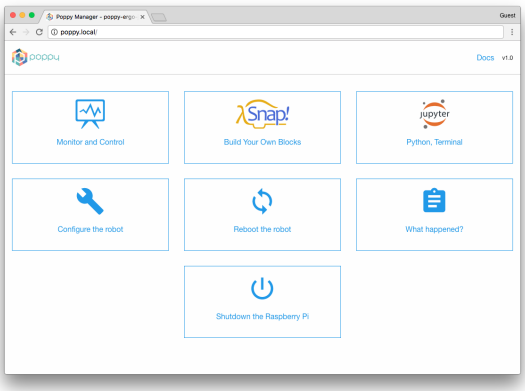


Figure 31. Home page on http://poppy.local

- Poppy Humanoid, a robust and complete robotics platform designed for genuine experiments in the real world and that can be adapted to specific user needs.
- Poppy Torso, a variant of Poppy Humanoid that can be easily installed on any flat support.

- Ergo Jr, a robotic arm. Durable and inexpensive, it is perfect to be used in class. It can be programmed in Python, directly from a web browser, using Ipython notebooks (an interactive terminal, in a web interface for the Python Programming Language).
- Snap. The visual programming system Snap (see figure 32), which is a variant of Scratch. Its features allow a thorough introduction of information technology. Several specific "blocks" have been developed for this.

Figure 32. The visual programming system Snap

- C++, Java, Matlab, Ruby, Javascript, etc. thanks to a REST API that allows you to send commands and receive information from the robot with simple HTTP requests.
- Virtual robots (Poppy Humanoid, Torso and Ergo) can be simulated with the free simulator V-REP (see figure 33). It is possible in the classroom to work on the simulated model and then allow students to run their program on the physical robot.
- Virtual robots (Poppy Ergo) can also be simulated with a 3D web viewer (see figure 34).

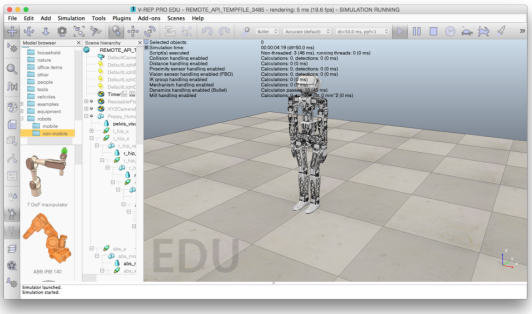


Figure 33. V-rep

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

The robots are designed with the final users in mind. The pedagogical tools of the project (robots and resources) are being created directly with the users and evaluated in real life by experiments. So teachers and researchers co-create activities, test them with students in class-room, share their experience and develop the platform as needed [126].

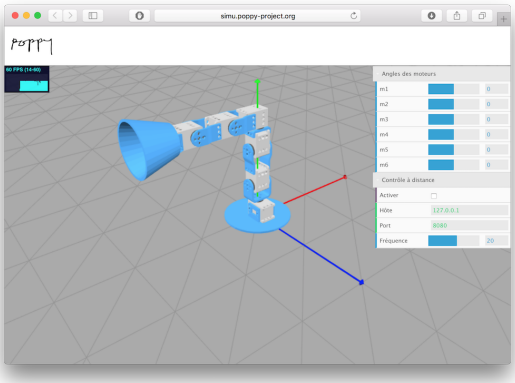


Figure 34. 3D viewer

The activities were designed mainly with Snap! and Python. Most activities use Poppy Ergo Jr, but some use Poppy Torso (mostly in higher school due to its cost).
The pedagogical experiments in classroom carried out during the first year of the project notably allowed to create and experiment many robotic activities. These activities are designed as pedagogical resources introducing robotics. The main objective of the second year was to make all the activities and resources reusable (with description, documentation and illustration) easily and accessible while continuing the experiments and the diffusion of the robotic kits.

Figure 35. Experiment robots and pedagogical activities in classroom

- Pedagogical working group : the teacher partners continued to use the robots in the classroom and to create and test new classroom activities. We organized some training to help them to discover and learn how to use the robotics platform. Also, an engineer of the Poppy Education team went to visit the teachers in their school to see and to evaluate the pedagogical tools (robots and activities) in a real context of use.

Five meetings have been organized during the year including all teachers part of the project as well as the Poppy Education team in order to exchange about their experience using the robots as a pedagogical tool, to understand their need and to get some feedback from them. This is helping us to understand better the educational needs, to create and improve the pedagogical tools.
You can see the videos of pedagogical robotics activities here:
https://www.youtube.com/playlist?list=PLdX8RO6QsgB7hM_7SQNLvyp2QjDAkkzLn

7.6.3.2. Pedagogical documents and resources

- We continued to improve the documentation of the robotic platform Poppy (https://docs.poppy-project.org/en/) and the documentation has been translated into French (https://docs.poppy-project.org/fr/.
  We configured a professional platform to manage the translation of the documentation (https://crowdin.com/project/poppy-docs. This platforms allows anybody to participate in the translation of the documentation to the language of their choice.
- To complete the pedagogical booklet [125] that provides guided activities and small challenges to become familiar with Poppy Ergo Jr robot and the Programming language Snap! (https://hal.inria.fr/hal-01384649/document) we provided a list of Education projects. Educational projects have been written for each activity carried out and tested in class. Each project has its own web page including resources allowing any teacher to carry out the activity (description, pedagogical sheet, photos / videos, pupil's sheet, teacher's sheet with correction etc.).
  The activities are available here:
  https://www.poppy-education.org/activites/activites-lycee
  The pedagogical activities are also available on the Poppy project forum where everyone is invited to comment and create new ones:
  https://forum.poppy-project.org/t/liste-dactivites-pedagogiques-avec-les-robots-poppy/2305

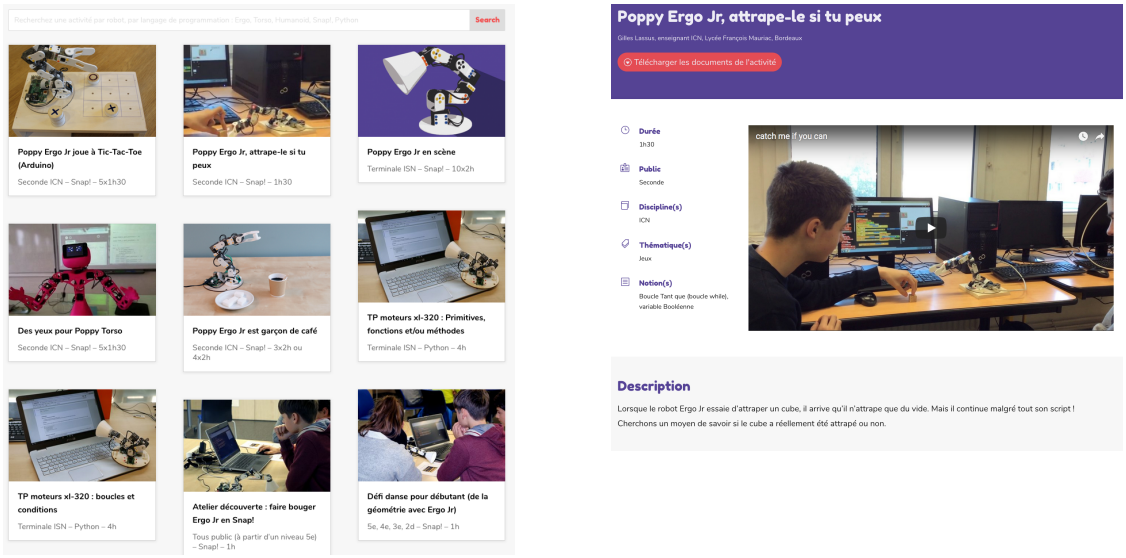


Figure 36. Open-source educational activities with Poppy robots are available on Poppy-Education.org

- A FAQ have been written with the most frequents questions to help the users: https://www.poppy-education.org/aide/
- A website has been created to present the project and to share all resources and activities.
  https://www.poppy-education.org/

7.6.3.3. Evaluation of the pedagogical kits

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

- Population

  Our sample is made up of 28 teachers and 146 students from the region Nouvelle Aquitaine. Each subject completed an online survey in June 2017. Here, we study several groups of individuals: teachers and students. Among the students we are interested in those who practiced classroom activities with the Ergo Jr kit during the school year 2016 - 2017 (N = 68) (age = 16, std = 2.44). Among these students, 37 where High School students following the "Computer Science and Digital Sciences" stream (BAC S option ISN), 12 followed the stream "Computer and Digital Creation" (BAC S option ICN) and 18 where in Middle School.

