RESEARCH CENTER

FIELD
Digital Health, Biology and Earth

Activity Report 2013

Section New Results

Edition: 2014-03-20

COMPUTATIONAL BIOLOGY

COMPUTATIONAL MEDICINE AND NEUROSCIENCES

COMPUTATIONAL NEUROSCIENCE AND MEDECINE

EARTH, ENVIRONMENTAL AND ENERGY SCIENCES

MODELING AND CONTROL FOR LIFE SCIENCES

ABS Project-Team

6. New Results

6.1. Modeling Interfaces and Contacts

Docking, scoring, interfaces, protein complexes, Voronoi diagrams, arrangements of balls.

The work undertaken in this vein in 2013 will be finalized in 2014.

6.2. Modeling Macro-molecular Assemblies

Macro-molecular assembly, reconstruction by data integration, proteomics, modeling with uncertainties, curved Voronoi diagrams, topological persistence.

6.2.1. Connectivity Inference in Mass Spectrometry based Structure Determination

Participants: Frédéric Cazals, Deepesh Agarwal.

In collaboration with J. Araujo, and C. Caillouet, and D. Coudert, and S. Pérennes, from the COATI project-team (Inria-CNRS).

In [14], we consider the following MINIMUM CONNECTIVITY INFERENCE problem (MCI), which arises in structural biology: given vertex sets $V_i \subseteq V, i \in I$, find the graph $G = (V, E)$ minimizing the size of the edge set E, such that the sub-graph of G induced by each V_i is connected. This problem arises in structural biology, when one aims at finding the pairwise contacts between the proteins of a protein assembly, given the lists of proteins involved in sub-complexes. We present four contributions.

First, using a reduction of the set cover problem, we establish that MCI is APX-hard. Second, we show how to solve the problem to optimality using a mixed integer linear programming formulation (`MILP`). Third, we develop a greedy algorithm based on union-find data structures (`Greedy`), yielding a $2(\log_2 |V| + \log_2 \kappa)$-approximation, with κ the maximum number of subsets V_i a vertex belongs to. Fourth, application-wise, we use the MILP and the greedy heuristic to solve the aforementioned connectivity inference problem in structural biology. We show that the solutions of `MILP` and `Greedy` are more parsimonious than those reported by the algorithm initially developed in biophysics, which are not qualified in terms of optimality. Since MILP outputs a set of optimal solutions, we introduce the notion of *consensus solution*. Using assemblies whose pairwise contacts are known exhaustively, we show an almost perfect agreement between the contacts predicted by our algorithms and the experimentally determined ones, especially for consensus solutions.

6.3. Algorithmic Foundations

Computational geometry, Computational topology, Voronoi diagrams, α-shapes, Morse theory.

6.3.1. Greedy Geometric Algorithms for Collection of Balls, with Applications to Geometric Approximation and Molecular Coarse-Graining

Participants: Frédéric Cazals, Tom Dreyfus.

In collaboration with S. Sachdeva (Princeton University, USA), and N. Shah (Carnegie Mellon University, USA).

Choosing balls to best approximate a 3D object is a non trivial problem. To answer it, in [18], we first address the *inner approximation* problem, which consists of approximating an object $\mathcal{F}_{\mathcal{O}}$ defined by a union of n balls with $k < n$ balls defining a region $\mathcal{F}_{\mathcal{S}} \subset \mathcal{F}_{\mathcal{O}}$. This solution is further used to construct an *outer approximation* enclosing the initial shape, and an *interpolated approximation* sandwiched between the inner and outer approximations.

The inner approximation problem is reduced to a geometric generalization of weighted max k-cover, solved with the greedy strategy which achieves the classical $1 - 1/e$ lower bound. The outer approximation is reduced to exploiting the partition of the boundary of $\mathcal{F}_{\mathcal{O}}$ by the Apollonius Voronoi diagram of the balls defining the inner approximation.

Implementation-wise, we present robust software incorporating the calculation of the exact Delaunay triangulation of points with degree two algebraic coordinates, of the exact medial axis of a union of balls, and of a certified estimate of the volume of a union of balls. Application-wise, we exhibit accurate coarse-grain molecular models using a number of balls 20 times smaller than the number of atoms, a key requirement to simulate crowded cellular environments.

6.3.2. Towards Morse Theory for Point Cloud Data

Participants: Frédéric Cazals, Christine Roth.

In collaboration with C. Robert (IBPC / CNRS, Paris, France), and C. Mueller (ETH, Zurich).

Morse theory provides a powerful framework to study the topology of a manifold from a function defined on it, but discrete constructions have remained elusive due to the difficulty of translating smooth concepts to the discrete setting.

Consider the problem of approximating the Morse-Smale (MS) complex of a Morse function from a point cloud and an associated nearest neighbor graph (NNG). While following the constructive proof of the Morse homology theorem, we present novel concepts for critical points of any index, and the associated Morse-Smale diagram [19].

Our framework has three key advantages. First, it requires elementary data structures and operations, and is thus suitable for high-dimensional data processing. Second, it is gradient free, which makes it suitable to investigate functions whose gradient is unknown or expensive to compute. Third, in case of under-sampling and even if the exact (unknown) MS diagram is not found, the output conveys information in terms of ambiguous flow, and the Morse theoretical version of topological persistence, which consists in canceling critical points by flow reversal, applies.
On the experimental side, we present a comprehensive analysis of a large panel of bi-variate and tri-variate Morse functions whose Morse-Smale diagrams are known perfectly, and show that these diagrams are recovered perfectly.

In a broader perspective, we see our framework as a first step to study complex dynamical systems from mere samplings consisting of point clouds.

6.4. Misc

Computational Biology, Biomedicine.

6.4.1. Book

Participant: Frédéric Cazals.

Edited in collaboration with P. Kornprobst, from the Neuromathcomp project-team.

Biology and biomedicine currently undergo spectacular progresses due to a synergy between technological advances and inputs from physics, chemistry, mathematics, statistics and computer science. The goal of the book [15] is to evidence this synergy, by describing selected developments in the following fields: bioinformatics, biomedicine, neuroscience.

This book is unique in two respects. First, by the variety and scales of systems studied. Second, by its presentation, as each chapter presents the biological or medical context, follows up with mathematical or algorithmic developments triggered by a specific problem, and concludes with one or two success stories, namely new insights gained thanks to these methodological developments. It also highlights some unsolved and outstanding theoretical questions, with potentially high impact on these disciplines.

Two communities will be particularly interested. The first one is the vast community of applied mathematicians and computer scientists, whose interests should be captured by the added value generated by the application of advanced concepts and algorithms to challenging biological or medical problems. The second is the equally vast community of biologists. Whether scientists or engineers, they will find in this book a clear and self-contained account of concepts and techniques from mathematics and computer science, together with success stories on their favorite systems. The variety of systems described will act as an eye opener on a panoply of complementary conceptual tools. Practically, the resources listed at the end of each chapter (databases, software) will prove invaluable to get started on a specific topic.

AMIB Project-Team

5. New Results

5.1. RNA

`../../../../projets/amib/IMG/designconstrained.png`

Figure 3. Language-theoretical constructs for the constrained design of RNAs. Starting from a secondary structure, the language of sequences compatible with base-pairing constraints is modeled as a context-free grammar (Left), while forced and forbidden motifs (here, $\{\mathsf{AA}\}$ *is forbidden, and* $\{\mathsf{AGC}, \mathsf{GG}\}$ *are forced) can be modeled by a dedicated automaton (Right).*

5.1.1. RNA design through random generation

Extensive experiments revealed a drift of existing software towards sequences with a high `G+C`-content. Relying on our random generation methods, we showed how to control this distributional bias in sequences using a multidimensional Boltzmann sampling [30], [22]. We also explored the combination of random generation (global sampling) and local search into a novel category of *glocal* approaches, yielding promising results.

Finally, we explored language-theoretic constructs, namely products of finite-state automata and context-free languages, to force or forbid the presence of identified functional motifs within designed sequences [33].

5.1.2. Towards 3D modeling of large molecules

Ab initio research benefited from our works on research and classification of RNA structural motifs [63]. Significant progress towards the *ab initio* prediction of the 3D structure of large RNAs were achieved. This problem is beyond the scope of current approaches and we proposed a promising coarse-grained approach based on game theory [13] that scales up to several hundreds of bases.

5.1.3. Fast-fourier transform for riboswitch

In the field of RNA computational biology, many algorithms use dynamic programming to partition the folding landscape according to a set of structural parameters. More precisely, the goal is to compute the number (resp. cumulated Boltzmann weight) $c_{p_1,p_2,p_3\ldots}$ of secondary structures having p_i occurrences of some structural parameter P_i, where P_i may denote the distance to a reference structure, the number of # helices, base-pairs...The resulting algorithms, although polynomial in theory, are usually unusable in practice, particularly due to their unreasonable complexities (typically $\Theta(n^{3+2k})/\Theta(n^{2+k})$ time/memory for k parameters) and the intrinsic difficulties one encounters while trying to distribute their computation over multiple processors (highly connected dependency graph).

In collaboration with P. Clote's group (Boston College), we have described generic algorithmic principles to dramatically decrease these complexities, and make this class of algorithms practical. The main idea is to capture the partitioned space within a large polynomial, which can typically be efficiently evaluated (typically in $\Theta(n^3)$) as soon as the parameters are additive. One can then perform (possibly in parallel) $\Theta(n^k)$ independent evaluations of the polynomial, and use the Discrete Fourier Transform to recover the coefficients in $\Theta(k \cdot n^k \cdot \log(n))$ time. Applying these principles to the `RNAbor` algorithm, whose complexities were in $\Theta(n^5)/\Theta(n^3)$, we obtained an novel $\Theta(n^4)/\Theta(n^2)$ (parallelizable in $\Theta(n^3)/\Theta(n^2)$ time/memory on $m \to \infty$ processors), we obtained a novel algorithm to detect bistable thermodynamic structures, such as riboswitches, which we presented at `Recomb`'13 [32].

5.2. Sequences

5.2.1. Random generation

The random generation of decomposable combinatorial structures, pioneered by P. Flajolet in the 80s, provides an elegant, yet powerful, framework to model and sample the objects which appear in computational biology. Random samples can then be used to assert the significance of a given observable when closed form formulae are difficult to obtain.

Messenger RNAs (mRNAs) encode proteins, but may also independently feature structured motifs which are crucial to recoding and alternative splicing mechanisms. In order to predict such motifs, the stability of smaller regions within a given mRNA must be compared to that of sequences generated with respect to a **background model** which, at the same time, preserves the encoded amino-acid sequence and the capacity of the overall sequence to form a stable fold (proxy-ed by the dinucleotide composition). Using multidimensional Boltzmann sampling, we have revisited the underlying – well-defined, yet never solved exactly – random generation problem, and provided the first unbiased and practical algorithm for the problem [27]. The algorithm, developed in collaboration with McGill and Université de Montréal (Canada), has linear time complexity as soon as a small tolerance (typically $\Theta(1/\sqrt{n})$) on the composition is allowed.

Some other biological objects, such as RNA secondary structures, naturally appear with probabilities which are poorly modeled by the uniform distribution. To better model such objects, Denise *et al* [3] have introduced the **weighted distribution**, and adapted classic random generation algorithms such that each object within a given combinatorial family can be generated with respect to it. However, the exponentially increasing probability ratio between the most and least probable object sometimes leads to a large degree of redundancy within generated sets . To work around this issue, and generate non-redundant sets of objects, we have proposed a sequential algorithm with deterministically avoids any previously generated word, without introducing any bias in the generation [17].

../../../../projets/amib/IMG/sparcs.png

Figure 4. Workflow of our `NASP` *pipeline [27]: An assessment of significantly (un)-structured regions in protein-coding RNAs can be achieved through a dinucleotide-preserving random generation of sequences encoding the same protein.*

`../../../../projets/amib/IMG/nonredgen.png`

Figure 5. A uniform random generation of words avoiding a predefined set of words can be achieved using a dedicated data structure, leading to a careful correction of the emission probabilities. Enriching the set of forbidden words after each generation, one obtains a non-redundant generation algorithm [17].

Besides, in collaboration with the Fortesse group at LRI, we developed a new divide and conquer algorithm for the random generation of words of regular languages, and we performed a complete benchmarking of all state-of-the-art methods dedicates to this problem [56].

`../../../../projets/amib/IMG/RNAPyro.png`

Figure 6. While simultaneously sequencing the genome of a (microbial) community, Next-Generation Sequencing techniques produce small genomic fragments, whose diversity arises from a combination of genetic variants and sequencing errors. We used knowledge of the RNA secondary structure to develop a pre-filter that detects and corrects post-mapping sequencing errors.

5.2.2. Next Generation Sequencing (NGS)

As a side-product of our previous collaborative studies with J. Waldispühl (McGill, Canada), focusing on sequence/structure relationship in RNA, we revisited the problem of detecting and correcting RNA sequences obtained using pyrrosequencing techniques. Indeed, ribosomal RNAs are often used to estimate the population diversity within a microbiome, and sequencing errors may lead to biased estimates. In this context, we investigated whether a complete knowledge of the RNA secondary structure could be exploited to detect and correct errors in NGS reads.

To that end, we introduced a probabilistic model, defined over all sequences at maximal distance d from the input read and their respective folding. This model captures both the stability of the induced fold and its compatibility with a reference multiple sequence alignment. We designed a linear-time inside/outside algorithm to compute exactly the probability that a given position is mutated in the ensemble. Our initial implementation, presented at RECOMB'13 [29] and published an extended version in *Journal of Computational Biology* [23], revealed encouraging results, and we plan to combine it with a population diversity estimator to test its potential in a metagenomics context.

5.2.3. Combinatorics of motifs

An algorithm for pvalue computation has been proposed in [44] that takes into account a Hiddden Markov Model and an implementation, SUFPREF, has been realized (http://server2.lpm.org.ru/bio).

Combinatorics of clumps have been extensively studied, leading to the definition of the so-called *canonic clumps*. It is shown in [28] that they contain the necessary information needed to calculate, approximate, and study probabilities of occurrences and asymptotics. This motivates the development of a *clump automaton*. It allows for a derivation of pvalues, decreasing the space and time complexity of the generating function approach or previous weighted automata.

Large deviations approximations are needed for very rare events, e.g. very small pvalues, as Gaussian approximations are known not to be applicable. In [21], combinatorial properties of words allow to provide an explicit and tractable formula for the tail distribution with a low space and time complexity and a guaranteed tightness. Double strands counting problem is addressed where dependencies between a sequence and its complement plays a fundamental role. A large deviation result is also provided for a set of small sequences, with non-identical distributions. Possible applications are the search of cis-acting elements in regulatory sequences that may be known, for example from ChIP-chip or ChipSeq experiments, as being under a similar regulatory control. In a recent internship at LIX, F. Pirot detected a Chi-like motif in Archae genome.

In a collaboration with AlFarabi University, where M. Régnier acts as a foreign co-advisor), word statistics were used to identify mRNA targets for miRNAs involved in various cancers [8], [9].

5.3. 3D Modelling and Interactions

Transmembrane beta-barrel proteins (TMB) account for 20 to 30% of identified proteins in a genome but, due to difficulties with standard experimental techniques, they are only 2% of the RCSB Protein Data Bank. Therefore, we study and design algorithmic solutions addressing the secondary structure, an abstraction of the 3D conformation of a molecule, that only retains the contacts between its residues. Although this representation may disregard some of the fine details of the molecule conformation, it still retains the general architecture of molecules, and is especially useful in the study of RiboNucleic Acids (RNAs) and transmembrane beta-barrel proteins (TMB). The latter class of proteins accounts for 20 to 30% of identified proteins in a genome but, due to difficulties with standard experimental techniques, they constitute only 2As TMB perform many vital functions, the prediction of their structure is a challenge for life sciences, while the small number of known structures prohibits knowledge-based methods for structure prediction. As TMBs are strongly structured objects, model based methodologies [26], [25] are an interesting alternative to these conventional methods. The efficiently obtained 3D structures provide a good model for further 3D and interaction analyses.

In a recent work [34], we focused on the identification of protein-protein complexes based on the putative interaction between pairs of proteins as the sole source of information. From the results obtained on *E. coli*, we started working on the prediction of multi-body protein complexes from sequence information alone.

In our protein-RNA project, we managed to obtain the first learning results. We optimized the RosettaDock scores and showed that such an optimization cannot be done efficiently without expert knowledge. The first results are to be presented at EGC in 2014 [61].

5.3.1. *Large scale cross-docking study of the specificity of protein-protein interactions*

The year 2013 saw the conclusion of a long-term collaboration, involving A. Carbone (UPMC) and A. Lopes (IGM, Paris XI). In a recent paper published in the prestigious *Plos Computational Biology* [16] journal, we showed that combining coarse-grain molecular cross-docking simulations and binding site predictions based on evolutionary sequence analysis is a viable route to identify true interacting partners for hundreds of proteins with a variate set of protein structures and interfaces. Also, we realized a large-scale analysis of protein binding promiscuity and provided a numerical characterization of partner competition and level of interaction strength for about 28000 false-partner interactions. Finally, we demonstrated that binding site prediction is useful to discriminate native partners, but also to scale up the approach to thousands of protein interactions. This study was based on a large computational effort made by thousands of internet users helping the World Community Grid over a period of 7 months.

5.4. Data Integration

Work performed in the Data Integration axis this year has been dedicated to the design and implementation of a new approach to reduce the complexity of scientific workflow structures. More precisely, we focused on the presence of “anti-patterns” in the workflow structures, idiomatic structures that lead to over-complicated design. We have then proposed the *DistilFlow* method and a tool for automatically detecting such anti-patterns and replacing them with different patterns which result in a reduction in the workflow’s overall structural complexity [10] (BMC Journal paper accepted, published early 2014). This work has been performed in close collaboration with the Taverna group from the University of Manchester.

DistilFlow is part of J. Chen’s thesis who has defended his PhD on October 11th, 2013 [7] and is now back to China as a research assistant in Lanzhou University.

5.5. Systems Biology

Systems Biology includes the study of interaction networks such as gene regulatory, metabolic, or signaling networks. It involves both designing the topology of the networks and predicting their dynamic and spatiotemporal aspects. It requires the import of concepts from across various disciplines and crosstalk between theory, benchwork, modelling and simulation.

5.5.1. *Topological analysis of metabolic networks*

In [73] we have developed a biclustering algorithm for elementary flux modes that is based on the Agglomeration of Common Motifs (ACoM). This allows a drastic diminution of the number of less significant fluxes and a kind of factorization of most important fluxes, yielding an algorithm running in quadratic time in the number of elementary flux modes.

We applied this algorithm to describe the decomposition into elementary flux modes of the central carbon metabolism in *Bacillus subtilis* and of the yeast mitochondrial energy metabolism. For *Bacillus subtilis*, a specific inhibition on the second domain of the lipoamide dehydrogenase (pdhD) component of pyruvate dehydrogenase complex that leads to the loss of all fluxes was exhibited [20]. Such a conclusion is not predictable in the classical approach.

5.5.2. *Evolution of metabolic networks*

A collaboration with IGM on the evolution of metabolic networks is ungoing. We aim at understanding how such networks would emerge over time among the variety of species, and how these changes could be responsible for characteristic life traits. Our methodology to characterize the evolutionary origin of the enzymatic repertoire of different fungal groups relies on machine learning. Preliminary results were presented at JOBIM 2013 [35].

5.5.3. Signaling networks

Our goal is to help the understanding of signaling pathways involving (GPCR) and to provide means to semi-automatically construct the signaling networks. Our method takes into account various kinds of biological experiments and their origin and automatically builds and draws the inferred network. Comparing the automatically deduced network with an already known fragment of the FSHR network allowed us to obtain new interesting hypotheses that are currently experimentally tested by biologists, our collaborators from INRA-BIOSin Tours. In the next months, experimental data for some GPCR (FSH, 5HT2 et 5HT4) will be prepared by BIOS and IGF (Montpellier), in the context of a GPCRNET ANR project.

Besides, in collaboration with K. Inoue, through the NII International Internship Program, we have studied the System Biology Graphical Notation language, a standard for expressing molecular networks, especially signaling networks, and proposed a translation of SBGN-AF into a logical formalism [31].

5.5.3.1. Modelling with Hsim

In a collaboration of P. Amar with microbiologists, the group of Marie-Joëlle Virolle from the *Institut de Génétique et de Microbiologie*, a first explicative model was proposed for the sigmoidicity of the shape of the survival curve of bacteria (*S. lividans*) having a antibiotic resistance gene, expressed at different levels, in presence of a constant concentration of antibiotics [24], [6], [18], [41].

This is particularly important since this method of inclusion of an antibiotics resistance gene to report the activity of its promoter is widely used in the streptomyces community.

5.5.3.2. Cancer and metabolism

It is shown in M. Behzadi's PhD thesis that most systems have very stable behaviours and that even large variations of their chemical characteristics do not affect the nature of the equilibria. This very general situation has been discovered by simulation but in some cases it is even possible to prove it mathematically.

Our collaborators M. Israël and L. Schwartz have listed more than a hundred tentative such bifurcations that we intend to study systematically. A preliminary study of the mitotic cycle with L. Paulevé has also put in evidence the strong influence of the pH of the cell on its capacity to duplicate. The PhD thesis of E. Bigan, co-directed by S. Daoudi (Univ. Denis Diderot) and J.-M. Steyaert investigates the generic properties of such complex systems and confirms that the ones we have already studied are not exceptions [43]. Some prospective cases are studied in [14].

BAMBOO Project-Team

6. New Results

6.1. Symbiont genome evolution and dynamics

The objective of this part of our work was to analyse genome rearrangements and dynamics. The results obtained were both algorithmic and biological.

In terms of algorithms, we developed a new method for repeat identification (RIME) [12], as well as an algorithm for finding the minimum number of three constrained versions of inversions that transform one given genome into another [25]. The constrained versions concerned symmetric, almost-symmetric and unitary inversions. The genome rearrangement algorithm is not exact: it is based on a greedy randomized search procedure to find such minimum number of constrained inversions.

The main set of biological results [4], [14] concerned trypanosomatids of the genera *Angomonas* and *Strigomonas* that live in a mutualistic association characterised by extensive metabolic cooperation with obligate endosymbiotic Betaproteobacteria. In contrast to their counterparts lacking symbionts, such trypanosomatids exhibit lower nutritional requirements and are autotrophic for essential amino acids and vitamins. Phylogenetic analyses showed that the cooperation in the first case is complemented by multiple horizontal gene transfers, from bacterial lineages to trypanosomatids, that appear to have occurred several times in the course of evolution. In contrast, but for three exceptions, such transfers are absent as concerns vitamin biosynthesis.

The above work was made possible in part because of the sequencing and annotation of the genomes whose metabolic pathways could then be inferred. We participated in these for some of the genomes involved in the above study [17].

6.2. Host-symbiont metabolic dialog

The methodological work done has covered one main question concerning what we called metabolic stories. Given a subset of metabolites representing those monitored as being under- or over-produced in some condition (e.g., interaction with a parasite) and a metabolic network represented as a compound graph, metabolic stories are maximal directed acyclic graphs (DAGs) that cover all the metabolites in the subset of interest, and have all sources and targets among these metabolites. One exact algorithm (TOUCHÉ, [24]) was developed to enumerate all metabolic stories that improved on our previous method (GOBBOLINO).

The algorithm above was validated on biological data [16] in a study of the response of yeast to cadmium exposure. We used this system as a proof of concept for our method and we showed that we are able to find a story that reproduces very well the current knowledge about the yeast response to cadmium. We further showed that this response is mostly based on enzyme activation. We also provided a framework for exploring the alternative pathways or side effects this local response is expected to have in the rest of the network. Finally, we discussed several interpretations for the changes we see and we suggest hypotheses which could in principle be experimentally tested.

6.3. Host-symbiont genetic dialog

Two sets of problems were addressed: (i) the development of algorithms for analysing NGS data especially RNAseq, and (ii) the development of algorithms for identifying small RNAs, notably microRNAs, and their targets.

The computational work on NGS is described in another section.

Computational work on small RNAs, initially miRNAs, led to the development of a new algorithmic method. This builds upon previously developed approaches, one which was applied to *Anopheles darlingi* for inferring miRNAs that however had a high rate of false positives, and a second that provided a way for navigating among all the candidates found. Recently however, we arrived at a better model for such inference in the double sense that the rate of false positives is smaller without losing in sensitivity, while the method is much faster. The paper presenting this work and the algorithm (MIRINHO) was submitted and is currently in revision.

6.4. Symbiont-host co-cladogenesis and co-evolution at the sequence and network levels

The problem here was to: (i) study the co-evolution of a set of hosts and their symbionts, and (ii) to understand the genetic architecture of a parasitic invasion by investigating the different phenotypes such invasion produces in the host.

Work on the first point took longer than initially planned but two papers are now submitted. In the first, titled “Co-phylogeny Reconstruction via an Approximate Bayesian Computation", we describe an algorithm (COALA) for estimating the frequency of co-evolutionary events based on a likelihood-free approach. The benefits of this method are twofold: (1) it provides more confidence in the set of costs to be used in a reconciliation, and (2) it allows to estimate the frequency of the events in cases where the dataset consists of trees with a large number of taxa. We evaluate our method on simulated and on real datasets. We show that in both cases, for a same pair of host and parasite trees, different sets of frequencies for the events constitute equally probable solutions. Moreover, sometimes these sets lead to different parsimonious optimal reconciliations, in the sense of presenting a different number of the events. For this reason, it appears crucial to take this into account before attempting any further biological interpretation of such reconciliations. More generally, we also show that the set of frequencies can vary widely depending on the input host and parasite trees. Indiscriminately applying a standard vector of costs may thus not be a good strategy.

In the second submitted paper related to the study of co-evolution and titled “EUCALYPT: Efficient tree reconciliation enumerator", we present a polynomial-delay algorithm for enumerating all optimal reconciliations. We show that in general many optimal solutions exist. We give an example where, for two pairs of host-parasite trees having each less than 40 leaves, the number of solutions is 2309, even when only time feasible solutions are kept. To facilitate their interpretation, those solutions are also classified in terms of the number of each event that they contain. This often enables to reduce considerably the number of different classes of solutions to examine further, but the number may remain high enough (16 for the same example). Depending on the cost vector, both numbers may increase considerably (for the same instance, to respectively 4080384 and 275).

Concerning the second question (genetic architecture of a parasitic invasion), one such phenotype is called “cytoplasmic incompatibility” (CI). Briefly, when a parasite invades a male host, it induces the death of the host’s offspring unless the female is also infected. This has been explained by a toxin/antitoxin model that involves a toxin deposited by the parasites in the male’s sperm inducing the death of the zygote unless neutralised by an antidote produced by the parasites in the egg. One toxin/antitoxin pair is usually linked to one genetic factor. Given a set of observed CIs, the question is how many genetic factors explain it. In its simplest form, this mathematically translates into, given a bipartite graph, finding its minimum biclique edge cover. One biclique corresponds to one factor. We had previously analysed the complexity of the problem and proposed an algorithm that was this year applied to a set of CI data from *Culex pipiens* [18].

6.5. NGS for biodiversity

In collaboration with the Laboratoire d’Écologie Alpine (LECA) at Grenoble where there is a strong expertise on DNA meta-barcoding, we had devised several tools for barcode design and analysis. ECOPRIMERS thus identified new barcode markers and their associated PCR primers within a DNA meta-barcoding approach. The algorithm was optimised two quality indexes measuring taxonomical range and discrimination to select the most efficient markers from a set of reference sequences, according to some experimental constraints such as marker length or specifically targeted taxa. We had also devised assembler algorithms directed to organelles (mitochondria or chloroplasts). This year, in collaboration with the Inria project-team MISTIS, we developed a statistical modelling approach to investigate the spatial cross-correlations between different taxa identified by meta-barcoding of soil sample from French Guiana (this was selected as a conference paper at the “45ème Journées de Statistiques” 2013 that took place at Toulouse and is organised by the Société Française de Statistique. This approach allows to visualise the co-occurrence pattern as a “species interaction graph”, and to study the mutual exclusion (competition) or inclusion (symbiosis) of different plant species.

6.6. NGS for genotypic variation detection

The computational work on NGS data concerned both algorithmic design and complexity analysis.

Based on the idea that each genotypic variation will correspond to a recognisable pattern in a de Bruijn graph constructed from a set of sequence reads, we had proposed a generic model for SNPs in DNA data, and then generalised it to the analysis of RNA. In this case, not only SNPs are present but also alternative splicing (AS) events, which, once again, generate a recognisable pattern in the de Bruijn Graph. We had therefore proposed a general model for all these variations (SNPs, indels and AS events) and introduced an exact algorithm (KISSPLICE) to extract all alternative splicing events. The algorithm also outputs candidate SNPs and indels. This year, we improved the algorithm [26]. As the problem relates to an old one in algorithmics (cycle enumeration), we also revisited it from a theoretical point of view [23].

The improved version of KISSPLICE [26] was used to analyse RNAseq data from two lines of *Asobara tabida* exhibiting different ovarian phenotypes in the absence of its endosymbiont *Wolbachia*. Although infected individuals of the two lines have similar phenotypes, numerous genes are differentially expressed between the two infected conditions. This could mean that two divergent strategies of tolerance have evolved. Preliminary results on the analysis of polymorphisms between these two lines suggest that differentially expressed genes tend to accumulate more variation. We are currently, via experiments done by the biologists in our team, testing the hypothesis that such genes are under strong selection pressure and may evolve through mutation accumulation, a process that could be related to assimilation.

A preliminary analysis of human data from the ENCODE project performed with KISSPLICE showed that an assembly-based method (without reference genome) is able to recover AS events that are missed by mapping-based methods (with a reference genome). Some of these events were experimentally validated, which represents the best type of proof we can provide to the biologists. The experimental part is made by our collaborator from the Inserm, Didier Auboeuf, in his team at the Centre National de Cancérologie of Lyon (CNCL), with whom we had an Inserm project, EXOMIC, funded for three years starting from 2012.

The identification of SNPs is also getting renewed interest even in the presence of a reference genome thanks to the possibility of re-sequencing many times the genome of a same or of very closely related species. The difficulty in the case of SNPs is to distinguish them from sequencing errors and from inexact repeats. We proposed a statistical test enabling to identify variations that are condition-specific, which enables to greatly enrich the list of potential SNP candidates. The paper on this test is in preparation. Its results as applied to the RNAseq data from two lines of *Asobara tabida* (see above) and to Drosophila species having diverged very recently were validated by, respectively, Fabrice Vavre and Cristina Vieira, both members of BAMBOO.

We also started addressing the problem that repeats (such as transposable elements for instance but not only) represent more in general for both local and global assemblers. We are thus developing a method that would enable to identify, in a de Bruijn graph built from RNAseq data, the vertices potentially corresponding to the borders of a repeated sequence. Preliminary results on simulated and real data show that the approach is promising (paper in preparation).

BEAGLE Project-Team

6. New Results

6.1. Stochastic dynamics of gene expression

A number of studies have established that stochasticity in gene expression may play an important role in many biological phenomena but the molecular mechanisms at stake are still poorly understood. By joint experimental and computational approaches, we explored the role played by chromatin dynamics in the regulation of stochastic gene expression in higher eukaryotic cells [31]. For this purpose, our biological partner generated isogenic chicken-cell populations expressing a fluorescent reporter integrated in one copy per clone. Although the clones differed only in the genetic locus at which the reporter was inserted, they showed markedly different fluorescence distributions, revealing different levels of stochastic gene expression. Use of chromatin-modifying agents then showed that direct manipulation of chromatin dynamics had a marked effect on the extent of stochastic gene expression. We then fitted the experimental data to a two-state model describing the opening/closing process of the chromatin. The model showed that the differences between clones seemed to be due mainly to the duration of the closed state, and that the agents we used mainly seem to act on the opening probability. These results highlight the importance of chromatin dynamics in stochastic gene expression. They shed a new light on the mechanisms of gene expression in higher eukaryotic cells, and argues in favor of relatively slow dynamics with long (hours to days) periods of quiet state.

This work was part of Gaël Kaneko's PhD and results from our long-lasting collaboration with Olivier Grandrillon and his BM2A team in the CGphyMC (Centre de Génétique et de Physiologie Moléculaire et Cellulaire, Lyon).

6.2. The impact of anomalous diffusion on cell signaling

This year, we published two papers describing the impact of diffusion and clustering in membrane domains for ubiquitous biological pathways. In the first paper [18], we showed that clustering of receptors in membrane domains affect severely the activation of 'hit and run' type of pathways (eg. IRS1, G proteins). This is a pure diffusion result and it is obtained without modifying molecular affinity. This impairment is dramatically important when receptors are highly clustered such are the cases for insulin and adrenergic receptors. In the same direction, we studied the impact of modified diffusion on the other ubiquitous pathway: enzyme/substrate equilibrium [10]. In that case diffusion was modified either using subdiffusion - obstacles and Continuous Time Random Walk - or using space-based inhomogeneous diffusion. We showed that while impairing diffusion all three mechanisms behave differently in the stationary regime. Therefore it it not possible to assume simple space-dependent diffusion for subdiffusion at the equilibrium limit (as it is always assumed). Furthermore, we showed that in the case of space-dependent diffusion - the shape of the diffusion profile can drastically affect the equilibrium and modify the pathway. Since these three phenomenons are thought to occur either in the membrane or the cytosplasm, our results show they can have non trivial effect of all chemical reaction occurring within these medium.