  Among the 68 students, 13 declared having used the educational booklet provided in the kit and 16 declared having used other robotic kits. Concerning the time resource dedicated to activities with the robot, 30 students declared having spent less than 6 hours, 22 declared between 6 and 25 hours, and 16 declared having spent more than 25 hours.

  have practiced less than 6 hours of activity with the robot (N = 30), between 6 and 25 hours (N = 22) or more than 25 hours (N = 16); having built the robot (N = 12); have used the visual programming language Snap! (N = 46), the language of Python textual programming (N = 21), both (N = 8) or none (N = 9), it should be noted that these two languages are directly accessible via the main interface of the robot.

- Evaluation of the tool

  We have selected two standardized surveys dealing with this issue: SUS (The Systeme Usability Scales) [62] and The AttrakDiff [100]. These two surveys are complementary and allow to identify the design problems and to account for the perception of the user during the activities. The results of these surveys are available in the article (in French) [76] published at the conference Didapro (Lausanne Feb, 2018). Figures 37 and 38 show the averages of the 96 respondents (68 students + 28 teachers) for each of the 10 statements from the SUS and 28 pairs of antonyms to be scored on a scale of 1 to 5 and a 7-point scale, respectively.

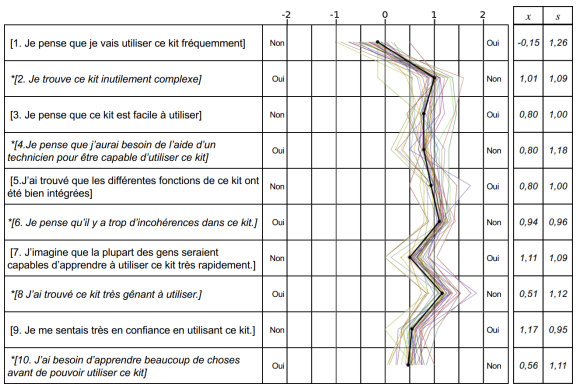


Figure 37. Result of SUS survey

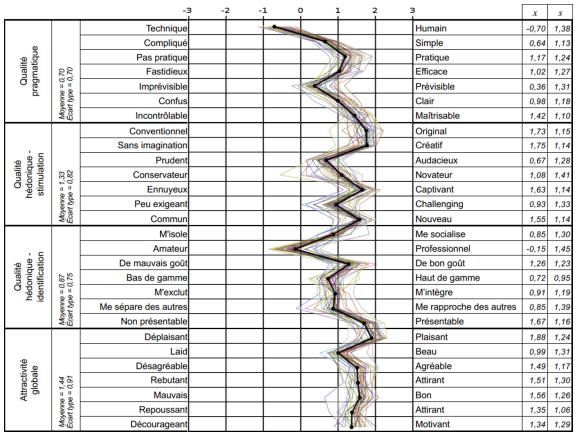


Figure 38. Result of AttrakDiff survey

- Evaluation of impact on learner

  One of the objectives of the integration of digital sciences in school is to allow students to have a better understanding of the technological tools that surround them daily (i.e. web, data, algorithm, connected object, etc.). So, we wanted to measure how the practice of activities with ErgoJr robot had changed this apprehension; especially towards robots. For that, we used a standardized survey: "attitude towards robot" *EuroBarometer 382* originally distributed in 2012 to more than 1000 people in each country of the European Union. On the one hand, we sought to establish whether there had been a change in response between 2012 and 2017, and secondly whether there was an impact on the responses of 2017 according to the participation, or not, in educational activities with ErgoJr robot. The analysis of the results is in progress and will be published in 2019.

- Web page for the experimentations

  To facilitate the storage of documents, their availability, and to highlight some information and news, a page dedicated to the experimentations is now available on the website. https://www.poppy-education.org/evaluation/

7.6.3.4. Partnership on education projects

- Ensam

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

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

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

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

Poppy Education). During the project, they met the choreographer Eric Minh Cuong Castaing and the engineer Segonds Theo (Inria - Poppy Education).

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

- DANE

  The Academic Delegation for Digital Educational is in charge of supporting the development of digital uses for pedagogy. It implements the educational digital policy of the academy in partnership with local authorities. She accompanies institutions daily, encourages innovations and participates in their dissemination.

- RobotCup Junior

  RoboCupJunior OnStage invites teams to develop a creative stage performance using autonomous robots that they have designed, built and programmed. The objective is to create a robotic performance between 1 to 2 minutes that uses technology to engage an audience. The challenge is intended to be open-ended. This includes a whole range of possible performances, for example dance, storytelling, theatre or an art installation. The performance may involve music but this is optional. Teams are encouraged to be as creative, innovative and entertaining, in both the design of the robots and in the design of the overall performance.

7.7. Other applications

7.7.1. Applications in Robotic myoelectric prostheses

Participants: Pierre-Yves Oudeyer [correspondant], Aymar de Rugy, Daniel Cattaert, Mick Sebastien.

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

7.7.1.1. Reachy, a 3D-printed Human-like Robotic Arm as a Test Bed for Prosthesis Control Strategies

To this day, despite the increasing motor capability of robotic prostheses, elaborating efficient control strategies is still a key challenge for their design. To provide an amputee with efficient ways to drive a prosthesis, this task requires thorough testing prior to integration into finished products. To preserve consistency with prosthetic applications, employing an actual robot for such testing requires it to show human-like features. To fulfill this need for a biomimetic test platform, we developed the Reachy robotic platform, a seven-joint human-like robotic arm that can emulate a prosthesis. Although it does not include an articulated hand and is therefore more suitable for studying reaching than manipulation, a robotic hand from available research prototypes could be integrated to Reachy. Its 3D-printed structure and off-the-shelf actuators make it inexpensive relatively to the price of a genuine prosthesis. Using an open-source architecture, its design makes it broadly connectable and customizable, so it can be integrated into many applications. To illustrate how Reachy can connect to external devices, we developed several proofs of concept where it is operated with various control strategies, such as tele- operation or vision-driven control. In this way, Reachy can help researchers to develop and test innovative control strategies on a human-like robot.

7.7.2. Ship Motion estimation from sea wave vision

Participants: David Filliat [correspondant], Natalia Díaz Rodríguez, Zhi Zhou, Manuel Cortés-Batet, Nazar-Mykola Kaminskyi.

Together with Naval Group, ENSTA Paris has been working on a set of software tools for simulating sea waves and motion estimation from images. The objective is predicting variables of interest in order to compensate the position and inclination of large boats at deep sea, seconds ahead of time to preserve stability. Work being currently done in partnership with Abo Akademi University (Turku, Finland) will validate the soon to be published Blender wave generator and machine learning algorithms, with real data gathered from the Baltic Sea archipelago.

MANAO Project-Team

7. New Results

7.1. Analysis and Simulation

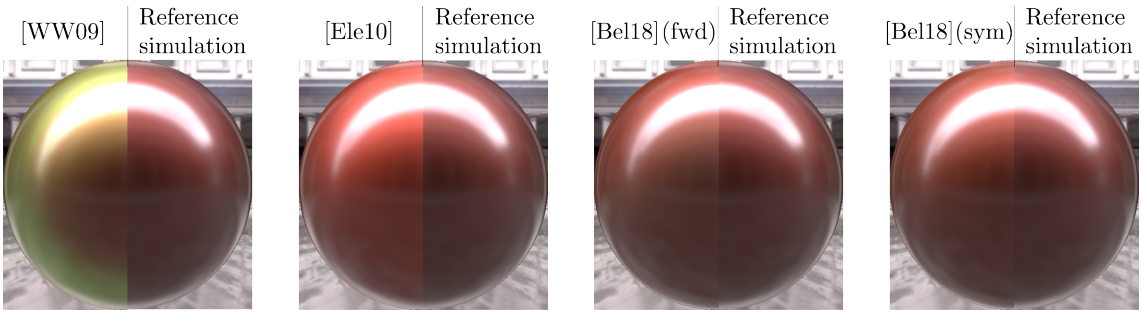


Figure 8. We study how the approximations made by layered material models impact their accuracy, and ultimately material appearance. Here we compare four models side by side with our reference simulation on a frosted metal – one of the 60 material configurations we have considered in our study. This specific choice is particularly problematic for the model of Weidlich and Wilkie [WW09], which creates oddly-colored reflections away from normal incidence. The variant of Elek [Ele10] is devoid of these artefacts, but clearly overestimates the intensity of the metallic base. Belcour's models [Bel18] (forward and symmetric) produce more accurate results, even though the intensity of the metallic base remains slightly higher. They still deviate from the reference simulation, especially at grazing angles as seen for instance at the bottom of the spheres. Our analysis in BRDF (and BTDF) space provides explanations for such departures from the reference.

7.1.1. Numerical Analysis of Layered Materials Models ,

Publications: [12], [14]

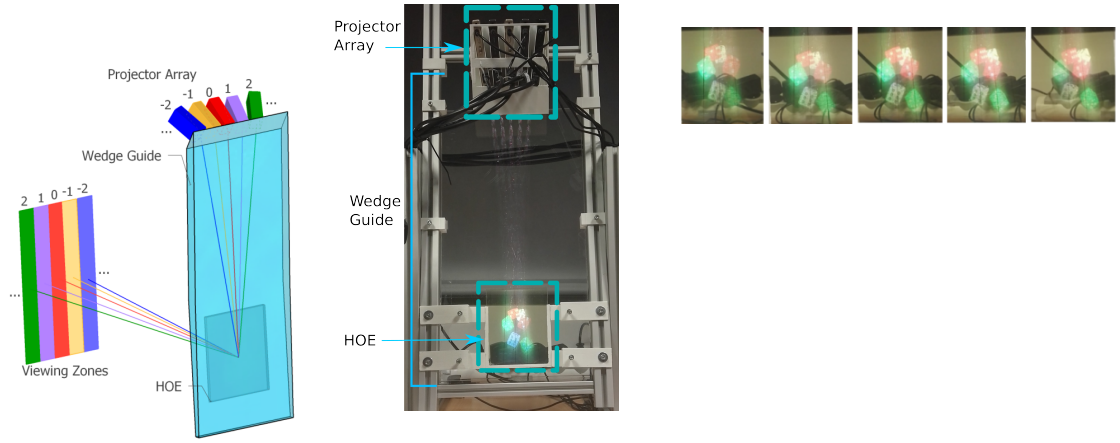
Most real-world materials are composed of multiple layers, whose physical properties impact the appearance of objects. The accurate reproduction of layered material properties is thus an important part of physically-based rendering applications. Since no exact analytical model exists for arbitrary configurations of layer stacks, available models make a number of approximations. In this technical report, we propose to evaluate these approximations with a numerical approach: we simulate BRDFs and BTDFs for layered materials in order to compare existing models against a common reference. More specifically, we consider 60 layered material configurations organized in three categories: plastics, metals and transparent slabs. Our results (see Figure 8) show that: (1) no single model systematically outperforms the others on all categories; and (2) significant discrepancies remain between simulated and modeled materials. We analyse the reasons for these discrepancies and introduce immediate corrections that improve models accuracy with little effort. Finally, we provide a few challenging cases for future layered material models.

7.1.2. A systematic approach to testing and predicting light-material interactions

Publication: [11]

Photographers and lighting designers set up lighting environments that best depict objects and human figures to convey key aspects of the visual appearance of various materials, following rules drawn from experience. Understanding which lighting environment is best adapted to convey which key aspects of materials is an important question in the field of human vision. The endless range of natural materials and lighting environments poses a major problem in this respect. Here we present a systematic approach to make this problem tractable for lighting–material interactions, using optics-based models composed of canonical lighting and material modes. In two psychophysical experiments, different groups of inexperienced observers judged the material qualities of the objects depicted in the stimulus images. In the first experiment, we took photographs of real objects as stimuli under canonical lightings. In a second experiment, we selected three generic natural lighting environments on the basis of their predicted lighting effects and made computer renderings of the objects. The selected natural lighting environments have characteristics similar to the canonical lightings, as computed using a spherical harmonic analysis. Results from the two experiments correlate strongly, showing (a) how canonical material and lighting modes associate with perceived material qualities; and (b) which lighting is best adapted to evoke perceived material qualities, such as softness, smoothness, and glossiness. Our results demonstrate that a system of canonical modes spanning the natural range of lighting and materials provides a good basis to study lighting–material interactions in their full natural ecology.

7.2. From Acquisition to Display



(a) Illustration (b) Photograph (c) 5 views

Figure 9. The autostereoscopic transparent display: light beams from multiple laser beam steering picoprojectors are coupled into a transparent wedge guide, and then the light from each projector is redirected to separate viewing zones using a transparent HOE.

7.2.1. Autostereoscopic transparent display using a wedge light guide and a holographic optical element

Publications: [8], [13]

We designed and developed a novel transparent autostereoscopic display consisting of laser picoprojectors, a wedge light guide, and a custom holographic optical element (HOE). Such a display can superimpose 3D data on the real world without any wearable.

The principle of our display, as depicted in Figure 9 , is to couple beams from multiple laser beam steering picoprojectors into a transparent wedge guide and then to redirect each beam to separate viewing zones using a transparent HOE. The HOE is wavelength-multiplexed for full-color efficiency, but only one angular grating is recorded and multiple viewing zones are reconstructed with several projector positions due to the high angular bandwidth. Our current prototype has 5 views but is theoretically able to generate 9 views. The views are located 50cm in front of the display, they are 3cm wide and 10cm high. These values are fixed once the HOE is recorded; they result from our choices and can be changed in the recording step.

This display has great potential for augmented reality applications such as augmented exhibitions in museums or shops, head-up displays for vehicles or aeronautics, and industrial maintenance, among others.

7.2.2. Wedge cameras for minimally invasive archaeology

Publication: [9]

Acquiring images of archaeological artifacts is an essential step for the study and preservation of cultural heritage. In constrained environments, traditional acquisition techniques may fail or be too invasive. We present an optical device including a camera and a wedge waveguide that is optimized for imaging within confined spaces in archeology. The major idea is to redirect light by total internal reflection to circumvent the lack of room, and to compute the final image from the raw data. We tested various applications onsite during an archaeological mission in Medamoud (Egypt). Our device was able to successfully record images of the underground from slim trenches, including underwater trenches, and between rocks composing a wall temple. Experts agreed that the acquired images were good enough to get useful information that cannot be obtained as easily with traditional techniques.