This research will be carried on in the group in 2014 and expanded by the more mathematical approaches initiated in 2013 by collaborations with V Calvez (Numed Inria Lyon), T Lepoutre (Dracula, INIA Lyon) and S. Fedotov (Univ Manchester, UK).

6.3. Localization of protein aggregates in E. coli

Aggregates of misfolded proteins are a hallmark of many age-related diseases. Recently, they have been linked to aging of *Escherichia coli* (*E. coli*) where protein aggregates accumulate at the old pole region of the aging bacterium. Because of the potential of *E. coli* as a model organism, elucidating aging and protein aggregation in this bacterium may pave the way to significant advances in our global understanding of aging. A first obstacle along this path is to decipher the mechanisms by which protein aggregates are targeted to specific intercellular locations. Here, using an integrated approach based on individual-based modeling, time-lapse fluorescence microscopy and automated image analysis, we show that the movement of aging-related protein aggregates in *E. coli* is purely diffusive (Brownian). Using single-particle tracking of protein aggregates in live *E. coli* cells, we estimated the average size and diffusion constant of the aggregates. Our results provide evidence that the aggregates passively diffuse within the cell, with diffusion constants that depend on their size in agreement with the Stokes-Einstein law. However, the aggregate displacements along the cell long axis are confined to a region that roughly corresponds to the nucleoid-free space in the cell pole, thus confirming the importance of increased macromolecular crowding in the nucleoids. We thus used 3D individual-based modeling to show that these three ingredients (diffusion, aggregation and diffusion hindrance in the nucleoids) are sufficient and necessary to reproduce the available experimental data on aggregate localization in the cells. Taken together, our results strongly support the hypothesis that the localization of aging-related protein aggregates in the poles of *E. coli* results from the coupling of passive diffusion-aggregation with spatially non-homogeneous macromolecular crowding. They further support the importance of “soft” intracellular structuring (based on macromolecular crowding) in diffusion-based protein localization in *E. coli*.

This work is a collaboration with the microbiology group led by A. Lindner (INSERM U1001, Cochin Med School, Paris). It has been published in [3] as part of A.S. Coquel’s PhD (defended Nov 2012, co-supervision H. Berry-A. Lindner).

6.4. The molecular signaling basis of neuronal plasticity

Many of the cell-level properties of the neurons vary as a function of the signals from other neurons or past activity. These modifications are often maintained in the long term, giving rise to cell memory. We have developed models of how the implicated signaling networks self-organize to support a memory and how this leads to cell-level responses such as changes of the firing threshold [24] or the spike-timing dependence [34]. The latter, for instance, corresponds to the observation that the probability of transfer of an electrical signal (spike) between two connected neurons (the synaptic weight) adapts depending on the timing between previous consecutive presynaptic and postsynaptic spikes. Combining a model of the implicated signaling networks with experimental measurements, we have uncovered the molecular mechanisms supporting this memory.

This work is developed in collaboration with both with applied mathematicians (B. Cessac, Inria Neuromath-comp, Sophia-Antipolis) and experimental neurobiologists (L. Venance, Collège de France, Paris).

6.5. A model for adipocyte size based on size-dependent lipid fluxes

We proposed in a paper published this year [28] a novel model that explains some of the peculiarities in the fat tissue storage cells. Indeed, adipocytes, as they are called, come in various size – with up to one order of magnitude in amplitude – but do not possess any characteristic size. The cellularity, the cell size distribution, is bimodal. We showed that a simple model of size-dependent lipid fluxes (using data from Carmen Lab) can explain this bimodality and allow us to retrieve any target cell distribution. Our result also provides an elegant and testable hypothesis for the triggering of adipocytes proliferation. The amount of unstored free fatty acid is actually a marker that the population has reached its maximal volume. This amount could serve as an index to start the proliferation.

This was a joint work with experimentalists from the CARMEN Institute (INSERM UMR1060, Lyon), namely C. Soulage and A. Géloën, and was part of H. Julienne master’s thesis.

6.6. Evolution of antibiotic resistance

The emergence of antibiotic resistant bacteria is a major threat to public health and there is a constant need for education to limit dangerous practices. Here, we propose to use alife software to develop training media for the public and the physicians. On the basis of the Aevol model we have been developing for more than six years, we built a game in which players fight bacterial infections using antibiotics. In this game the bacteria can evolve resistance traits, making the infection more and more difficult to cure. The game has been tested with automatic treatment procedures, showing that it behaves correctly. It was demonstrated during the French "Nuit des Chercheurs" in October 2012 and was published in 2013 in the ECAL conference [2].

This is a joint work with Dominique Schneider from the Laboratoire Adaptation et Pathogénie des Microorganismes (LAPM, UMR CNRS 5163, Grenoble).

6.7. Spontaneous dynamics of genome size

Even though numerous genome sequences are now available, evolutionary mechanisms that determine genome size, notably their fraction of non-coding DNA, are still debated. In particular, although several mechanisms responsible for genome growth (proliferation of transposable elements, gene duplication and divergence, etc.) were clearly identified, mechanisms limiting the overall genome size remain unclear. By using a matrix population model, we showed that genome size can be simply limited by the spontaneous dynamics of duplications and large deletions, which tends to make genomes shrink even if the two types of rearrangements occur at the same rate. In the absence of Darwinian selection, we proved the existence of a stationary distribution of genome size even if duplications are twice as frequent as large deletions. To test whether selection can overcome this spontaneous dynamics, we also simulated our model numerically and chose a fitness function that directly favors genomes containing more genes, while keeping duplications twice as frequent as large deletions. In this scenario where, at first sight, everything seems to favor infinite genome growth, we showed that genome size remains nonetheless bounded. As a result, our study reveals a new pressure that could help limiting genome growth.

This work was part of Stephan Fischer's PhD thesis, which was defended in December 2013. A manuscript is currently under review. Stephan's PhD was co-supervised by Samuel Bernard (Inria Dracula team and Institut Camille Jordan, UMR CNRS 5208, Lyon).

6.8. Inference of evolutionary molecular events at different scales

We have progressed in the integration of several evolutionary events at different scales of genomes in a single model used for inference of ancient events from the observation of extant genomes. We handle nucleotide substitutions, gene duplications, losses, lateral transfers and rearrangements. We have tested the framework on 36 cyanobacteria species, reconstructing up to 80% of ancestral chromosomes in some clades [8]. The inference algorithm is still mainly sequencial, in the sense that it first accounts for nucleotide substitutions, gene duplications, losses, lateral transfers [29], and then for rearrangements. But we also developed a way to provide a feedback of the result on rearrangements to the inference of substitutions by correcting gene trees [22], [38]. We have used these methods to reconstruct a nucleotide-scale sequence of the genome of the medieval black death agent [9],[44]. It includes a chromosome and three plasmids, and is different in structure from any extant strain. We follow the first ancient bacterial genome sequencing in 2011 and complete and order the genome with computational predictions. We then dispose of a complete view of the molecular evolution in the *Yersinia pestis* clade.

This work was part of Murray Patterson's post-doctoral fellowship. It also involved collaborations with L. Gueguen and V. Daubin from the Laboratoire de Biométrie et Biologie Evolutive (UMR CNRS 5558, Lyon), with Nadia El-Mabrouk from the Département d'Informatique et de Recherche Opérationnelle in Montréal (Canada), with Cédric Chauve from the Department of Mathematics of Simon Fraser University (Burnaby, Canada), and with G. Szollosi from the Biophysics Research Group in Budapest (Hungary).

BIGS Project-Team

6. New Results

6.1. Modern methods of data analysis

Participants: R. Bar, S. Ferrigno, B. Lalloué, J-M. Monnez, A. Muller-Gueudin, S. Tindel

6.1.1. Help to medical decision and telemedecine in the monitoring of heart failure

We describe here a project started in 2013, for which we expect some concrete output in 2014. This project fits in the general framework of telemedecine and more precisely in the monitoring of heart failure patients. From measurements performed automatically and daily on a patient at home through a new process under development at the Pluri-Thematic Clinical Investigation Center of the University Hospital of Nancy, the aim is to propose therapeutic adjustments to improve the prognosis of patient in order to increase his chances of survival or to avoid his rehospitalization.

The patient's condition and its evolution are determined by the initial values of his biological or clinical parameters as well as those collected throughout his follow-up. The treatments are intended to stabilize or change the values of parameters in order to avoid the occurrence of adverse events, in particular the death of the patient. This is why the first part of the study will consist in building survival scores or rehospitalization scores according to the values of biological or clinical parameters.

In a second part, we will seek to build models of the evolution of the values of biological or clinical parameters depending on treatments (average or cumulative drug doses, drug combinations) and patients' characteristics. This will allow to predict the potential effect of an adjustment proposal or modification of treatment and then predict a new survival score to conclude the relevance or not of the proposed medication.The physician will have this help to confirm or change his decision which belongs finally to him.

We will use to carry out this study a wide range of classic and recent methods of data analysis, in particular discriminant analysis, without a priori: several methods will be used, compared and selected according to their performance in the treated applications.

6.1.2. Online factorial data analysis methods

Nowadays data analysts are often faced with the problem of dealing with a rapid and infinite flow of data. Examples include web, telecommunications, process control or financial data. We made first the assumption that the data are generated at random according to a stationary distribution, but in many cases this assumption does not hold true. We developed in [13] the online adaptation of principal component analysis and other dimension reduction statistical algorithms by using stochastic approximation. An R package was developed by Romain Bar.

6.1.3. Data analysis techniques and Bayesian models applied to the context of social inequalities and environmental exposures

The aim of [10] is to improve the knowledge about and apply data mining techniques and some Bayesian model in the field of social and environmental health inequalities. The health event considered is infant mortality. We try to explain its risk with socio-economic data retrieved from the national census and environmental exposures such as air pollution, noise, proximity to traffic, green spaces and industries. The data mining part details the development of a procedure of creation of multi-dimensional socio-economic indices and of an R package that implements it.

6.1.4. A simultaneous stepwise variable selection and clustering algorithm to discriminate a class variable with numerous levels

In supervised learning the number of levels of a categorical variable to explain can be high. When some of its levels are of low frequency, clustering them in order to reduce the number of classes can be useful to perform relevant discriminant analyses. On the other hand selecting relevant predictors is a crucial step to build robust and efficient classification rules, especially when too many variables are available in regard to the overall sample size. We are currently carrying out an extension of an algorithm we had devised to solve both these problems using an alternate minimization of Wilks' Lambda. We show through simulations the interest of adding Akaike Information Criterion as another optimality criterion. We also moved forward to stepwise selection and applied this new version of our algorithm to real allergology datasets.

6.1.5. Local polynomial regression. Application to the estimation of the fetal growth.

This topic is an ongoing collaboration with M. Maumy-Bertrand, for which we expect a publication in 2014. We have established exact rate of strong uniform consistency for the local linear estimator of the conditional distribution function. We want to extend our results to obtain exact rates of strong uniform consistency for the local linear estimator of other conditional quantities: the conditional mean $\mathbb{E}(Y|X)$, and the conditional quantiles $q_\alpha(x) = \inf\{y : F(y|x) \geq \alpha\}$, for $\alpha \in (0, 1)$.

Another crucial problem with the non parametric regression methods is the choice of the bandwidth parameter h. It is common in practice to choose $h > 0$ so to minimize asymptotically the mean square error (MSE) or the mean integrated square error (MISE). This minimization leads to an optimal choice of h of the form $h_n = C(X_1, ..., X_n)n^{-1/5}$, where n is the sample size, and $X_1, ..., X_n$ are the n independent copies of the random variable X. This bandwidth is called a *data-driven bandwidth* to enhance its dependence to the data. Our current project in this direction consists in establishing the consistency of the local linear estimator when the bandwidth h is allowed to range in a small interval which may decrease in length with the sample size. Such a result would be immediately applicable to prove uniform consistency of the local linear estimator when the bandwidth is a data-driven bandwidth $h_n = C(X_1, ..., X_n)n^{-1/5}$.

Turning to applications, note that we have a contact with Professor Bernard Foliguet at the Maternité Régionale de Nancy. We will continue to collaborate with him, to estimate growing curves of the fetal weight, and other fetal quantities thanks to the techniques mentioned above.

6.1.6. Cohort analysis

In an ongoing work with the INSERM team of P. Guéant, we aim at describing the complex interactions between genetic, phenotypic and biologic variables that are available in medical cohorts, in different contexts (cognitive decline; inflammatory intestinal diseases; liver cancer).

A firt step in our analysis, which should be finished in 2014, consists in giving an overview of the existing methods given in the literature, for the analysis of qualitative and quantitative data. Indeed, we have to describe links between qualitative and quantitative data:

1. with exploratory methods, or factorial models,
2. with regression models to predict qualitative variable by the use of qualitative or quantitative factors.

In the sequel, we will test non association or independence between the variables. The objective is to develop new methods, adapted to the studied cohorts (matching cases/controls, high number of individuals, high number of explicative variables, missing data problem). The particularity of our work is to combine statistical and symbolical methods.

After having identified and choice the relevant variables, we will have to give a model for classifying the data. The proposed models will allow us to identify subgroups of invidious, with common genetic, biologic and phenotypic characteristics.

6.1.7. Local polynomial estimation and goodness-of-fit tests

We describe here an ongoing work with Marie-José Martinez, assistant professor at the IUT of Grenoble and member of the Inria MISTIS team. A related publication should be finished at the end of 2014. Many clinical trials and other medical studies involve responses that might be considered to have a normal distribution. However, this is not invariably the case and models based on this distribution are often indiscriminately applied to data which might be better handled otherwise. This is especially true for discrete data. An approach which may yields models that are more biologically reasonable in many situations is to use generalized linear models (GLM).

In statistical theory, generalized linear models were formulated by John Nelder and Robert Wedderburn (1972) as a way of unifying various other statistical models including for examples linear regression, logistic regression and poisson regression. Such a technique was developed by McCullagh and Nelder (1989). It is an extension of the linear model, in the sense that its satisfies a relation of the form $Y = g(X) + \epsilon$ where:

- The stochastic component ϵ follows other distributions than the Gaussian.
- The function g can be non linear.

Notice that those models are well-suited to analyze dependences between variables following distributions in the so-called exponential family, like Poisson, Binomial and Gamma distributions. In practice, link functions are chosen such that the inverse link, $\mu = g^{-1}(\eta)$ is easily computed. For instance, for binomial data, logit and probit link functions are commonly used.

Our aim in this project is to use generalized linear models in order to extend our global test of goodness-of-fit to a wide range of models used in biological and medical applications. We wish to use the cumulative conditional distribution $F(y|X = x)$ again, which embodies all the information about the joint behavior of two random variables. The expected outcome is a global goodness of fit test for the relationship between two random variables in the exponential family. The test will compare a nonparametric estimator of the cumulative distribution function with the value of the cumulative distribution function under the null hypothesis.

6.1.8. Model selection for SVM

Support vector machines provide a very powerful method of data classification, for which model selection is one of the key issues. For a support vector machine, model selection consists in selecting the kernel function, the values of its parameters, and the amount of regularization. To set the value of the regularization parameter, one can minimize an appropriate objective function over the regularization path. A priori, this requires the availability of two elements: the objective function and an algorithm computing the regularization path at a reduced cost. The literature provides us with several upper bounds and estimates for the leave-one-out cross-validation error of the ℓ_2-SVM. However, no algorithm was available so far for fitting the entire regularization path of this machine. In our contribution [3], we introduce the first algorithm of this kind. It is involved in the specification of new methods to tune the corresponding penalization coefficient, whose objective function is a leave-one-out error bound or estimate. From a computational point of view, these methods appear especially appropriate when the Gram matrix is of low rank. A comparative study involving state-of-the-art alternatives provides us with an empirical confirmation of this advantage.