7.2.3. Study of contrast variations with depth in focused plenoptic cameras

Publication: [10]

A focused plenoptic camera has the ability to record and separate spatial and directional information of the incoming light. Combined with the appropriate algorithm, a 3D scene could be reconstructed from a single acquisition, over a depth range called plenoptic depth-of-field. We have studied the contrast variations with depth as a way to assess plenoptic depth-of-field. We take into account the impact of diffraction, defocus, and magnification on the resulting contrast. We measure the contrast directly on both simulated and acquired images. We demonstrate the importance of diffraction and magnification in the final contrast. Contrary to classical optics, the maximum of contrast is not centered around the main object plane, but around a shifted position, with a fast and nonsymmetric decrease of contrast.

7.2.4. Unifying the refocusing algorithms and parameterizations for traditional and focused plenoptic cameras

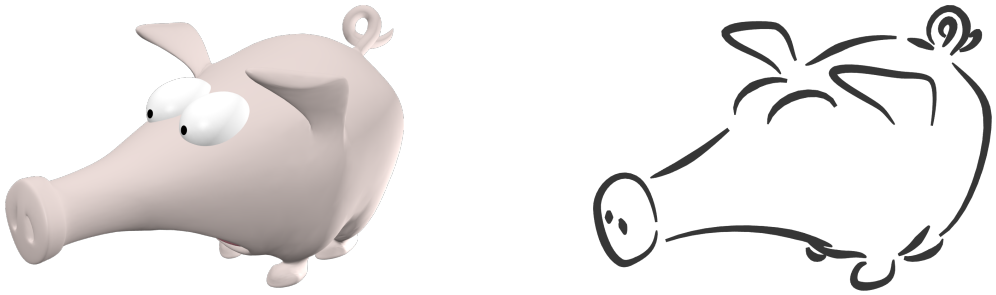
Publication: [16]

We propose a unique parameterization of the light rays in a plenoptic setup, allowing the development of a unique refocusing algorithm valid for any plenoptic configurations, based on this parameterization. With this method we aim at refocusing images at any distances from the camera, without previous discontinuity due to change of optical configuration. We aim to obtain reconstructed images visually similar to the results of the other algorithms, but quantitatively more accurate.

7.3. Rendering, Visualization and Illustration

7.3.1. Line drawings from 3D models: a tutorial

Publication: [7]



(a) 3D object with diffuse shading (b) Stylized curves

Figure 10. The occluding contours of the 3D model "Origins of the Pig" by Keenan Crane, shown in (a) with diffuse shading, are depicted in (b) with calligraphic brush strokes.

This tutorial describes the geometry and algorithms for generating line drawings from 3D models, focusing on occluding contours. The geometry of occluding contours on meshes and on smooth surfaces is described in detail, together with algorithms for extracting contours, computing their visibility, and creating stylized renderings and animations. Exact methods and hardware-accelerated fast methods are both described, and the trade-offs between different methods are discussed. The tutorial brings together and organizes material that, at present, is scattered throughout the literature. It also includes some novel explanations, and implementation tips. A thorough survey of the field of non-photorealistic 3D rendering is also included, covering other kinds of line drawings and artistic shading (Figure 10). In addition, we provide an interactive viewer at https://benardp.github.io/contours_viewer/.

7.4. Editing and Modeling

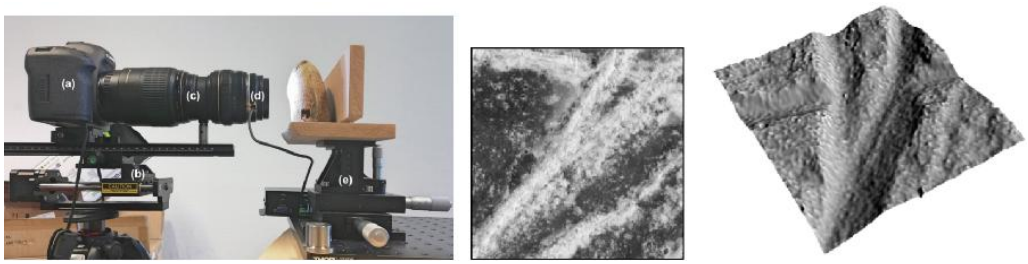


Figure 11. Left: our low-cost depth-from-focus acquisition setup. Right: result of our reconstruction algorithm for the sample shown in the middle image. The width of grooves are about 500 micrometers.

7.4.1. Depth from focus stacks at micrometer scale

In this work we designed a low-cost acquisition setup and a new algorithm for the digitalization of micro reliefs. The setup is based on a common digital camera equipped with a special assembly of different lenses designed to enable a $\times 2$ magnification factor with a very shallow depth of field (fig. 11). A micro-metric motor rail allows us to acquire dense focus stacks with depth information that can be reconstructed through image processing and analysis techniques. To enhance the accuracy of this reconstruction step, we designed

novel focus estimators as well as novel focus-point analysis algorithms exploiting novel 3D invariants. Our initial results show that we are able to reconstruct depth maps with sub-step length accuracy.

POTIOC Project-Team

7. New Results

7.1. Mixed-Reality System for Collaborative Learning at School

Participants: Philippe Giraudeau, Théo Segonds, Solène Lambert, Martin Hachet

External collaborators: Université de Lorraine

Traditional computer systems based on the WIMP paradigm (Window, Icon, Menu, Pointer) have shown potential benefits at school (e.g. for web browsing). On the other hand, they are not well suited as soon as hands-on and collaborative activities are targeted. To face this problem, we have designed and developed CARDS, a Mixed-Reality system that combines together physical and digital objects in a seamless workspace to foster active and collaborative learning (Figure 3). In [23], we describe the design process based on a participatory approach with researchers, teachers, and pupils. We then present and discuss the results of a user study that tends to show that CARDS has a good educational potential for the targeted activities.

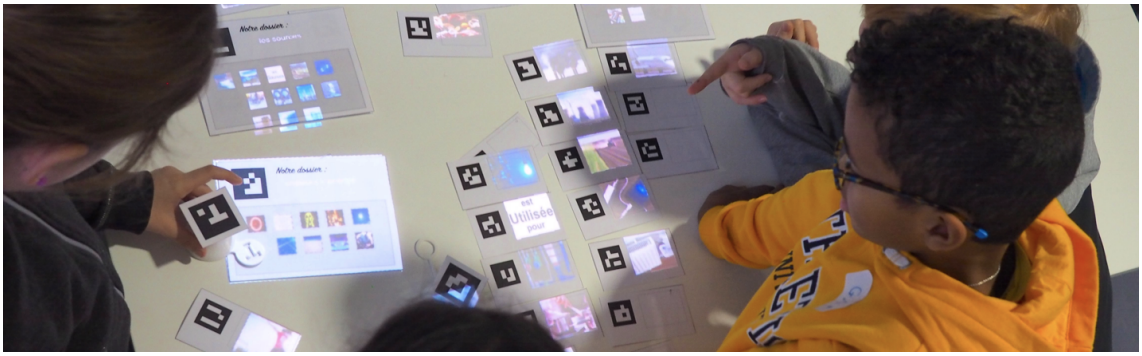


Figure 3. CARDS: Collaborative Activities based on the Real and the Digital Superimposition.