6.2. Estimation for complex and biological systems

Participants: T. Bastogne, C. Lacaux, S. Mézières, S. Tindel, P. Vallois

6.2.1. Tumor growth modeling

This project is an extension of our article [15], which will be written in 2014. A cancer tumor can be represented for simplicity as an aggregate of cancer cells, each cell behaving according to the same discrete model and independently of the others. Therefore to measure its size evolution, it seems natural to use tools coming from dynamics of population, for instance the logistic model. This deterministic framework is well-known but the stochastic one is worthy of interest. We are currently working on a model in which we suppose that the size V_t at time t of the tumor is a diffusion process of the type :

$$\begin{cases} dV_t = r\,V_t\,\left(1-\frac{V_t}{\kappa}\right) - c\,V_t + \beta\,V_t\,dB_t \\ V_0 = v > 0 \end{cases} \tag{1}$$

where $(B_t)_{t\geq 0}$ is a standard brownian motion starting from zero.Then (i) We define a family of time continuous Markov chains which models the evolution of the rate of malignant cells and approximate (under some conditions) the diffusion process (V_t). (ii) We study in depth the solution to equation (1). This diffusion process lives between two frontiers : 0 and κ. In this stochastic setting, the role of κ is not so clear and we contribute to understand it. We describe the asymptotic behavior of the diffusion according to the values of the parameters. The tools we resort to are boundary classification criteria and Laplace transform of the hitting time to biological worthwhile level. We believe we are able in particular to express the mean of the hitting time.

The next step in this project can be summarized as follows: at this point in our investigations on tumor growth modelization, we have identified a pertinent and consistent model. Nevertheless our study remains theoretical. A statistical estimation of the parameters $r, \kappa, c\beta$ is thus in order. This would permit to apply our model to real data. A further objective could be to consider a more complex form for the logistic term, see e.g. Schurtz (2007).

6.2.2. *Local score associated with long biological sequences*

Statistical properties of the distribution of the local score is largely used by molecular biologists to extract important features in biological sequences and in particular to determine the most significant one among a collection of biological sequences. The probabilistic model which is commonly used is the following. Associated with a sequence $(\epsilon_i)_{i\geq 1}$ of independent, centered and reduced random variables, consider $S_n = \epsilon_1 + \cdots + \epsilon_n$ and

$$\underline{S}_n = \min_{0\leq i\leq n} S_i, \quad U_n = S_n - \underline{S}_n = S_n - \min_{i\leq n} S_i \quad n \geq 0.$$

In biological sequence analysis, (ϵ_i) can for example correspond to the physical or chemical properties of the i-th amino acid or nucleotid of the sequence ; it can also reflect the similarity between components of two sequences. The local score $\overline{U}_n$ is the supremum of (U_n) up to time n. Molecular biologists are interested by this supremum as it highlights the best part of the studied sequence, the eventual segment of DNA transmitted by a common ancestor for sequence comparison or the best hydrophobic segment of a protein that would thus naturally move in a transmembrane place. It is clear that the trajectory of (U_n) can be decomposed in a succession of 0 and excursions above 0. These excursions have an important biological interpretation and in particular the highest one corresponds to the best segment due to the physico chemical property or similarity scores that have been chosen. Note that the local score $\overline{U}_n$ can be viewed as the maximum of the heights of all the excursions up to time n. In the article [22], we are interested in complete excursions up to a fixed time. This leads us to introduce the maximum U_n^* of the heights of all the excursions up to time n. The second variable which will play an important role is θ_n^* the time necessary to reach its maximal height U_n^*.
We believe that the knowledge of the joint distribution of the pair (U_n^*, θ_n^*) would permit to get more efficient statistical tests than the ones only based on the local score. This point should be developed in a forthcoming paper.

However, it seems difficult to determine explicitly the law of $(U_n^*,\ \theta_n^*)$ for a fixed n. This difficulty can be overcome considering biological sequences which have a large number of bases and approximating the initial random walk (S_n) by a Brownian motion (B_t) started at 0. Using the functional theorem of convergence of Donsker, the process (U_k) can be compared to

$$\widehat{U}(t) := B(t) - \inf_{0\leq s\leq t} B(s), \quad t \geq 0. \tag{2}$$

This leads us to consider:

1. the local score $\overline{U}(t)$ which is the maximum of the heights of all the excursions of $U(s)$ up to time t,
2. the maximum $U^*(t)$ of the heights of all the complete excursions up to time t,
3. the time $\theta^*(t)$ taken by $U(s)$ starting from the beginning of the largest excursion to hit the maximal level $U^*(t)$.

The approximation of (U_n) by $(\widehat{U}_t)$ permits to prove that the asymptotic distribution of $\left(\frac{U_n^*}{\sqrt{n}}, \frac{\theta_n^*}{n}\right)$ as $n \to \infty$ is the one of $(U^*(1), \theta^*(1))$. Consequently, our initial problem in the discrete setting reduces to determine the joint law of $(U^*(t), \theta^*(t))$, where $t > 0$ is given. We determine in [22] the distribution and the density functions of $(U^*(t), \theta^*(t))$.

6.2.3. Bacteriophage therapies

In the last years Bacteriophage therapies are attracting the attention of several scientific studies. They can be a new and powerful tool to treat bacterial infections or to prevent them applying the treatment to animals such as poultry or swine. Very roughly speaking, they consist in inoculating a (benign) virus in order to kill the bacteria known to be responsible of a certain disease. This kind of treatment is known since the beginning of the 20th century, but has been in disuse in the Western world, erased by antibiotic therapies. However, a small activity in this domain has survived in the USSR, and it is now re-emerging (at least at an experimental level). Among the reasons of this re-emersion we can find the progressive slowdown in antibiotic efficiency (antibiotic resistance). Reported recent experiments include animal diseases like hemorrhagic septicemia in cattle or atrophic rhinitis in swine, and a need for suitable mathematical models is now expressed by the community.

At a mathematical level, whenever the mobility of the different biological actors is high enough, bacteriophage systems can be modeled by a kind of predator-prey equation. Namely, set S_t (resp. Q_t) for the bacteria (resp. bacteriophages) concentration at time t. Then a model for the evolution of the couple (S, Q) is as follows:

$$\begin{cases} \mathrm{d}S_t &= [\alpha - k\,Q_t]\,S_t \mathrm{d}t + \varepsilon\,S_t\,\mathrm{d}W_t^1 \\ \mathrm{d}Q_t &= \left[d - mQ_t - k\,Q_t\,S_t + k\,b\,e^{-\mu\zeta}\,Q_{t-\zeta}\,S_{t-\zeta}\right]\mathrm{d}t + \varepsilon\,Q_t\,\mathrm{d}W_t^2, \end{cases} \tag{3}$$

where α is the reproducing rate of the bacteria and k is the adsorption rate. In equation (3), d also stands for the quantity of bacteriophages inoculated per unit of time, m is their death rate, we denote by b the number of bacteriophages which is released after replication within the bacteria cell, ζ is the delay necessary to the reproduction of bacteriophages (called latency time) and the coefficient $e^{-\mu\zeta}$ represents an attenuation in the release of bacteriophages (given by the expected number of bacteria cell's deaths during the latency time, where μ is the bacteria's death rate). A given initial condition (S_0, Q_0) is also specified, and the term $\varepsilon\,\mathrm{d}W_t$ takes into account a small external noise standing for both uncertainties on the measures and the experiment conditions. One should be aware of the fact that the latency time ζ (which can be seen as the reproduction time of the bacteriophages within the bacteria) cannot be neglected, and is generally of the same order (about 20mn) as the experiment length (about 60mn).

With this model in hand, our main results in this direction (see [1]) have been the following:

- Quantification of the exponential convergence to a bacteria-free equilibrium of equation (3) when d is large enough.
- Use of the previous result plus concentration inequalities in order to study the convergence of the noisy system to equilibrium in a reasonable time range.
- Simulation of the stochastic processes at stake in order to observe the convergence to equilibrium.

6.2.4. Light transport in tissues with probabilistic methods

Photodynamic therapy (PDT) is a type of phototherapy used for treating several diseases such as acne, bacterial infection, viruses and some cancers. The aim of this treatment is to kill pathological cells with a photosensitive drug that is absorbed by the target cells and that is then activated by light. For appropriate wavelength and power, the light beam makes the photosensitizer produce singlet oxygen at high doses and induces the apoptosis and necrosis of the malignant cells. Our project focuses on an innovative application: the interstitial PDT for the treatment of high-grade brain tumors. This strategy requires the installation of optical fibers to deliver the light directly into the tumor tissue to be treated, while nanoparticles are used to carry the photosensitizer into the cancer cells. In order to optimize the intra-cerebral position of our optical fiber, two fundamental questions have to be answered:

1. What is the optimal shape and position of the light source in order to optimize the damage on malignant cells?
2. Is there a way to identify the physical parameters of the tissue which drive the light propagation?

Notice that we are obviously not the first ones to address these issues, and there is nowadays a consensus in favor of the algorithm proposed by L. Wang and S. L. Jacques for the simulation of light transport in biological tissues. However, our starting point is the observation that the usual methods slightly lack of formalism and miss formal representations that answer the questions of identifiability. In [25], in the framework of homogeneous biological tissues, we propose an alternative MC method to Wang's algorithm. Then we also propose a variance reduction method. Interestingly enough, our formulation also allows us to design quite easily a Markov chain Monte Carlo (MCMC) method based on Metropolis-Hastings algorithm and to handle the inverse problem (of crucial importance for practitioners), consisting in estimating the optical coefficients of the tissue according to a series of measurements. We have compared the proposed MC and MCMC method and Wang's algorithm: we see that our MC method is much more consistent. However, MCMC methods induce quick mutations, which paves the way to very promising algorithms in the inhomogenous case. To handle the inverse problem, we derive a probabilistic representation of the variation of the fluence with respect to the absorption and scattering coefficients. This leads us to the implementation of a Levenberg-Marquardt type algorithm that gives an approximate solution to the inverse problem.

6.2.5. System identification of gap junctional intercellular communication channels of two cancer cell lines.

The main challenge addressed in this work [12], [14] was to assess the relevance of a proposed model-based approach to detect differences between gap junctional intercellular communication channels of two cancer cell lines. To that aim, KB and FaDu, two human head and neck carcinoma cell lines, were used. The former expresses connexin proteins (positive line) while the latter does not (negative line). Moreover, each cell line was tested on spheroid (3-D) and monolayer (2-D) slices and *in vitro* assays were repeated six times. Continuous-time system identification algorithms of the Matlab System Identification and CONTSID toolboxes are tested and applied to a set of *in vitro* data. Results firstly show an acceptable fit of the biological responses but they mainly emphasize the possibility to use several model parameters as statistics to discriminate different cancer cell lines. So, this study exemplifies the potential contribution of dynamic system identification methods and tools to the discovery of new diagnostic biomarkers in systems biology.

6.2.6. Photodynamic therapy modeling and control.

We have also carried on the development of methodological and technological innovations for the realtime control of the therapeutic efficiency in photodynamic therapy (Tylcz:2013). One part of the innovation has been protected by a patent submitted in 2012 (No.1261339, INPI) and extended in 2013. A demonstration platform is currently in development.

6.2.7. Bio-inspired system reliability method.

Based on previously developed works (Keinj, 2011, 2012), we have also proposed in [15] an extension of the target theory in biology applied to system reliability. In this development, we consider rough products produced by a factory. Each product coming from the plant has m vital elements and some elements can be damaged. To obtain a perfect product (i.e. all the constitutive m elements are safe) all the damaged elements are repaired and a test phase follows. The result of this two-steps procedure is random. We suppose that the number Z_k of non-damaged elements is a Markov chain valued in the set $\{0, 1, \cdots, m\}$, where k is the number of applied repairing-test phases. We have a qualitative result which says that if the repair phase is efficient then $P(Z_k = m)$ is close to 1. As for production of a large number n of products, the former result allows us to give conditions under which either the n elements or a fraction of these n elements are (is) safe after the application of k previous maintenance phases.

6.2.8. Dynamical Global Sensitivity Analysis as an Early Warning for System's Critical Transition.

In biology, systems with bifurcations may experience abrupt irreversible and often unwanted shifts in their performance, called critical transitions. For many systems like climate, economy, ecosystems it is highly desirable to identify indicators serving as early warnings of such regime shifts. Several statistical measures were recently proposed as early warnings of critical transitions including increased variance, autocorrelation and skewness of experimental or model-generated data. The lack of automatized tool for model-based prediction of critical transitions led to designing DyGloSA, a Matlab toolbox for dynamical global parameter sensitivity analysis (GPSA) of ordinary differential equations models. One part of our research activity in 2013 was focused on the implementation of a global sensitivity analysis method developed in (Dobre, 2011, 2012) into DyGloSA for dynamical global parameter sensitivity analysis (GPSA) of ordinary differential equations models. This work has been carried out in the context of a collaboration with the University of Luxembourg and more precisely the Thomas Sauter's team. We have shown in [2] that tools developed in this toolbox are efficient to analyze several models with bifurcations and predict the time periods when systems can still avoid going to a critical transition by manipulating certain parameter values, which is not detectable with the existing SA techniques.

6.3. Inference for gaussian systems

Participants: C. Lacaux, S. Tindel

6.3.1. Inference for dynamical systems driven by Gaussian noises.

As mentioned at the *Scientific Foundations* Section, the problem of estimating the coefficients of a general differential equation driven by a Gaussian process is still largely unsolved. To be more specific, the most general ($\mathbb{R}$-valued) equation handled up to now as far as parameter estimation is concerned is of the form:

$$X_t^\theta = a + \theta \int_0^t b(X_u)\, du + B_t,$$

where θ is the unknown parameter, b is a smooth enough coefficient and B is a one-dimensional fractional Brownian motion. In contrast with this simple situation, our applications of interest (see the *Application Domains* Section) require the analysis of the following $\mathbb{R}^n$-valued equation:

$$X_t^\theta = a + \int_0^t b(\theta; X_u)\, du + \int_0^t \sigma(\theta; X_u)\, dB_t, \tag{4}$$

where θ enters non linearly in the coefficient, where σ is a non-trivial diffusion term and B is a d-dimensional fractional Brownian motion. We have thus decided to tackle this important scientific challenge first.

To this aim, here are the steps we have focused on in 2013:

- A better understanding of the underlying rough path structure for equation (4), carried out in [6]. This step allows a proper definition of stochastic integrals with respect to fractional Brownian motion in a wide range of contexts.
- Extension of pathwise stochastic integration to processes indexed by the plane in [19], which helps to the definition of noisy systems such as partial differential equations.
- Gaussian type bounds for equations driven by a fractional Brownian motion, obtained in [18], [7]. This is an important preliminary step for likelihood estimates for stochastic processes. Also notice the interesting central limit theorems exhibited in [24], in a context which is similar to our equation of interest.
- Numerical aspects of a maximum likelihood type procedure for an equation of the form (4), expressed in terms of Malliavin calculus tools (see [4]).

6.3.2. LAN property for fractional Brownian motion

We have first focused on an important statistical property of fractional Brownian paths on their own. Indeed, the local asymptotic normality (LAN) property is a fundamental concept in asymptotic statistics, which gives the asymptotic normality of certain estimators such as the maximum likelihood estimator for instance. In [5], we focus on the LAN property for the model where we observe a sample of n observations $\mathbf{X_n} = (X_1, ..., X_n)$ of a Gaussian stationary sequence. The sequence $(X_n)_{n\in\mathbb{N}}$, whose spectral density f_θ is indexed by a parameter θ, can admit antipersistence, long memory or short memory and be noninvertible. To be more specific, our main assumption is:

$$f_\theta(x) \sim_{x\to 0} |x|^{-\alpha(\theta)} L_\theta(x)$$

with L_θ a slowly varying function and $\alpha(\theta) \in (-\infty, 1)$. We prove the LAN property by studying an asymptotic expansion of the log likelihood and using some results on Toeplitz matrices. In particular, our assumptions are fulfilled by fractional Gaussian noises and autoregressive fractionally integrated moving average processes (ARFIMA(p, d, q)). We also obtain the LAN property for fractional Brownian motion.