7.2. DroneSAR: Extending Physical Spaces in Spatial Augmented Reality using Projection on a Drone

Participants: Rajkumar Darbar, Joan Sol Roo, Thibaut Lainé, Martin Hachet

Spatial Augmented Reality (SAR) transforms real-world objects into interactive displays by projecting digital content using video projectors. SAR enables co-located collaboration immediately between multiple viewers without the need to wear any special glasses. Unfortunately, one major limitation of SAR is that visual content can only be projected onto its physical supports. As a result, displaying User Interfaces (UI) widgets such as menus and pop-up windows in SAR is very challenging. We are trying to address this limitation by extending SAR space in mid-air. We propose Drone-SAR, which extends the physical space of SAR by projecting digital information dynamically on the tracked panels mounted on a drone (see Figure 4 . DroneSAR is a proof of concept of novel SAR User Interface (UI) , which provides support for 2D widgets (i.e., label, menu, interactive tools, etc.) to enrich SAR interactive experience. We describe this concept, as well as implementation details of our proposed approach in [22].

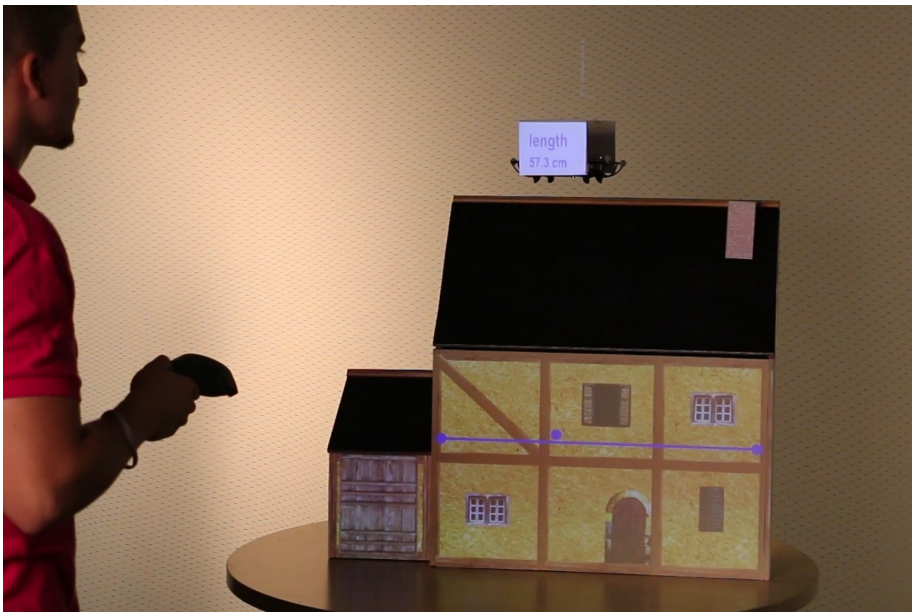


Figure 4. DroneSAR: Projection on a drone allows us to extend the physical space for interacting spatial augmented reality.

7.3. Tangible and modular devices for supporting communication

Participants: Joan Sol Roo, Pierre-Antoine Cinquin, Martin Hachet

External collaborators: Ullo

Our physiological activity reflects our inner workings. However, we are not always aware of it in full detail. Physiological devices allow us to monitor and create adaptive systems and support introspection. Given that these devices have access to sensitive data, it is vital that users have a clear understanding of the internal mechanisms (extrospection), yet the underlying processes are hard to understand and control, resulting in a loss of agency. In this work, we focus on bringing the agency back to the user, by using design guidelines based on principles of honest communication and driven by positive activities. To this end, we conceived a tangible, modular approach for the construction of physiological interfaces (see Figure 5). We are exploring the potential of such an approach with a set of examples, supporting introspection, dialog, music creation, and play.

7.4. Accessible Interactive Audio-Tactile Drawings

Participants: Lauren Thevin, Anke Brock, Martin Hachet

External collaborators: CNRS, Univesité Paul Sabatier, ENAC

Interactive tactile graphics have shown a true potential for people with visual impairments, for instance for acquiring spatial knowledge. Until today, however, they are not well adopted in real-life settings (e.g. special education schools). One obstacle consists in the creation of these media, which requires specific skills, such as the use of vector-graphic software for drawing and inserting interactive zones, which is challenging for stakeholders (social workers, teachers, families of people with visual impairments, etc.). We explored how a Spatial Augmented Reality approach can enhance the creation of interactive tactile graphics by sighted users. We developed the system using a participatory design method. A user study showed that the augmented

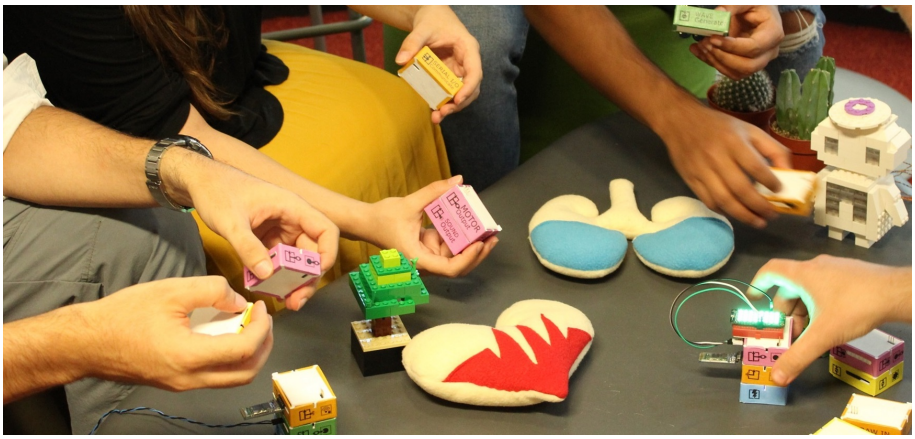


Figure 5. modular bricks that support the easy creation and interfacing with physiological applications.

reality device allowed stakeholders (N=28) to create interactive tactile graphics more efficiently than with a regular vector-drawing software (baseline), independently of their technical background. This work illustrated in Figure 6 is described in [29].

Following the same approach, we are currently exploring how physical board games can be moved into accessible ones for people with visual impairments.

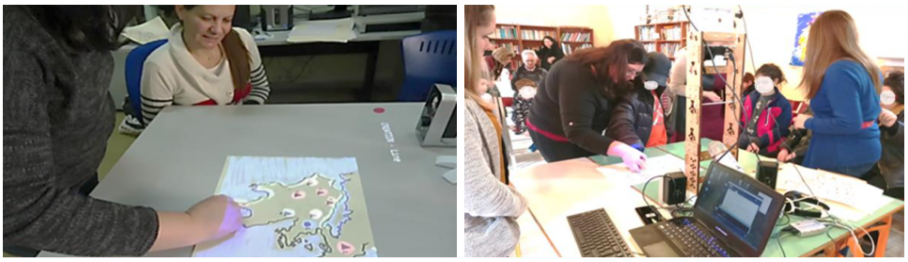


Figure 6. Combining touch and audio for accessible drawings at school.

7.5. Accessibility of e-learning systems

Participants: Pierre-Antoine Cinquin, Damien Caselli and Pascal Guitton

External collaborators: Hélène Sauzéon

In 2019, we continued to work on new digital teaching systems such as MOOCs. Unfortunately, accessibility for people with disabilities is often forgotten, which excludes them, particularly those with cognitive impairments for whom accessibility standards are fare from being established. We have shown in [12] that very few research activities deal with this issue.

In past years, we have proposed new design principles based on knowledge in the areas of accessibility (Ability-based Design and Universal Design), digital pedagogy (Instruction Design with functionalities that reduce the cognitive load : navigation by concept, slowing of the flow…), specialized pedagogy (Universal Design for Learning, eg, automatic note-taking, and Self Determination Theory, e.g., configuration of the interface according to users needs and preferences) and psychoducational interventions (eg, support the joint teacher-learner attention), but also through a participatory design approach involving students with disabilities and experts in the field of disability (Figure 7). From these framework, we have designed interaction features which have been implemented in a specific MOOC player called Aïana. Moreover, we have produced a MOOC on digital accessibility which is published on the national MOOC platform (FUN) using Aïana (4 sessions since 2016 with more than 11 000 registered participants). https://mooc-francophone.com/cours/mooc-accessibilite-numerique/. Our first field studies demonstrate the benefits of using Aïana for disabled participants [32].

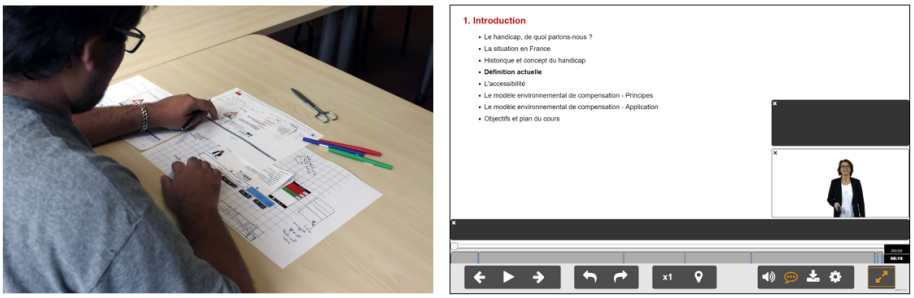


Figure 7. Design of Aïana MOOC player.

7.6. A hybrid setup for an artistic experience

Participants: Vincent Da Silva Pinto, Martin Hachet

External collaborators: Léna d'azy

In 2019, we have worked with the stenographer Cécile Léna to conceive and build a hybrid setup that combines a real physical mock-up and a virtual environments. This allows the participants to explore the sky by pointing at planets with a telescope in miniature, and observing a virtual view of the pointed planet by looking through an immersive stereoscopic installation (Figure 8).

7.7. Mental state decoding from EEG signals using robust machine learning

Participants: Aurélien Appriou, Smeethy Pramij, Khadijeh Sadatnejad, Aline Roc, Léa Pillette, Thibaut Monseigne, Fabien Lotte

External collaborators: Andrzej Cichocki, Pierre-Yves Oudeyer, Edith Law, Jessie Ceha, Frédéric Dehais, Alban Duprès, Sarah Blum, Nicolas Drougard, Sébastien Scannella, Raphaëlle N. Roy

Modern machine learning algorithms to classify cognitive and affective states from electroencephalography signals: Estimating cognitive or affective states from brain signals is a key but challenging step in the creation of passive brain-computer interface (BCI) applications. So far, estimating mental workload or emotions from EEG signals is only feasible with modest classification accuracies, thus leading to unreliable neuroadaptive applications. However, recent machine learning algorithms, notably Riemannian geometry based classifiers (RGC) and convolutional neural networks (CNN), have shown to be promising for other BCI systems, e.g., motor imagery-BCIs. However, they have not been formally studied and compared together for

Figure 8. Echelles célestes allows the participant to explore the sky.

cognitive or affective states classification. We have thus explored such machine learning algorithms, proposed new variants of them, and benchmarked them with classical methods to estimate both mental workload and affective states (Valence/Arousal) from EEG signals. We studied these approaches with both subject-specific and subject-independent calibration, to go towards calibration-free systems. Our results suggested that a CNN obtained the highest mean accuracy, although not significantly so, in both conditions for the mental workload study, followed by RGCs. However, this same CNN underperformed in both conditions for the emotion data set, a data set with little training data. On the contrary, RGCs proved to have the highest mean accuracy with the Filter Bank Tangent Space classifier (FBTSC) we introduced in this paper. Our results thus contributed to improve the reliability of cognitive and affective states classification from EEG. They also provide guidelines about when to use which machine learning algorithm. This work was just accepted for publication in the IEEE System Man and Cybernetics magazine.

Towards decoding curiosity from Brain and physiological signals: The neurophysiological mechanisms underlying curiosity and intrinsic motivation are currently not well understood. However, being able to identify objectively, from neurophysiological signals, the curiosity level of a user, would bring a very useful tool both to neuroscientists and psychologists, to understand curiosity deeper, as well as to designers of human-computer interaction, in order to trigger curiosity or to adapt an interaction to the curiosity levels of its users. A first step to do that, is to collect neurophysiological signals during known states of curiosity, in order to develop signal processing/machine learning tools to recognize those states from such signals. We designed and ran an experimental protocol to measure both brain activity through Electroencephalography (EEG) and physiological responses (heart rate, skin conductance, Electrocardiogram) when subjects were induced into different states of curiosity. During the experiment, fun facts were presented to subjects to induce different levels of curiosity. We obtained those fun facts using the Google functionality "I'm feeling curious" as well as crowdsourcing. A subject could choose a fun fact that made him curious, and push forward with a 4-to-10 questions chain on this theme. For each question on a given theme, a subject could choose to reveal the answer (interpreted as a curious state) or to skip it (interpreted as a non-curious state). Skipping an answer will automatically break the chain and will point the subject to the next fun fact. Neurophysiological signals were collected from 28 subjects, between a question and the choice of revealing the answer. Then those subjects graded the question on a 1-to-7 curiosity level scale. We are currently working on finding biological markers of curiosity by analyzing the collected signals using machine learning.

Channel Selection over Riemannian manifold with non-stationarity consideration for Brain-Computer interface applications: EEG signals are essentially non-stationary. Such non-stationarities, including cross-trial, cross-session, and cross-subject non-stationarities, are the result of various neurophysiological and extra-physiological causes. Such non-stationarities lead to variations in BCI users' performance. To handle this problem, we designed and compared multiple criteria for selecting EEG channels over the Riemannian manifold, for EEG classification. These criteria aim to promote EEG covariance matrix classifiers to generalize well by considering EEG data non-stationarity. Our approach consists of both increasing the discriminative information between classes over the manifold and reducing the dispersion within classes. We also reduce the influence of outliers in both discriminative and dispersion measures. The criteria were evaluated on EEG signals recorded from a tetraplegic subject and dataset IVa from BCI competition III. Experimental evidences confirm that considering the dispersion within each class as a measure for quantifying the effects of non-stationarity and removing the most affected channels can improve BCI performance (see Figure 9). This work was submitted to ICASSP 2020.

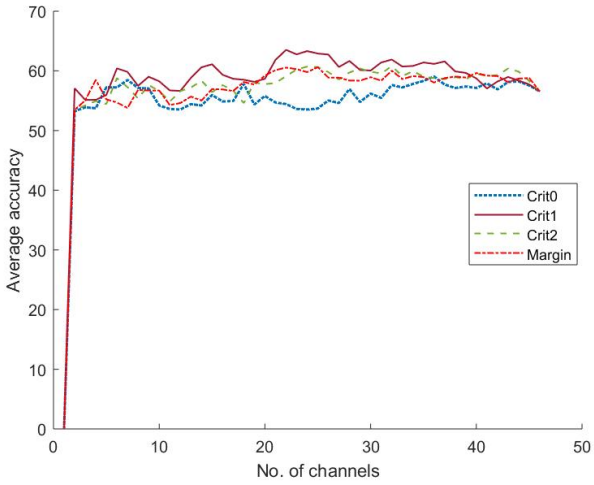


Figure 9. Results of BCI classification accuracy for various number of selected channels, for different channel selection algorithms. Our proposed algorithms - termed here Crit1, Crit2 and Margin - all improved upon the state-of-the-art (Crit0).