6.3.3. Self-similarity properties and stable or Gaussian random fields

In 2009, C. Lacaux and H. Biermé carried on the study of some sample paths properties for an important class of anisotropic random fields called operator scaling random fields, which had been previously introduced by H. Biermé, M. Meerschaert and P. Scheffler (2007). To be more specific, the classical self-similarity property is replaced by the following operator scaling property:

$$\forall c > 0,\ \left(X(c^E x)\right)_{x\in\mathbb{R}^d} \overset{(d)}{=} c\,(X(x))_{x\in\mathbb{R}^d}\,, \tag{5}$$

where $c^E := \exp\left(E\,\ln(c)\right)$. In particular, the Hölder regularity properties of operator scaling Gaussian or stable harmonizable random fields have been expressed in terms of the matrix E. The method they used can be applied to study the modulus of continuity of many stable or Gaussian random fields. As example in 2011, with P. Scheffler, they have followed it to study multi-operator harmonizable stable random fields, which satisfy a local version of the operator scaling property and enjoy a regularity which may vary along the trajectories. In [20], it has been developed in the more general framework of conditionally sub-Gaussian random series. This allows to also study for example some multistable random fields, which have been introduced in (Falconer & al, 2009); for such a field X, the marginal $X(x)$ is a stable random variable whose index of stability can depend on x. In this paper, some conditions have been proposed to establish the uniform convergence of the series (on an eventually random ball), an upper bound for the modulus of continuity of its limit, an uniform control of the partial series ones and an explicit rate of convergence. Focusing on LePage random series, upper bounds of the modulus of continuity of some harmonizable stable or multistable random fields are provided. In the conference paper [11], [20] has then been applied to study the class of linear multifractional multistable motions. In particular, the upper bound obtained for the modulus of linear multifractional stable motion is the sharpest available.

We are also interested in self-similar processes indexed by manifolds in [8]. This study is motivated by the fact various spatial data are indexed by a manifold and not by the Euclidean space $\mathbb{R}^d$ in practical situations such as image analysis.

BONSAI Project-Team

6. New Results

6.1. High-throughtput sequence processing

- Within our collaboration with Montpellier (IRB and LIRMM) we published a paper on CRAC, a software for analysing short RNA sequences and detecting variations among them [5].
- We have been invited to contribute an invited book chapter on metatranscriptomic data analysis (*Methods in Molecular Biology*, in press). This chapter covers the complete bioinformatic analysis from raw reads to taxonomic assignation, and introduces our software SortMeRNA (see Paragraph 5.6). This is a joint work with team LABIS in Genoscope.
- Evguenia Kopylova defended her thesis on December, the 11th (*"New algorithmic and bioinformatic approaches for the analysis of data from high throughput sequencing"*, [1]). The second part of her work deals with a new read mapper for metagenomic sequence data.
- Within our collaboration with the Lille hospital, we developped a seed-based heuristics for the detection of lymphocyte rearrangements from high-throughput data. This method is implemented in the software Vidjil (see Section 5.7). Our results were presented at the Jobim conference [8], and a journal article was submitted.

6.2. RNA algorithms

- We have started a new collabarative project with Bielefeld Universität on an extension of *Algebraic Dynamic Programming*. We introduced a generic specification framework, called *inverted coupled rewrite systems* [9], that can deal with optimization problems on strings, trees, and arc-annotated sequences. It is based on the following ideas: the solutions of combinatorial optimization problems are the inverse image of a term rewrite relation that reduces problem solutions to problem inputs. A tree grammar is used to further refine the search space, and optimization objectives are specified as interpretations of these terms. All these constituents provide a mathematically precise and complete problem specification, leading to concise yet translucent specifications of dynamic programming algorithms.

6.3. Genomic rearrangements

- Within a collaboration with LIAFA (CNRS UMR 7089, and University Paris 7) we published a method for the assembling of ancestral gene orders from contiguous ancestral fragments [4].

6.4. Nonribosomal peptides

- Yoann Dufresne is starting a PhD thesis on computational biology for nonribosomal peptides (NRPs) under the supervision of Maude Pupin and Laurent Noe, after doing his master thesis with them. He already worked on the translation of the chemical structure of the NRPs into their monomeric structure. NRPs can be represented by their chemical structure that is a graph where the atoms are represented by nodes and the chemical bonds by arcs; or by their monomeric structure that is a graph where the monomers are represented by nodes and the chemical bonds between monomers by arcs. We designed a novel algorithm capable of localizing the monomers from a reference list in the chemical structures of peptides [7]. It is based on a heuristic that utilizes chemical information of NRPs. The preliminary results are encouraging, and should lead to further studies.

DYLISS Project-Team

6. New Results

6.1. Data integration

Participants: Jacques Nicolas, Andres Aravena, Charles Bettembourg, Jérémie Bourdon, Jeanne Cambefort, Guillaume Collet, Olivier Dameron, Damien Eveillard, Julie Laniau, Sylvain Prigent, Anne Siegel, Sven Thiele, Valentin Wucher.

Metabolic network reconstruction: combinatorial gap-filling method We introduced an exhaustive gap-filling procedure on the first metabolic network for a macroalgea (Ectocarpus Siliculosus). As this species is a non benchmark model, this issue is related to hard combinatorial optimization problems. To that matter, we took advantages of the latest improvement of Answer Set Programming solvers (combination of clasp and unclasp) and introduced a new model of the network expansion problem. [*G. Collet, D. Eveillard, S. Prigent, A. Siegel, S. Thiele*] [27]

Identification of functional gene units in non benchmark models We introduced the concept of "shortest genome segments" (SGS) to detect functional units on exotic species, such as extremophiles, that are by nature unrefined. They correspond to genome portion which contain a large density of genes coding for enzymes which regulate successive reactions of metabolic pathways. There identification is a hard optimization combinatorial problem. We relied on the declarative modeling power of answer set programming (ASP) to encode the identification of shortest genome segments and prove that SGS are stable in (i) computational time and (ii) ability to predict functional units when one deteriorates the biological knowledge [*D. Eveillard, A. Siegel, S. Thiele*] [26]

Refinement of regulatory network from genomic, expression data and functional unit data We integrated heterogeneous information from two types of network predictions to determine a causal explanation for the observed gene co-expression. We modeled this integration as a combinatorial optimization problem. We demonstrated that this problem belongs to the NP-hard complexity class. We proposed an heuristic approach to have an approximate solution in a practical execution time. Our evaluation showed that the E.coli regulatory network resulting from the application of this method has higher accuracy than the putative one built with traditional tools. Applications to the mining bacterium *Acidithiobacillus ferrooxidans* allowed analyzing the relevance of central regulators. [*A. Aravena, D. Eveillard, A. Siegel*] [23], [13] [Thesis]

Reconstruction of a protein interaction network for archaebacteria To gain insights into genomic maintenance processes in hyperthermophilic archaea, a protein-interaction network centered on informational processes of *Pyrococcus abyssi* was generated by affinity purification coupled with mass spectrometry. We have proposed a graph theoretic analysis of this network including statistical (e.g. clusterisation coefficients) and topological aspects (bicluster analysis, search of a maximal interaction skeleton), which helps network interpretation in terms of formation of complexes or interaction dynamics. [*J. Nicolas*] [20] [Online publication]

Knowledge evolution in ontologies We studied the impact of an ontology evolution on its structural complexity. As a case study we used sixty monthly releases of the Gene Ontology and its three independent branches i.e. biological processes (BP), cellular components (CC) and molecular functions (MF). For each release, we measured complexity by computing metrics related to the size, the nodes connectivity and the hierarchical structure. We showed that the variation of the number of classes and relations in an ontology does not provide enough information about the evolution of its complexity. However, connectivity and hierarchy-related metrics revealed different patterns of values as well as of evolution for the three branches of the Gene Ontology [*O. Dameron, C. Bettembourg*] [17], [14] [Online publication][Thesis]

Treatment process representation for breast cancer patients. The general cancer registry of Poitou-Charentes developed a multiple source information system covering diseases, anatomical structures and cytopathology. We proposed an algorithm for representing and analyzing the patient's treatment process. An expert compared the original data with our representation and computed a score of dissimilarity. The results showed that an integrated information system can successfully analyze the data to determine whether they comply with the guidelines [*O. Dameron*] [31].

AphidAtlas project We began a collaboration with the AphidAtlas project for defining the structure of an ontology of aphids anatomy and development [*O. Dameron*] [30].

6.2. Asymptotic dynamics

Participants: Anne Siegel, Oumarou Abdou-Arbi, Geoffroy Andrieux, Jérémie Bourdon, Jeanne Cambefort, Damien Eveillard, Michel Le Borgne, Vincent Picard, Sven Thiele, Santiago Videla.

Learning families of boolean signaling networks We propose the use of ASP to explore the space of feasible logic models of a signaling network. To that matter, we exhaustively enumerate the set of sub-optimal boolean logical models which are compatible with both the topology of a knowledge-based influence graph and the observed response of the system to several perturbations (phosphorylation datasets). We illustrate the importance of characterizing such a family of models in a global and exhaustive manner by revisiting a model of pro-growth and inflammatory pathways in human liver cells and studying the variability with the set of compatible models. [*A. Siegel, S. Thiele, S. Videla*] [18] [Online publication]

Control the steady-state response of qualitative signaling networks: intervention sets The minimal intervention set problem roughly consists in identifying the perturbation that can be undergone over a signaling network to predict a fixed expected behavior. We have provided a precise characterization of the minimal intervention set problem relying on three-valued logic and fixpoint semantics. We address this problem within ASP and using real-world biological benchmarks we show that it greatly outperforms previous work using dedicated algorithms. [*A. Siegel, S. Videla*] [19] [Online publication]

Reachability in dynamical signaling networks: cut sets In the scope of discrete finite-state models of interacting components, we present a novel algorithm for identifying sets of local states of components whose activity is necessary for the reachability of a given local state. Those sets are referred to as cut sets; they provide potential therapeutic targets that are proven to prevent molecules of interest to become active, up to the correctness of the model. Our method is based on the so-called Graph of Local Causality and form an under-approximation of the complete minimal cut sets of the dynamics. It makes tractable the formal analysis of very large scale networks. [*G. Andrieux, M. Le Borgne*] [28], [12] [Online publication][Thesis]

Exploring metabolism flexibility through quantitative study of precursor sets for system outputs We extended a Flux-Balanced-Analysis approach to quantify the precursor composition of each system output and to discuss the biological relevance of a set of flux in a given metabolic network. The composition is called contribution of inputs over outputs [AIO]. In order to further investigate metabolic network flexibility, we have proposed an efficient local search algorithm computing the extremal values of AIO coefficients. This approach enables to discriminate diets without making any assumption on the internal behaviour of the system. [*J. Bourdon, O. Abdou-Arbi, A. Siegel*] [15], [11] [Thesis]

6.3. Sequence annotation

Participants: François Coste, Aymeric Antoine-Lorquin, Catherine Belleannée, Guillaume Collet, Gaëlle Garet, Clovis Galiez, Laurent Miclet, Olivier Quenez, Jacques Nicolas, Valentin Wucher.

Refinement of mi-RNA regulation network thanks to concept analysis MicroRNAs (miRNAs) are small RNA molecules that bind messenger RNAs (mRNAs) to silence their expression. To improve the discrimination between true and false interactions during their prediction, we defined a repair process based on the hypothesis that the true graph is formed by interaction modules represented by formal concepts, i.e. set of miRNAs having the same regulation profile. To validate our hypothesis and method, we have extracted parameters from a biological miRNA/mRNA network and used them to build random networks. Each repaired network can be evaluated with a score balancing the number of edge changes and the conceptual adequacy in the spirit of the minimum description length principle. [*J. Nicolas, V. Wucher*] [32]

Analogical proportions and the factorization of information in distributive lattices. We have conducted theoretical studies to elucidate whether formal concept lattices can have properties that could be used in further studies. In this direction, analogical proportions are statements involving four entities, of the form 'A is to B as C is to D'. They play an important role in analogical reasoning. They have been formalized in both a propositional logic setting and an algebraic setting. We define and study analogical proportions in the general setting of lattices, and more particularly of distributive lattices. We discussed the decomposition of analogical proportions in canonical proportions as well as the resolution of analogical proportion equations, and illustrate especially on the case of Boolean lattices, which reflects the logical modeling. [*L. Miclet*] [24], [29]

Bioinformatics and Artificial Intelligence In this book chapter, we introduce the main objects studied in Bioinformatics at different levels (the macromolecules, their interactions as well as the knowledge formalization or extraction) and present meanwhile a survey of the contribution and influence of Artificial Intelligence to this research field on related key tasks (gene prediction, functional annotation, structure prediction, transcriptomics analysis, network acquisition and analysis, knowledge integration and formalization, information retrieval and extraction from documents, ...). [*F. Coste*] [33]

Genome studies: fast assembly and SNP identification This work is a follow-up of collaborations with the GenScale team and the GenOuest platform. We reported the first identification of a set of SNPs isolated from the genome of *I. ricinus* - an important vector of pathogens in Europe, by applying a reduction of genomic complexity, pyrosequencing and new bioinformatics tools[21] [Online publication]. We also contributed to show that the genome assembly program MINIA is successfully able to assemble a 100 Mbp genome on a very low-end, low-power system with 512 MB RAM and a 32 GB flash drive such as a Raspberry Pi. [*G. Collet, O. Quenez*] [34][Online publication]

GENSCALE Project-Team

6. New Results

6.1. NGS methodology

Participants: Dominique Lavenier, Claire Lemaitre, Pierre Peterlongo, Guillaume Rizk, Anaïs Gouin, Fabrice Legeai.

- **Efficient Kmer counting**: Counting all the substrings of length k (k-mers) in DNA/RNA sequencing reads is the preliminary step of many bioinformatics applications. However, state of the art k-mer counting methods require that a large data structure resides in memory. Such structure typically grows with the number of distinct k-mers to count. We have developed a new streaming algorithm for that purpose which only requires a fixed user-defined amount of memory and disk space. This approach realizes a memory, time and disk trade-off. DSK is the first approach that is able to count all the 27-mers of a human genome dataset using only 4.0 GB of memory and moderate disk space (160 GB), in 17.9 h. DSK can replace a popular k-mer counting software (Jellyfish) on small-memory servers. [24]
- **Questionning the classical re-sequencing analyses approach**: Classical re-sequencing analyses are based on a first step of read mapping, then only mapped reads are taken into account in following analyses such as variant calling. We investigated the sources of unmapped reads in aphid re-sequencing data of 33 individuals, and we demonstrated that these reads contain valuable information that should not be discarded as usually done in such analyses. We proposed also an approach to extract this information, based on assembly and re-mapping. [34]
- **Repeat detection** A new algorithm was developed for detecting long similar fragments occurring at least twice in a set of biological sequences. The problem becomes computationally challenging when the frequency of a repeat is allowed to increase and when a non-negligible number of insertions, deletions and substitutions are allowed. The proposed algorithm, called Rime (for Repeat Identification: long, Multiple, and with Edits) performs this task, and manages instances whose size and combination of parameters cannot be handled by other currently existing methods. To the best of our knowledge, Rime is the first algorithm that can accurately deal with very long repeats (up to a few thousands), occurring possibly several times, and with a rate of differences (substitutions and indels) allowed among copies of a same repeat of 10-15% or even more. [17]

6.2. NGS applications

Participants: Dominique Lavenier, Claire Lemaitre, Pierre Peterlongo, Guillaume Rizk, Fabrice Legeai.

- **Participation to an international competition of assembly**: The process of generating raw genome sequence data continues to become cheaper, faster, and more accurate. However, assembly of such data into high-quality finished genome sequences remains challenging. Many genome assembly tools are available, but they differ greatly in terms of their performance and in their final output. More importantly, it remains largely unclear how to best assess the quality of assembled genome sequences. In this context, we have participated to the Assemblathon-2 competitions, which purpose was to assess current state-of-the-art methods in genome assembly. Globally, the cumulative z-scores of different assembly criteria set our assembly strategy in the 4th position compared to other competitors (21 groups). [12]
- **Assembly on Raspberri Pi**: Current Assembly tools require computers with large memory configuration. In order to demonstrate the efficiency of our low memory footprint assembly tools, we assemble the genome of C. Elegans (100 Mbp) on the raspberry PI computer, a small system equipped with only 512 MB RAM and 32 GB flash drive. [42]

- **SNP detection on the tick** We took part of a population genetic study on the tick species *Ixodes ricinus*, the main vector species of human and animal vector-borne diseases in Europe. In this framework, we proposed the first identification of a set of SNPs isolated from the genome of *I. ricinus*, by applying, among others a new tool developed in the GenScale team: discoSnp. The main advantage of this tool is to be able to detect SNPs without the use of a reference genome, which is crucially lacking for the tick species. Among the detected SNPs, 384 were selected, according to their minimal and maximal coverage and context sequences for experimental validation. Among them, 368 (95.8%) were biologically validated, demonstrating the precision of discoSNP.[23]
- **NGS analyses on insect models** We achieved the transcriptome assembly and analyzed the differential expression of an important noctuid pest. [22], [18]. Using gene expression data (RNA-Seq) in males, sexual females and asexual females of the pea aphid, we confirm theoretical models suggesting that the evolution of sex-biased gene expression may restrict the product of a sexually antagonistic allele to the sex it benefits.[19]
- **Genome sequencing and annotation**: We participated in the sequencing and annotation of several bacterial species of the Mollicute group. These bacteria are important pathogens of ruminants. The sequencing and annotation of their genomes confirmed their pathogenic features and phylogenetic location in the tree of Mollicutes. This is the first step before comparative genome analyses to unravel the genetic basis of mycoplasma pathogenicity and host specificity. [15], [16], [21]

6.3. HPC and parallelism

Participants: Dominique Lavenier, Rumen Andonov, Guillaume Chapuis, François Moreews, Charles Deltel.