Monitoring Pilot's Mental Workload Using ERPs and Spectral Power with a Six-Dry-Electrode EEG System in Real Flight Conditions: Recent technological progress has allowed the development of low-cost and highly portable brain sensors such as pre-amplified dry-electrodes to measure cognitive activity out of the laboratory. This technology opens promising perspectives to monitor the "brain at work" in complex real-life situations such as while operating aircraft. However, there is a need to benchmark these sensors in real operational conditions. We therefore designed a scenario in which twenty-two pilots equipped with a six-dry-electrode EEG system had to perform one low load and one high load traffic pattern along with a passive auditory oddball. In the low load condition, the participants were monitoring the flight handled by a flight instructor, whereas they were flying the aircraft in the high load condition. At the group level, statistical analyses disclosed higher P300 amplitude for the auditory target (Pz, P4 and Oz electrodes) along with higher alpha band power (Pz electrode), and higher theta band power (Oz electrode) in the low load condition as compared to the high load one. Single trial classification accuracy using both event-related potentials and event-related frequency features at the same time did not exceed chance level to discriminate the two load conditions. However, when considering only the frequency features computed over the continuous signal,

classification accuracy reached around 70% on average. This study demonstrates the potential of dry-EEG to monitor cognition in a highly ecological and noisy environment, but also reveals that hardware improvement is still needed before it can be used for everyday flight operations. This work was published in the journal Sensors in [13].

7.8. Understand and modeling Mental-Imagery BCI user training

Participants: Camille Benaroch, Aline Roc, Léa Pillette, Fabien Lotte

External collaborators: Camille Jeunet, Bernard N'Kaoua

Computational models of performance: Mental-Imagery based Brain-Computer Interfaces (MI-BCIs) make use of brain signals produced during mental imagery tasks to control a computerised system. The currently low reliability of MI-BCIs could be due, at least in part, to the use of inappropriate user-training procedures. In order to improve these procedures, it is necessary first to understand the mechanisms underlying MI-BCI user-training, notably through the identification of the factors influencing it. Thus, we first aimed at creating a statistical model that could explain/predict the performances and the progression of MI-BCI users using their traits (e.g., personality). We used the data of 42 participants (i.e., 180 MI-BCI sessions in total) collected from three different studies that were based on the same MI-BCI paradigm. We used machine learning regressions with a leave-one-subject-out cross validation to build different models. A first results showed that using the users' traits only may enable the prediction of performances for a single multiple-session experiment, but might not be sufficient to reliably predict MI-BCI performances across different experiments. A second result showed that using the users' traits and the users' past performances may enable the prediction of the progression of one user as reliable models were found for two of the three studies. Part of this work was published at the International Graz BCI conference in [21].

Would Motor-Imagery based BCI user training benefit from more women experimenters?: Throughout MI-BCI use, human supervision (e.g., experimenter or caregiver) plays a central role. While providing emotional and social feedback, people present BCIs to users and ensure smooth users' progress with BCI use. Though, very little is known about the influence experimenters might have on the results obtained. Such influence is to be expected as social and emotional feedback were shown to influence MI-BCI performances. Furthermore, literature from different fields showed an experimenter effect, and specifically of their gender, on experimental outcome. We assessed the impact of the interaction between experimenter and participant gender on MI-BCI performances and progress throughout a session. Our results revealed an interaction between participants gender, experimenter gender and progress over runs. It seems to suggest that women experimenters may positively influence participants' progress compared to men experimenters. This work was published at the International Graz BCI conference in [28].

7.9. Redefining and optimizing BCI user training tasks, stimulations and feedback

Participants: Jelena Mladenovic, Smeethy Pramij, Léa Pillette, Romain Sabau, Fabien Lotte

External collaborators: Jérémy Frey, Jérémie Mattout, Matteus Joffily, Emmanuel Maby, Bertrand Glize, Bernard N'Kaoua, Pierre-Alain Joseph, Camille Jeunet, Roger N'Kambou, Boris Mansencal

Active inference as a unifying, generic and adaptive framework for a P300-based BCI: We proposed the use of a generic, computational framework – Active (Bayesian) Inference to automatically lead the adaptation process in a P300 speller BCI. It adapts through the use of a probabilistic model of the user built upon user's reactions to flashing/spelled letters. Using such observations, at each iteration it updates its beliefs about user intentions, and converges towards a predefined goal, i.e. correctly spelled letters. Active Inference is a recent computational neuroscience approach that models learning and decision making of the brain. As such, by endowing such model to the BCI machine, it enables the machine to adapt in a similar fashion as the brain would. We demonstrate an implementation of Active Inference on a simulated P300-Speller BCI, with real EEG data from 18 subjects. Results demonstrate the ability of Active Inference to yield a significant increase

in bit rate (17%) over state-of-the-art approaches. This work was published in Journal of Neural Engineering in [18].

Towards adaptive and adapted difficulty for MI-BCI user training: We investigated the relationship between the human factors and BCI performance during MI-BCI training. Additionally, we investigated the influence of user personality traits and states on learning the MI skill, i.e., evolution of performance over a session. We conducted a MI experiment in which we influence the user through task difficulty. We acquire data to build a predictive model that could unveil which kind of task is optimal for what kind of user. Moreover, depending on what we set to be predicted, be it a flow state or performance, it can serve as a guide for overall adaptation, i.e., it can serve as an optimization criteria to wager between user experience and system accuracy for instance. We then used priors on user traits and states acquired from the prediction models to perform a simple adaptive method which provides optimal task difficulty to each user. To demonstrate the usefulness of the model for maximizing performance, we perform a simulation using real data from the MI-BCI experiment mentioned above. This work was presented in the PhD thesis of Jelena Mladenovic, that was successfully defended on September 10th, 2019.

Impact of MI-BCI feedback for post-stroke and neurotypical people: We investigated how the modality of the feedback could be adapted to the learners. First, based on a review of the literature, we argued that somatosensory abilities of post-stroke patients have not, but should be, taken into account for BCI-based motor therapies. Indeed, somatosensory abilities play an important role in motor rehabilitation in general, and in BCI-based therapies in particular. It is assumed that during BCI based therapies the co-activation of ascending (i.e., somatosensory) and descending (i.e., sensorimotor) networks enables significant functional motor improvement, together with significant sensorimotor-related neurophysiological changes. Somatosensory abilities seem essential for the patients to benefit from the feedback provided by the BCI system. Yet, around half of post-stroke patients suffer from somatosensory deficits. We hypothesize that these deficits alter their ability to benefit from BCI-based therapies. Our review of the literature on BCI-based motor rehabilitation post-stroke of 14 randomized clinical trials indicates that somatosensory abilities were rarely considered and/or reported. Only two studies over the fourteen reported using them as inclusion/exclusion criteria. Though, none of these two studies reported how they assess the somatosensory abilities. We argue that assessing the somatosensory abilities of the patients is necessary to avoid any bias and enable reliable comparison between-subject and between-study. It could also be leveraged to improve our understanding of the underlying mechanisms of motor recovery and adapt the therapy to the patients' abilities.

Our review of the literature also informed us that a multimodal feedback composed of both somatosensory and visual feedback enables better performances than an unimodal visual feedback, at least in the short term. Though, the long term influence of such feedback remained unknown. Therefore, we assessed the long term effects of a multimodal feedback composed of both vibrotactile and realistic visual stimulations (presented in [43], see also Figure 10), and a unimodal feedback with only realistic visual stimulations. We found that the beneficial impact of a multimodal feedback composed of both visual and somatosensory stimulation compared to a visual feedback alone remains true even for long term training, which had not been tested before. Also, the order of presentation of the different modalities of feedback might have an influence. Using an unimodal visual feedback only seems to be better suited for untrained participants. We hypothesis that integrating information arising from two modalities of feedback while performing the task could be particularly challenging for a novice learner. Both these works were presented in the PhD thesis of Léa Pillette, that was successfully defended on December 16th, 2019.

A physical learning companion for Mental-Imagery BCI User Training: We continued our work on PEANUT, that we designed, implemented and tested, and which is the first learning companion dedicated to providing social presence and emotional feedback during MI-BCI user training. PEANUT provided social presence and emotional support, depending on the performance and progress of the user, through interventions combining both pronounced sentences and facial expressions. It was designed based on the literature, data analyses and user-studies. We notably conducted several online user surveys to identify the desired characteristics of our learning companion in terms of appearance and supporting speech content. From the results of these surveys we notably deduced which should be the characteristics (personal/non-personal, exclamatory/declarative) of the sentences to be used depending on the performance and progression of a

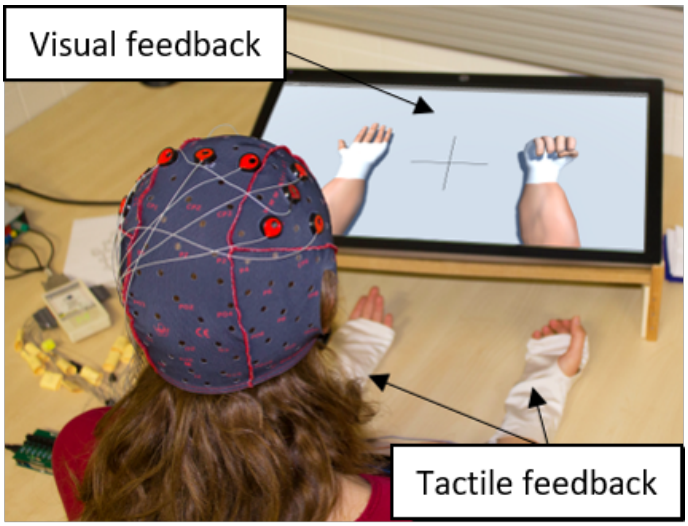


Figure 10. Our multimodal (realistic visual + vibrotactile) feedback for BCI training.

learner. We also found that eyebrows could increase expressiveness of cartoon-like faces. Then, once this companion was implemented, we evaluated it during real online MI-BCI use. We found that non-autonomous people, who are more inclined to work in a group and are usually disadvantaged when using MI-BCI, were advantaged compared to autonomous people when PEANUT was present with an increase of 3.9% of peak performances. Furthermore, in terms of user experience, PEANUT seems to have improved how people felt about their ability to learn and memorize how to use an MI-BCI by 7.4%, which is a dimension of the user experience we assessed. This work was published in the International Journal of Human-Computer Studies in [19].

Long-term mental imagery BCI training of a tetraplegic user: We participated to the Cybathlon BCI series 2019 competition in Graz (https://www.tugraz.at/institutes/ine/graz-bci-conferences/8th-graz-bci-conference-2019/cybathlon-bci-series-2019/), as team NITRO (Neurotechnology Inria Team Racing Odyssey), during which we trained a tetraplegic user over several months, with up to 3 training sessions per week, to learn to control a 4-class and self-paced mental imagery BCI connected to a racing video game (see Figure 11). This training and the resulting BCI design used several of our recent research and development works, notably new OpenViBE development on the feedback, progressive user training and adaptive Riemannian EEG classifiers [41].

7.10. Turning negative into positives! Exploiting “negative” results in Brain-Computer Interface research

Participants: Fabien Lotte

External collaborators: Laurent Bougrain, Ricardo Chavarriaga, Camille Jeunet, Karen Dijkstra, Andrea Kübler, Reinhold Scherer, Moritz Grosse-Wentrup, Natalie Dayan, Dave Thompson, Md Rakibul Mowla

Results that do not confirm expectations are generally referred to as “negative” results. While essential for scientific progress, they are too rarely reported in the literature - BCI research is no exception. This led us to organize a workshop on BCI negative results during the 2018 International BCI meeting. First, we demonstrated why (valid) negative results are useful, and even necessary for BCIs. These results can be used to confirm or disprove current BCI knowledge, or to refine current theories. Second, we provided concrete

Figure 11. BCI-based control of a racing video game by a tetraplegic user during the Cybathlon BCI series in Graz, Austria.

examples of such useful negative results, including the limits in BCI-control for complete locked-in users and predictors of motor imagery BCI performances. Finally, we suggested levers to promote the diffusion of (valid) BCI negative results, e.g., promoting hypothesis-driven research using valid statistical tools, organizing special issues dedicated to BCI negative results, or convincing institutions and editors that negative results are valuable. This work was published in the Brain-Computer Interface journal, in [16].

7.11. Speed of rapid serial visual presentation of pictures, numbers and words affects event-related potential-based detection accuracy

Participants: Fabien Lotte

External collaborators: Stephanie Lees, Paul McCullagh, Liam Maguire, Damien Coyle

Rapid serial visual presentation (RSVP) based brain-computer interfaces (BCIs) can detect target images among a continuous stream of rapidly presented images, by classifying a viewer's event related potentials (ERPs) associated with the target and non-targets images. Whilst the majority of RSVP-BCI studies to date have concentrated on the identification of a single type of image, namely pictures, here we studied the capability of RSVP-BCI to detect three different target image types: pictures, numbers and words. The impact of presentation duration (speed) i.e., 100-200ms (5-10Hz), 200-300ms (3.3-5Hz) or 300-400ms (2.5-3.3Hz), was also investigated. 2-way repeated measure ANOVA on accuracies of detecting targets from non-target stimuli (ratio 1:9) measured via area under the receiver operator characteristics curve (AUC) for N=15 subjects revealed a significant effect of factor Stimulus-Type (pictures, numbers, words) ($F(2,28) = 7.243$, $p = 0.003$) and for Stimulus-Duration ($F(2,28) = 5.591$, $p = 0.011$). Furthermore, there was an interaction between stimulus type and duration: $F(4,56) = 4.419$, $p = 0.004$). The results indicated that when designing RSVP-BCI paradigms, the content of the images and the rate at which images are presented impact on the accuracy of detection and hence these parameters are key experimental variables in protocol design and applications, which apply RSVP for multimodal image datasets. This work was published in IEEE Transactions on Neural Systems and Rehabilitation Engineering, in [15].

7.12. Design and preliminary study of a neurofeedback protocol to self-regulate an EEG marker of drowsiness

Participants: Thibaut Monseigne, Fabien Lotte

External collaborators: Stéphanie Bioulac, Pierre Philip, Jean-Arthur Micoulaud-Franchi

Neurofeedback (NF) consists in using EEG measurements to guide users to perform a cognitive learning using information coming from their own brain activity, by means of a real-time sensory feedback (e.g., visual or auditory). Many NF approaches have been studied to improve attentional abilities, notably for attention deficit hyper activity disorder. However, to our knowledge, no NF solution has been proposed to specifically reduce drowsiness. Thus, we propose an EEG-NF solution to train users to self-regulate an EEG marker of drowsiness, and evaluate it with a preliminary study. Results with five healthy subjects showed that three of them could learn to self-regulate this EEG marker with a relatively short number of NF sessions (up to 8 sessions of 40 min). This work was published at the International Graz BCI conference in [27].