- **Improving time performances of Mapping quantitative trait loci (QTL)** : we have developed a fast implementation of QTLMap, which takes advantage of the data parallel nature of the problem by offsetting heavy computations to a graphics processing unit (GPU). This new implementation performs up to 75 times faster than the previous multicore implementation, while maintaining the same results and level of precision . This speedup allows one to perform more complex analyses, such as linkage disequilibrium linkage analyses (LDLA) and multiQTL analyses, in a reasonable time frame. [13]
- **Integration of parallelism in bioinformatics workflows**: We propose a Model-Driven Architecture approach for capturing the complete design process of bioinformatics workflows. This approach is applied to graphical workflow editors and allows to quickly convert a workflow prototype in a parallel implementation. This work can have an impact on the way bioinformaticians implement their analysis and increase their productivity.[30]
- **Parallel assembly on FPGAs**: This research work proposes a method to reduce the overall time for assembly by using pre-processing of the short read data on FPGAs and processing its output using Velvet. We demonstrate significant speed-ups with slight or no compromise on the quality of the assembled output.[32]
- **All-Pairs Shortest Paths with multi-GPU** We propose a new algorithm for the All-Pairs Shortest Paths problem for graphs with good partitioning properties and its multi-GPU implementation. Our implementation targets large graphs (up to 10^6 vertices) and allows graphs with negative edges to be computed. [35]

6.4. Protein structures

Participants: Rumen Andonov, Guillaume Chapuis, Dominique Lavenier, Mathilde Le Boudic-Jamin, Antonio Mucherino, Douglas Goncalves.

- **A book on distance geometry problems (DGP).** This is a collection of invited papers on the topic "distance geometry" [38]. Among the other contributions, it contains a survey on "distance geometry" and "structural biology", which tries to function as a bridge between two scientific communities: computer science and biology. It presents some recent developments in the field by using a language common to the two communities [37]. In another contribution, the complexity of the DGP is discussed: even if this problem is NP-hard in general, we noticed a polynomial complexity on instances of DGP related to protein conformations (in the case all the available distances are exact)[36].

- **DGP with interval data.** In our preliminary works on the discretization of the Distance Geometry Problem (DGP), we considered instances where all distances were supposed to be exactly known. When biological molecules are concerned, however, this is not generally the case. We worked therefore for considering the full-atom representation of the protein backbone, where some of the distances are subject to uncertainty within a given nonnegative interval. We showed that the discretization is still possible in this case, and proposed the iBP algorithm to solve the discretized DGP. [20]
- **New pruning device for DGP.** After the discretization, DGPs can be solved by a branch-and-prune (BP) algorithm, which is potentially able to enumerate the entire solution set. This solution set, however, can be very large for some instances, while only the most energetically stable conformations are of interest. We worked therefore for integrating the BP algorithm with two new energy-based pruning devices. Our computational experiments showed that the newly added pruning devices were actually able to improve the performance of the algorithm, as well as the quality (in terms of energy) of the conformations in the solution set. [28]
- **Discretization orders for the DGP.** The main assumption that allows for the discretization of DGPs is strongly based on the order in which the atoms of the molecule are considered. The "natural" order of the atoms in the amino acid chain does not always allow for the discretization. We tried to find discretization orders in several ways, based on different approaches. In [31], we extended a previously proposed greedy algorithm that is able to deal with interval data (inexact distances). In [27], we handcrafted some discretization orders for the side chains of the amino acids involved in the protein synthesis. In [29], we proposed a heuristic, which outperforms, on large instances, the greedy algorithm previously proposed.
- **DGP with Clifford Algebra** The BP algorithm for the DGP is based on a search on the tree, where nodes of the tree belonging to a common layer provide the possible positions for the same atom of the molecule. When interval data are given, a curve in 3d (containing the possible positions for the atom) can be associated to one of such nodes. Since it is generally not necessary to have protein conformations with a precision higher than 1A, sample points on these curves can be chosen. The way to choose these sample points is not, however, a simple task. This is the reason why we are trying to make this selection process adaptive, by exploiting Clifford Algebra to this purpose. Preliminary studies in this direction were presented in [25]
- **Parallel seed-based approach to protein structure similarity detection** We have developed a new parallel heuristic-based approach to structural similarity detection between proteins that discovers multiple pairs of similar regions. We prove that returned alignments have RMSDc and RMSDd lower than a given threshold. Computational complexity is addressed by taking advantage of both fine- and coarse-grain parallelism. [26]
- **Datamining.** The selection of features that describe samples in sets of data is a typical problem in data mining. A crucial issue is to select a maximal set of pertinent features, because the scarce knowledge of the problem under study often leads to consider features which do not provide a good description of the corresponding samples. The concept of consistent biclustering of a set of data has been introduced to identify such a maximal set. The problem can be modeled as a 0–1 linear fractional program, which is NP-hard. We reformulated this optimization problem as a bilevel program, and we proposed a heuristic for its solution [39].

IBIS Project-Team

5. New Results

5.1. Analysis of gene regulatory networks by means of piecewise-linear (PL) models

GENETIC NETWORK ANALYZER (GNA) is a tool for the qualitative modeling and simulation of the dynamics of gene regulatory networks by means of PLDE models, as described in Section 4.1 . GNA has been integrated with the other bioinformatics tools distributed by Genostar (http://www.genostar.com/). Version 8.5 of GNA was released by IBIS and Genostar this year. This version is an update of version 8.0, deposited at the Agence pour la Protection des Programmes (APP). Some bugs have been corrected in the new version and the program has been adapted to the latest versions of Java and the software platform of Genostar. Version 8.5 supports the SBML standard and is also capable of exporting its models to the newly-developed standard for qualitative models, SBML Qual. This standard has been elaborated by the community of developers of logical and related modeling tools (CoLoMoTo), in which the GNA developers participate, and has been described in detail in a paper published in *BMC Systems Biology* [6].

The predictions obtained with the help of GNA are purely qualitative, describing the dynamics of the network by means of a state transition graph. While a qualitative analysis is appropriate for certain problems, the absence of precise quantitative predictions may not be desirable in others, such as the analysis of a limit cycle or the design of a controller for a synthetic network. The quantitative study of PLDE models of gene regulatory networks is hindered by the fact that the step functions describing the logic of regulatory interactions lead to discontinuities in the right-hand side of the PLDE models (Section 3.1). This has motivated extensions of the PLDE models based on differential inclusions and Filippov solutions. As of now, no numerical simulation tool for the simulation of these Filippov extensions is available.

In collaboration with Vincent Acary and Bernard Brogliato of the BIPOP project-team, we have shown how tools developed for the simulation of nonsmooth mechanical, electrical and control systems can be adapted for this purpose, in a paper published in *Physica D* [2]. We have presented a method for the numerical analysis of one proposed extension, called Aizerman–Pyatnitskii (AP)-extension, by reformulating the PLDE models as mixed complementarity systems (MCSs). This allows the application of powerful methods developed for this class of nonsmooth dynamical systems, in particular those implemented in the SICONOS PLATFORM developed by BIPOP. We have also shown that under a set of reasonable biological assumptions, putting constraints on the right-hand side of the PLDE models, AP-extensions and classical Filippov extensions are equivalent. This means that the proposed numerical method is valid for a range of different solution concepts. We have illustrated the practical interest of our approach through the numerical analysis of three well-known networks developed in the field of synthetic biology.

5.2. Inference of bacterial regulatory networks from reporter gene data

The use of fluorescent and luminescent reporter genes allows real-time monitoring of gene expression, both at the level of individual cells and cell populations (Section 3.2). In order to fully exploit this technology, we need methods to rapidly construct reporter genes, both on plasmids and on the chromosome, mathematical models to infer biologically relevant quantities from the primary data, and computer tools to achieve this in an efficient and user-friendly manner. For instance, in a typical microplate experiment, 96 cultures are followed in parallel, over several hours, resulting in 10,000-100,000 measurements of absorbance and fluorescence and luminescence intensities. Over the past few years, we put into place an experimental platform and data analysis software, notably the WELLREADER program (Section 4.2), to allow biologists to make the most out of the information contained in reporter gene expression data. A web-based version of WELLREADER is currently in preparation. Valentin Zulkower has analyzed the measurement models underlying WELLREADER, work was presented at the *Journées Ouvertes Biologie, Informatique et Mathématiques (JOBIM'13)* [13] and submitted for publication.

The above tools have been used in a series of studies directed at the experimental mapping of gene regulatory networks in *E. coli*. A first example, which was carried out in the framework of the PhD thesis of former IBIS member Guillaume Baptist, concerns the development of a new screening methodology for identifying all genes that control the expression of a target gene through genetic or metabolic interactions. The screen combines mutant libraries with luciferase reporter constructs. Instead of a static picture of gene expression, this method allows dynamical monitoring in different environmental conditions. Mutants with interesting phenotypes can thus be selected based on multiple criteria, and the expression dynamics of the target gene can be extensively characterized. The method has been applied to the identification of the direct and indirect regulators of the gene *acs* in *Escherichia coli*. We confirmed known genetic regulators of the gene and identified new regulatory influences, many of which involve metabolic intermediates or metabolic sensing. An analysis of mutants involved in glycolysis and glucose transport demonstrates that the classical model of catabolite repression in *E. coli* needs to be amended. A paper describing the above work was published in *Nucleic Acids Research* this year [3].

A second example is a study, led by Stéphan Lacour in collaboration with Akira Ishihama and Hiroshi Ogasawara in Japan, on the lifestyle adaptation of *E. coli*. The study concerns the switch between swimming motility and biofilm formation in response to changes in environmental growth conditions. The stationary phase sigma factor RpoS is an important regulator of this switch since it stimulates adhesion and represses flagellar biosynthesis. By measuring the dynamics of gene expression, we show that RpoS inhibits the transcription of the flagellar sigma factor, FliA, in exponential growth phase. RpoS also partially controls the expression of CsgD and CpxR, two transcription factors important for bacterial adhesion. We have demonstrated that these two regulators repress the transcription of *fliA*, *flgM* and *tar*, and that this regulation is dependent on the growth medium. CsgD binds to the flgM and fliA promoters around their -10 promoter element, strongly suggesting direct repression. The results show that CsgD and CpxR also affect the expression of other known modulators of cell motility. An updated structure of the regulatory network controlling the choice between adhesion and motility was proposed in the paper based on this work, published in the *Journal of Bacteriology* [7].

A third study, published in *Research in Microbiology* [8], also focuses on the alternative sigma factor RpoS. The small protein Crl increases the interaction between RpoS and RNA polymerase and thereby activates certain RpoS-dependent promoters. However, the growth-phase dependence of the interaction of Crl with different forms of polymerase remains unknown. We have used 41 GFP transcriptional fusions to study the dynamics of gene regulation by RpoS and Crl during growth transition from exponential to stationary phase in *Escherichia coli*. This has confirmed that RpoS can regulate gene expression in exponential phase, both positively and negatively. Crl slightly stimulates transcription by RpoS in exponential phase and controls a subset of RpoS-dependent genes in stationary phase. Growth temperature strongly affects induction of specific promoters by RpoS, whereas its impact on gene regulation by Crl is much less significant. In addition, we have identified five new genes regulated by Crl (*ada*, *cbpA*, *glgS*, *sodC* and *flgM*), and demonstrated that Crl improves promoter binding and opening by RpoS-containing RNA polymerase at the *hdeA* promoter. The study also shows that Crl is a cognate enhancer of RpoS activity under different growth conditions, since its deletion has no effect on genes transcribed by other sigma factors.

In the framework of the PhD thesis of Diana Stefan, a network inference method developed by Eugenio Cinquemani and colleagues, first published in *Bioinformatics* in 2010, has been applied to reporter gene data from the network regulating motility of *E. coli*, described above. The results are currently being prepared for publication.

5.3. Models of carbon metabolism in bacteria

Kinetic models capture the dynamics of the large and complex networks of biochemical reactions that endow bacteria with the capacity to adapt their functioning to changes in the environment. In collaboration with Matteo Brilli and Daniel Kahn (Inra and Université Claude Bernard in Lyon), we previously developed an approximate model of central metabolism of *E. coli*, using linlog kinetics, and estimated the parameter values from metabolomics, transcriptome, proteomics data sets, as described in an article published in *Bioinformatics* in 2011. The results of this study revealed the fundamental role played by the identifiability of the model parameters, an issue often overlooked in systems biology. This prompted us for a thorough investigation of the concepts of structural identifiability (in presence of perfect, idealized data), practical identifiability (in presence of noisy and limited amounts of data), and the relations between the two. In addition, we looked into the implications of this analysis for the reduction of nonidentifable to identifiable models. While having a solid mathematical basis, the study was tailored to the actual experimental practice, and resulted in a practical model reduction method that improves upon our previous approach in case of large measurement noise. This study, and the results from its application to both *in-silico* case studies and state-of-the-art datasets, were reported in a paper that appeared in the *Journal of Mathematical Biology* this year [4]. Although the theoretical development has focused on linlog models and related classes of approximate kinetic models, it is important to note that the results also bear on more general classes of nonlinear models of metabolism.

A second line of work is based on the use of classical kinetic models that are, in comparison with the above-mentioned linlog models, much reduced in scope (the focus is on the metabolic and genetic regulation of the glycolysis pathway) and granularity (individual reactions are lumped together). The models, developed by Delphine Ropers, are being calibrated using experimental data from IBIS group and the group of Jean-Charles Portais at Inra/INSA in Toulouse, and will be used to understand some key mechanisms in the adaptation of *E. coli* to the exhaustion of glucose. The PhD thesis of Manon Morin, in the framework of a collaboration supported by a Contrat Jeune Scientifique Inra-Inria, will further develop these research directions. In the framework of their PhD theses, Stéphane Pinhal and Valentin Zulkower also study specific aspects of carbon metabolism, using both models and experimental data. In parallel, we collaborate with Myriam Ferro at CEA in Grenoble to investigate how state-of-the-art measurements of the absolute concentrations of enzymes in *E. coli* can be integrated with other high-throughput data sets and kinetic models. The first results of this collaboration were accepted for publication in *Molecular and Cellular Proteomics* early 2014.

5.4. Stochastic modeling and identification of gene regulatory networks in bacteria

At the single-cell level, the processes that govern gene expression are often better described by stochastic models. Modern techniques for the real-time monitoring of gene expression in single cells enable one to apply stochastic modelling to study the origins and consequences of random noise in response to various environmental stresses, and the emergence of phenotypic variability. The potential impact of single-cell stochastic analysis and modelling ranges from a better comprehension of the biochemical regulatory mechanisms underlying cellular phenotypes to the development of new strategies for the control of cell populations and even of single cells. General modeling paradigms, such as the Chemical Master Equation, exist for the description of stochastic dynamics at the single-cell level. However, due to the complexity of the interactions, current studies have often preferred to focus on specific cases of interest by *ad-hoc* modeling and analysis. In addition, theoretical and practical challenges inherent in the inference of stochastic models from biological experimental data have limited the development of general identification approaches.

Work in IBIS on the probabilistic modeling of gene expression and interaction dynamics at the level of individual cells is centered around two main challenges. On the one hand, we address identification from microscopy data and analysis of the arabinose uptake dynamics in *E. coli* upon glucose exhaustion. Starting from a reduced arabinose uptake model, Eugenio Cinquemani and Michel Page are working on methods for the estimation of unknown stochastic model parameters from statistical population snapshot data collected by fluorescence microscopy experiments. Analysis of the model focuses on the problem of model-based real-time single-cell state estimation, with feedback control applications in mind, in collaboration with Alfonso Carta (BIOCORE). Based on a stochastic model reflecting switch-like dynamics in the form of sigmoidal reaction rates, taking a Chemical Master Equation model with cell-dependent parameters as a gold standard desdcription of the system, a Chemical Langevin Equation approximation is proposed as a convenient approximation of the model for observer design purposes. On top of this model approximation, a so-called Square-Root Unscented Kalman filter (SRUKF) is designed. Based on simulations of a realistically tuned model, SRUKF is found to perform as good as much heavier particle filters based on the gold-standard model.

Results were presented at the *European Control Conference (ECC)* in 2013 [11], where we also showed that including extrinsic noise effects explicitly in the estimation process allows one to improve the knowledge of the hidden states.

On the other hand, we investigate the use mixed effects-modelling and identification techniques to characterize single-cell profiles in isogenic cell populations. Mixed-effects models are hierarchical models where parametric response profiles of individuals is subject to inter-individual parameter variability following a common population distribution. In collaboration with Gregory Batt (CONTRAINTES) and Giancarlo Ferrari-Trecate (University of Pavia, Italy), we are adapting and applying existing procedures from pharmacokinetics to the context of microfluidic data, with focus on the budding yeast response to osmolarity shocks. The first results of the work were presented at the *European Control Conference (ECC)* this year [12]. Rigorous model identification and validation steps are performed on data from real-time control experments performed in Pascal Hersen's lab at Université Paris Descartes, for both mixed-effects modelling and for the competing method of moment-based identification. Results show the tendency of mixed-effects modelling to avoid overfitting for this system, trading fitting performance for validation performance and hence predictive capabilities. The work is being further developed and the collaboration tightened by the ongoing visit of Andres Gonzalez, PhD candidate at the University of Pavia, to CONTRAINTES and IBIS. A first journal publication is in preparation, which will be followed by extensions and refinements of the method.

In parallel, work concerning the study of noise propagation in gene regulatory networks is carried out in collaboration with Irina Mihalcescu (Université Joseph Fourier). Finally, collaboration of Eugenio Cinquemani with Marianna Rapsomaniki, PhD student affiliated with Zoi Lygerou (University of Patras, Greece) and John Lygeros (ETH Zürich, Switzerland), has been devoted to the analysis of data from Fluorescence Recovery After Photobleaching (FRAP) experiments. It has given rise to a novel method for reconstructing physical diffusion and immobilization parameters at the level of single cells. The method has been applied to nuclear species of mammalian cells and results are part of a journal paper under revision.

5.5. Shared control of gene expression by global physiological effects and specific regulators

Gene expression is controlled by the joint effect of (1) the global physiological state of the cell, in particular the activity of the gene expression machinery, and (2) DNA-binding transcription factors and other specific regulators. While many studies have focused on networks of transcription factors, the analysis of the relative contributions of both transcription factors and global effects of the physiological state has received relatively little attention thus far.

In the framework of the PhD thesis of former IBIS member Sara Berthoumieux, we have developed a model-based approach to distinguish between these two effects using time-resolved measurements of promoter activities. We have demonstrated the strength of the approach by analyzing a circuit involved in the regulation of carbon metabolism in *E. coli*, consisting of two pleiotropic regulators of the cell (Crp and Fis), the gene *acs* encoding the enzyme acetyl-CoA synthetase (Acs), and the signaling metabolite cyclic AMP (cAMP) which activates Crp. *acs* is strongly expressed in the absence of glucose and is thus an excellent indicator of the transcriptional response of carbon metabolism to a growth-phase transition.

Our results show that the transcriptional response of the network is controlled by the physiological state of the cell and the signalling metabolite cAMP. The (surprising) absence of a strong regulatory effect of transcription factors suggests that they are not the main coordinators of gene expression changes during growth transitions, but rather that they complement the effect of global physiological control mechanisms. This change of perspective has important consequences for the interpretation of transcriptome data and the design of biological networks in biotechnology and synthetic biology. An article presenting the above results was published in *Molecular Systems Biology* this year [5] and selected as an Editor's choice in *Science* (http://ibis.inrialpes.fr/article1040.html).

In the above-mentioned work, the activity of the gene expression machinery was indirectly measured, by monitoring the activity of a constitutive gene, that is, a gene whose expression does not depend on any specific regulators but only on the activity of the gene expression machinery. There exists a huge literature on the molecular mechanisms coupling the activity of the gene expression machinery to changes in the nutritional quality of the environment, but a quantitative and dynamic picture of this very complicated regulatory system is still missing. Delphine Ropers and Edith Grac as well as Nils Giordano are developing models to achieve this, from bottom-up and top-down perspectives, respectively.

5.6. Control of regulatory networks in bacteria

A bacterial cell adapts its growth rate to the environment, notably to the availability of nutrients providing the molecular building blocks and the energy required for growth. Upon a change in the environment, the global physiology of the cell is adjusted in parallel with the adaptation of the growth rate. In the context of the PhD thesis of former IBIS member Jérôme Izard, we have studied the relation between the gene expression machinery, the global physiology of the cell, and the growth rate from a different perspective. Our aim was to change the mechanisms regulating the activity of the gene expression machinery in such a way so as to be able to externally control the growth rate of the cell.

More precisely, we have engineered an *E. coli* strain in which the transcription of an essential component of the global gene expression machinery, RNA polymerase, is under the tight control of an inducible promoter. By adjusting the inducer concentration in the medium we can adjust the RNA polymerase concentration and thereby reversibly tune the growth rate of the bacterium between zero and the maximal growth rate. The growth arrest is completely reversed when RNA polymerase is provided again. The analysis of the transcriptome at growth rates restricted by the concentration of RNA polymerase confirms that the concentration of RNA polymerase is the major determinant of changes in gene expression patterns. Our modified *E. coli* strain provides a novel way of setting growth rate in a tunable, reversible, modular, and medium-independent way. The strain, described in a paper submitted for publication, opens new perspectives for studying the mechanisms of growth control as well as for developing biotechnological applications, the subject of the post-doctoral fellowship of Cindy Gomez Balderas-Barillot. We have submitted a patent proposing such applications, which underlies the technology transfer activities undertaken in the recently-started Reset project (Section 7.2).

MAGNOME Project-Team

6. New Results

6.1. Adopting new computing paradigms

Participants: David James Sherman [correspondant], Pascal Durrens, Natalia Golenetskaya, Florian Lajus, Xavier Calcas.

Analyses in comparative genomics are characteristically forms of datamining in high-dimension sets of relations between genes and gene products. For every linear increase in genomic data, these relations can grow at worst geometrically.

Natalia Golenetskaya's thesis[12] developed an integrated architecture that we call *Tsvetok*, which combines a novel NoSQL storage schema, domain-specific Map-Reduce algorithms, and existing resources to efficiently handle the fundamentally data-parallel analyses encountered in comparative genomics [48], [42], [51]. Tsvetok components are deployed in MAGNOME's private cloud and have been extensively tested using data and use cases derived from log analyses of the Génolevures web resource. We designed Map-Reduce solutions for the principal whole-genome analyses used by MAGNOME for comparative genomics, in particular new distributed algorithms for systematic identification of gene fusion and fission events in eukaryote genomes, and large-scale consensus clustering for protein families. These examples illustrate two strategies that can be used to scale algorithms in a Map-Reduce setting[12].

1. Converting classical graph-based algorithms with message propagation: instead of traversing a graph, which would incur high latency, information is sent forward in waves, and synchronized later. Some of the intermediate computations may be redundant, but overall running time is minimized.
2. Iterative sampling strategies, which run the standard algorithm on carefully chosen subsets, and later compute a consensus of the intermediate results. The iterations may take some time to converge, but the individual instances can be run within one machine.

Florian Lajus extended the Magus software platform to use the NoSQL storage components in Tsvetok, and has validated it on a large collection of fungal genomes. Xavier Calcas is currently integrating the Galaxy platform [8] with Magus.

6.2. Improving inference of metabolic models

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

The Pantograph approach uses an annotated "scaffold" (reference) model and a collection of complementary predictions of homology between scaffold genes and target genes. The basis of the method is a weighing of the homology evidence to decide whether a reaction that is present in the scaffold ought be be present in the target.

We have improved on the method in two ways. First, we model the implicit knowledge represented in the boolean formula of each gene association, to derive hypotheses about the explicit role of individual genes; for example, a gene association $(S_1 \wedge S_2) \vee (S_1 \wedge S_3)$ may implicitly represent an enzyme complex formed from two subunits, the first encoded by gene S_1, and the second encoded by two paralogous genes S_2 and S_3 (figure 2). By using these hypotheses to rewrite gene associations, we improve the decision of whether a reaction is present in the target or not.

---

[8] http://usegalaxy.org

`../../../../projets/magnome/IMG/tsvetok.png`

Figure 1. General architecture of Tsvetok, showing the role of NoSQL (Apache Cassandra) and Map-Reduce (Apache Hadoop) paradigms

Second, we have adopted an abductive strategy for inferring reactions. In this strategy we consider that it is the reactions that explain the genes observed in the target genome. In the corresponding abductive logic program, the observations are the genes in the target, the integrity constraints are the rules that rewrite gene associations, and the hypotheses to be abduced are the reactions in the model. The scaffold model is compiled into a set of facts and predicates that express the reactions, their gene associations, and the integrity constraint rules; the abducibles generate assertions that specific reactions are in the target model. Combined with the facts of the genes observed in the target, this program generates, through abduction, the set of target reactions that explain the greatest number of genes.

The advantage of this approach is that it can invent, through specialization, reactions that are not present *per se* in the scaffold model.

`../../../../projets/magnome/IMG/sbgn_col.png`

Figure 2. An explicit model that is one possible explanation of the gene association $(S_1 \wedge S_2) \vee (S_1 \wedge S_3)$

6.3. Knowledge-based generalization of metabolic models

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

There is an inherent tension between detail and understandability in large metabolic networks: detailed description of individual reactions is needed for simulation, but high-level views of reactions are needed for describing pathways in human terms. We defined knowledge-based methods that factor similar reactions into “generic” reactions in order to visualize a whole pathway or compartment, while maintaining the underlying model so that the user can later “drill down” to the specific reactions if need be[22], [23], [26] This method is available as a Python library from http://metamogen.gforge.inria.fr/.

Figures 3 and 4 illustrate model generation for *Yarrowia lypolitica* fatty acid oxidation in the peroxisome. Molecular species are represented as circular nodes, and the reactions as square ones, connected by edges to their reactants and products. Ubiquitous species (e.g. *oxygen*, *water*, *ATP*) are of smaller size and colored gray. Non-ubiquitous species are divided into fifteen equivalence classes and colored accordingly (red/blue for trivial species/reaction equivalence classes, different colors for non-trivial equivalence classes). The size of the model does not allow for readability of the species labels, thus we do not show them (figure 3).

The generalization algorithm identifies equivalent molecular species using an ontology, and groups together reactions that operate on the same abstract species. It finds the greatest generalization the preserves stoichiometry. The generalized model represents quotient species and reactions. For example, the violet *unsaturated FA-CoA* node is a quotient of *hexadec-2-enoyl-CoA*, *oleoyl-CoA*, *tetradecenoyl-CoA*, *trans-dec-2-enoyl-CoA*, *trans-dodec-2-enoyl-CoA*, *trans-hexacos-2-enoyl-CoA*, *trans-octadec-2-enoyl-CoA*, and *trans-tetradec-2-enoyl-CoA* (colored violet in figure 3). In a similar manner, the light-green *acCoA oxidase* quotient reaction, that converts *fatty acyl-CoA* (yellow) into *unsaturated FA-CoA* (violet), generalizes six corresponding light-green reactions of the initial model (figure 3).

The generalized model describes β*-oxidation* in a more generic way: as a transformation of *fatty acyl-CoA* (yellow) into *unsaturated FA-CoA* (violet), then into *hydroxy FA-CoA* (dark green), *3-oxo FA-CoA* (magenta), and back to *fatty acyl-CoA* (with a shorter carbon chain); while the specific model describes the same process in more details, specifying those reactions for each of the *fatty acyl-CoA* species presented in the organisms' cell (e.g. *decanoyl-CoA*, *dodecanoyl-CoA*, etc.). That is why the *beta-oxidation* chain of the reactions in the initial model, transforming step-by-step the *fatty-acyl-CoA* with the longest carbon chain into the one with the shortest chain, in the generalized model appears as a cycle (generalizing all the *fatty-acyl-CoA*s into one species, regardless the chain-length).

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

6.4. Characterization of STAND protein families

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

In collaboration with Sven Saupe and Mathieu Paoletti from IBGC Bordeaux (ANR Mykimun), we worked on characterization of the STAND protein family in the fungal phylum. We established an *in silico* screen based on state-of-the-art bioinformatic tools, which – starting from experimentally studied sequences from *Podospora anserina* – allowed us to determine the first systematic picture of fungal STAND protein repertoire (ms. in preparation). Most notably, we found evidence of extensive modularity of domain associations, and signs of concerted evolution within the recognition domain. Both results support the hypothesis that fungal STAND proteins, originally described in the context of vegetative incompatibility, are involved in a general fungal immune system. In addition, we investigated improved protein domain representations and elaborated a grammatical modelling method [15], which will be used to elucidate mechanisms of formation and operation of the STAND proteins.

6.5. Avoiding stiffness in BioRica

Participants: David James Sherman [correspondant], Joaquin Fernandez.

We previously formalized two strategies for integrating discrete control with continuous models, coefficient switches that control the parameters of the continuous model, and strong switches that choose different models [29], [27]. While these strategies have proved useful for modeling hybrid systems in biotechnology [31] and medicine [28], the resulting system model can be inefficient when the different subsystems evolve at very different time scales. In order to improve the efficiency of the resulting simulations, we investigated the use of Kofman's Quantized State Systems (QSS), and demonstrated that the QSS approach can be adapted to BioRica [13]. On the strength of this demonstration, we invited Joaquin Fernandez from Kofman's lab to Magnome. Joaquin had previously implemented an efficient library for QSS simulation, and during his stay succeeded in adapting it to our hybrid modeling framework. In his approach, SBML models with events are compiled into a hybrid model, using a variant of Modelica for surface syntax and using the QSS library for efficient simulation.

`../../../../projets/magnome/IMG/pero_before_colored.png`

Figure 3. Yarrowia lypolitica fatty acid oxidation model before generalization. Reactions of the specific model are divided into fifteen equivalence classes, represented by different colours

`../../../../projets/magnome/IMG/pero_after.png`

Figure 4. Generalization of the Yarrowia lypolitica fatty acid oxidation model, described as a transformation of fatty acyl-CoA (yellow) into unsaturated FA-CoA (violet), then into hydroxy FA-CoA (dark green), 3-oxo FA-CoA (magenta), and back to fatty acyl-CoA (with a shorter carbon chain)

6.6. Applications in biotechnology and health

Participants: David James Sherman, Pascal Durrens [correspondant], Florian Lajus, Xavier Calcas.

Using MAGNOME's Magus system and YAGA software, we have successfully realized a full annotation and analysis of several groups of related genomes:

- Seven new genomes, provided to the Génolevures Consortium by the CEA–Génoscope (Évry), including two distant genomes from the *Saccharomycetales* were annotated using previously published Génolevures genomes.
- Twelve wine starter yeasts linked to fermentation efficiency.
- Five pathogenic (to human) and non pathogenic Nakaseomycetes.
- Two oleaginous strains with applications in biofuels.

Winemaking yeasts. In collaboration with partners in the ISVV, Bordeaux, we have assembled and analyzed 12 wine starter yeasts, with the goal of understanding genetic determinants of performance in wine fermentation. Analysis included identification of strain-specific gains and losses of genes linked both to niche specificity and to performance in industrial applications (article in prep.). A further combined analysis with 50 natural and industrial strains showed a pattern of introgression concentrated in industrial strains (article in prep.).

Oleaginous yeasts. In collaboration with Prof Jean-Marc Nicaud's lab at the INRA Grignon, we developed the first functional genome-scale metabolic model of *Yarrowia lipolytica*, an oleaginous yeast studied experimentally for its role as a food contaminant and its use in bioremediation and cell factory applications.

Using MAGNOME's Pantograph method (see section 5.2) we produced an accurate functional model for *Y. lipolytica*, MODEL1111190000 in BioModels [9], that has been qualitatively validated against gene knockouts. This model has been enriched by Anna Zhukova with ontology terms from ChEBI and GO.

Pathogenic yeasts. A further group of five species, comprised of pathogenic and nonpathogenic species, was analyzed with the goal of identifying virulence determinants [39]. By choosing species that are highly related but which differ in the particular traits that are targeted, in this case pathogenicity, we are able to focus of the few hundred genes related to the trait [16]. The approximately 40,000 new genes from these studies were classified into existing Génolevures families as well as branch-specific families.

---

[9] http://biomodels.net/

MORPHEME Project-Team

4. New Results

4.1. 3D reconstruction in fluorescence imaging

Participants: Emmanuel Soubies, Laure Blanc-Féraud, Sébastien Schaub.

This work was made in collaboration with Gilles Aubert, Laboratoire J.A. Dieudonné (CNRS,UNS).

We propose a new model for the reconstruction of biological structures using Multiple-Angle Total Internal Reflection Fluorescence Microscopy (MA-TIRFM). This recent microscopy technique allows the visualization of sub-cellular structures around the plasma membrane which is of fundamental importance in the comprehension of exchanges mechanisms of the cell. We present a 3D reconstruction method based on a shape prior information on the observed structures and robust to shot noise and background fluorescence. A novelty with respect to the state of the art is to propose a method allowing the recovery of multiple objects aligned along the axial axis.

TIRFM principle is based on the total internal reflection phenomenon of a light beam at the interface between two mediums of refractive indices n_i (incident) and n_t (transmitted) which produces an evanescent wave capable of exciting fluorophores that are near the coverslip surface. Excited fluorophores emit photons that are then collected by a CCD camera to produce a resulting 2D image (radial dimension). The 2D image formation is formulated as follows [29]:

$$S(x, y, \alpha) = I_0(\alpha) \int_0^{\infty} R(x, y, z) \exp\left(-\frac{z}{d(\alpha, \lambda)}\right) \, dz \tag{6}$$

where $S(x, y, \alpha)$ is the recorded image for the incident angle α, $R(x, y, z)$ denote the 3D unknown fluorophore density, $I_0(\alpha)$ is the intensity at the interface $d(\alpha, \lambda)$ is the penetration depth (theoretically known) and λ is the incident light wavelength. The problem is then to determine R in (1) from acquisitions S_α with different incident angles.

In order to solve this ill-posed inverse problem, we model the 3D unknown fluorophore density by a collection of parametrized objects defined on a state space $\mathcal{X} = P \times M$ by their location $\rho \in P$ and their marks (i.e geometric attributes $\omega \in M$). The optimization problem can be formulated as a minimization problem where both the number of objects in the model and their parameters have to be estimated. This difficult combinatorial optimization problem is tackled by using a Marked Point Process approach [36] which allows modelling interactions between the objects in order to regularize the inverse problem.

Figure 1 right shows the Root Mean Square Errors (RMSE) of each estimated parameter for different noise levels on simulated data. We obtain a hight accuracy reconstruction with an RMSE less than 10 nm for the radial position (x,y) and the radius. A larger RMSE (between 80 and 125 nm, depending on the noise level) is found on the axial position vesicles estimation. As we can see on figure 1 left, the error on the axial position estimation is due to the deepest objects ($>$ 300 nm), objects close to the glass interface are well estimated. Figure 1 right shows also the robustness of the model with respect to shot noise and background fluorescence since the errors remain almost constant with the increasing noise level. The proposed method have also been tested on a real sample of beads of known diameters in order to quantify the quality of the reconstruction. The obtained results are promising for feature estimation of predefined shape structures [17].

4.2. Depth-variant blind restoration for confocal microscopy

Participants: Saima Ben Hadj, Laure Blanc-Féraud.

../../../../projets/morpheme/IMG/result.jpg

Figure 1. Left : Reconstructions for different noise levels (colors represent z positions of objects). (a) Simulated sample, (b) → (c) → (d) : Reconstruction for an increasing level of noise. Right : RMSE for different noise levels.

3D images of confocal microscopy basically suffer from two types of distortions: a depth-variant (DV) blur due to the variation of the refractive index between the different mediums composing the system and the imaged specimen, and a Poisson noise due to photon counting process at the sensor.

The Point Spread Function (PSF) is depth-variant and its knowledge is crucial for the restoration of these images. Nevertheless, the PSF is inaccessible in practice since it depends on the optical characteristics of the biological specimen and thus needs to be estimated for each different specimen.

In our previous work [5], [4], we developed a method for the joint estimation of the specimen function (the sharp and clean image) and the 3D DV PSF by minimizing a criterion arising from the maximum a posteriori approach. The DV PSF is approximated by a convex combination of a set of space-invariant PSFs taken at different depths.

Recently, we proposed to consider additional constraints on the PSF coming from the optical system modeling [21], [6]. In fact, the confocal microscopy PSF is related to the magnitude of a complex function known as *complex valued-amplitude PSF* whose shape and support are given in the Fourier domain by the numerical aperture of the optical system [30], [35]. This latter is known as it is given by the system manufacturer. We incorporate this constraint in the joint PSF and image estimation algorithm [5] by using the Gerchberg-Saxton algorithm (GS) [31] since it allows to alternate constraints in the spatial and frequency domains. Numerical tests on a simulated image of a bead shell are encouraging (cf. figures 2 (a), (b), (c), and (d) presenting z-slices of the original image, simulated and reconstructed images). In particular, the added constraint allows to better estimate the PSF shape compared to the previous method [5] (cf. figures 2 (e), (f), and (g)).

(a) (b) (c) (d)

(e) (f) (g)

Figure 2. (Y, Z) slices of the simulated observation (a), true image (b) and PSF (e), estimated image (c) and PSF (f) with the additional PSF constraints, estimated image (d) and PSF (g) without the additional PSF constraints.

4.3. Head Tracking and Flagellum Tracing for Sperm Motility Analysis

Participants: Huei Fang Yang, Xavier Descombes, Grégoire Malandain, Sylvain Prigent.

This work is supported in part by ANR MOTIMO project.

Sperm quality assessment plays an important role in human fertility and animal breeding. One of the most important attributes for evaluating semen quality is sperm motility, according to the World Health Organization (WHO) report. When performed manually, semen analysis based on sperm motility is labor-intensive and subject to intra- and inter-observer variability. Computer-assisted sperm analysis (CASA) systems, in contrast, provide rapid and objective semen fertility assessment. In addition, they also offer a means of statistical analysis that may not be achieved by visual assessment. Hence, automated sperm motility analysis systems are highly desirable.

We present a computational framework designed to track the heads and trace the tails for quantitative analysis of sperm motility, which is illustrated in Figure 3 . Our framework includes 3 modules: head detection, head registration, and flagellum tracing. These modules are performed sequentially to obtain the head trajectories and flagellar beat patterns. First, the head detection module detects the sperm heads in the first image of the image data using a Multiple Birth and Cut (MBC) algorithm. The detections are the inputs to the head registration module for obtaining the head trajectories and angles of head rotation. We use a block matching method to register the heads in the subsequent images with respect to the positions and angles of those detected in the first image. This is different from other tracking methods that consider only the head positions. Finally, we propose a flagellum tracing algorithm, based on a Markov chain Monte Carlo (MCMC) sampling method, to obtain the flagellar beat patterns.

We validate our framework using two microscopy image sequences of ram semen samples that were imaged at two different conditions, at which the sperms behave differently. The results show the effectiveness of our framework [19].

4.4. Tree-like Shapes Distance Using the Elastic Shape Analysis Framework

Participants: Alejandro Mottini, Xavier Descombes, Florence Besse.

The analysis and comparison of tree-like shapes is of great importance since many structures in nature can be described by them. In the field of biomedical imaging, trees have been used to describe structures such as neurons, blood vessels and lung airways. Since it is known that axon morphology provides information on their functioning and allows the characterization of pathological states, it is of paramount importance to develop methods to analyze their shape and to quantify differences in structures

We have developed a new method for comparing tree-like shapes that takes into account both topological and geometrical information [14], [15]. Our metric combines the Elastic Shape Analysis Framework originally designed for comparing shapes of 3D closed curves in Euclidean spaces with a matching process between branches. Moreover, the method is able to compute the mean shape of a population of trees.

As a first application, we used our method for the comparison of axon morphology. The performance was tested on a group of 61 (20 normal, 24 type one mutant and 17 type two mutant) 3D images, each containing one axonal tree. We have calculated inter and intra class distances between them and implemented a classification scheme. We have compared our results with the ones obtained by three other methods. Results showed that the proposed method better distinguishes between the two populations than the other methods.

4.5. 3D Modeling of developing organisms

Participants: Gaël Michelin, Grégoire Malandain, Léo Guignard [Virtual Plants], Christophe Godin [Virtual Plants].

This work is made in collaboration with Patrick Lemaire (CRBM).

Image-based studies of developing organs or embryos produce a huge quantity of data. To handle such high-throughput experimental protocols, automated computer-assisted methods are highly desirable. We aim at designing an efficient cell segmentation method from microscopic images. Similary to another work [32], the proposed approach is twofold: first, cell membranes are enhanced or extracted by the means of structure-based filters, and then perceptual grouping (i.e. tensor voting) allows to correct for segmentation gaps (see figure 6). We assessed different structure-based filters as well as different perceptual grouping strategies to identify the most efficient combination, in term of result quality and computational cost [13].

../../../../projets/morpheme/IMG/sperm_analysis_workflow.png

Figure 3. Overview of the proposed framework. The input to our framework is an image sequence. The pre-processing step is to remove the inhomogeneous background and noise. The three main modules in our framework are head detection, head registration, and flagellum tracing. These three modules perform sequentially to obtain the head trajectories and flagellar beat patterns for sperm motility analysis. Note that the output of the head registration module is image sequences for each individual sperms in which the heads are registered. Here, we show the minimum intensity projection (MinIP) of the image sequence.

../../../../projets/morpheme/IMG/normal_tracing.jp

Figure 4. Original confocal microscopy image of an axonal tree (left) and its tracing (right) (maximum intensity projections).

../../../../projets/morpheme/IMG/means_normal_mut.

Figure 5. Mean normal (left) and mutant (right) axonal trees (2D projections).

(A) (B) (C) (D)

(E) (F) (G) (H)

Figure 6. Illustrations of the different steps of the algorithm: (A) a 2D slice of original image, (B) the resulting surface detector response, (C) the directional extrema of the response image, (D) the deduced binarisation of the cell membranes, (E) the result of Tensor Voting applied to binarised image, (F) the cells segmentation computed from (E), (G) a 3D view of original image, (H) a 3D view of the cells segmentation.

4.6. Spatio-temporal registration of embryo images

Participants: Grégoire Malandain, Léo Guignard [Virtual Plants], Christophe Godin [Virtual Plants].

This work is made in collaboration with Patrick Lemaire (CRBM).
Current imaging techniques can capture temporal sequences of 3D images with very high time resolution over several hours. Comparing sequences covering the same time period opens the way to the study of developmental variability. Stitching together sequences captured from different embryos may help producing a sequence covering the whole development of the animal of interest. For this, it is necessary to align two sequences in both time and space.

We developed a method to align two 3D+t time series, based on the detection and pairing of 3D+t landmarks. These landmarks, which correspond to periods of fast morphogenetic change, are deduced from the analysis of the non-linear transformations that allow to co-register pairs of consecutive 3D images in each sequence (see figure 7). [12].

4.7. Characterizing cell membrane properties

Participants: Sylvain Prigent, Xavier Descombes, Grégoire Malandain, Hélène Barelli [IPMC].

Some mammalian cells show striking differences in the acyl chain composition of their membrane phospholipids. In most cases, the majority of phospholipids bear one saturated and one monounsaturated acyl chains at positions 1 and 2 or the glycerol, respectively. However, some cells and notably neurons contain large amounts of phospholipids with a polyunsaturated fatty acyl chain, generally at position 2. The aim of this work is to compare the impact of the phospholipid polyunsaturation vs monounsaturation on the mechanical and functional properties of the plasma membrane.

For this task, we currently investigate how phospholipid insaturation affects the ability of specialized protein machineries involved in transport vesicle formation, by first detecting vesicles in 2D+t sequences of microscopic images of individual cells, and then tracking detected vesicles through the temporal sequences (see figure8) [23].

4.8. Tracking growing axons in 3D+t fluorescent two-photon microscopy images

Participants: Sylvain Prigent, Xavier Descombes [contact], Florence Besse, Caroline Medioni.

During the maturation of the nervous system, neuronal cells emit cellular extensions (dendrites, axons) allowing them to connect to other neurons, and thus, establish a network in which information is transmitted and/or stored. The formation of axonal extensions and directed migration of these extensions are two key processes controlling the morphology of neuronal cells, and then the number and nature of partners in a given network within a neuron. These two processes are controlled by both external factors to neuronal cell (guidance molecules, neurotrophic signals, ...) and internal factors (transcription factors, post-transcriptional regulators, regulators of the actin cytoskeleton or microtubules, ...). The goal of this work is to automatically extract axonal trajectories from images to then be able to model the processes controlling the morphology of neuronal cells.

The images we use are 3D+t images of growing drosophila brains obtained with a bi-photon confocal microscope. A single movie is about 200 3D frames that correspond to an acquisition every 5 minutes.

The developed method to extract axonal trajectories from 3D+t images is made of 3 main steps. The first step is to detect axonal tips on each 3D frame of the movie. This detection is performed using Marked Point Process. We designed a dedicated model based on an ellipse shape, a prior of no-overlapping between detected ellipses, and a data term calculated by:

$$d = \min(d_B(R_0, R_1), d_B(R_0, R_2), \max(d_B(R_0, R_3), d_B(R_0, R_4)))$$

`../../../../projets/morpheme/IMG/final_registration_bis_2.png`

Figure 7. Spatio-temporal registration of two time-series of embryo. Enlarged renderings indicate the registered timepoints. Notice that the temporal registration is not linear since the interval length between two registered time is different from one embryo to the next.

Figure 8. Left: detection of individual vesicles in one image of the sequence. Right: resulting paths of tracked vesicles through the 2D+t sequence.

where d_B denote the Bhattacharyya distance, R_0 the set of pixels inside the ellipse and R_1, R_2, R_3, R_4 four sets of pixels obtained by partitioning the ellipse contour into 4 regions around the ellipse vertex (see figure 9).

`../../../../projets/morpheme/IMG/axons_tip_zoom.png`

`../../../../projets/morpheme/IMG/axons_data_t`

(a) (b)

Figure 9. (a) Zoom on an axon tip. (b) Proposed ellipse model to detect an axon tip.

The second step consists to track the axons along the time frames by linking the tips detections. We designed an association tracking algorithm that builds a graph by connecting spatially close detections in neighboring frames. Negative costs have been introduced to favor long tracks. Then we ran sequentially the shortest path algorithm on this graph to obtain axons trajectories (see figure 10 (b)).

As the axonal tips size is close to the image resolution, the proposed method as the drawback to give false alarms trajectories. We then added a last processing step that aim to analyze the trajectories and remove those that do not correspond to axonal trajectories. This filtering removes three types of trajectories: 1) the trajectories that follow static regions of the image, 2) the short trajectories (less than 5 frames) and 3) random walk trajectories. An example of final automatic tracking is shown in figure 10 (c).

../../../../projets/morpheme/IMG/axo [illegible]

(a) (b) (c)

Figure 10. (a) 2D projection of the last frame. (b) Obtained tracks using the association tracking method. (c) Obtained tracks after false alarms removing.

4.9. A Hierarchical, Graph-cut-based Approach for Extending a Binary Classifier to Multiclass – Illustration with Support Vector Machines

Participants: Alexis Zubiolo, Eric Debreuve, Grégoire Malandain.

The problem of automatic data classification is to build a procedure that maps a datum to a class, or category, among a number of predefined classes. The building of such a procedure is the learning step. Using this procedure to map data to classes is referred to as classification or prediction. The procedure is therefore a classification, or prediction, rule. A datum (text document, sound, image, video, 3-dimensional mesh...) is usually converted to a vector of real values, possibly living in a high-dimensional space, also called signature. Offline, supervised learning relies on a learning set and a learning algorithm. A learning set is a set of signatures that have been tagged with their respective class by an expert. The learning algorithm input is formed by this set together with some parameters, its output being a prediction rule. Some learning algorithm, or method, apply only to the 2-class case. Yet, adapting such a binary classifier to a multiclass context might be preferred to using intrinsically multiclass algorithms, for example if it has strong theoretical grounds and/or nice properties; if free, fast and reliable implementations are available...The most common multiclass extensions of a binary classifier are the one-versus-all (OVA) (or one-versus-rest) and one-versus-one (OVO) approaches. In any extension, several binary classifiers are first learned between pairs of groups of classes. Then, all or some of these classifiers are called when predicting the classes of new samples. When the number of classes increases, the number of classifiers involved in the learning and the prediction steps becomes computationally prohibitive. Hierarchical combinations of classifiers can limit the prediction complexity to a logarithmic law in the number of classes (at best). Combinatorial approaches can be found in the literature. Because of their high learning complexity, these approaches are often disregarded in favor of an approximation trading optimality for computational feasibility. In our work, the high combinatorial complexity is overcome by formulating the hierarchical splitting problems as optimal graph partitionings solved with a minimal cut algorithm. In fact, as this algorithm performs only few additions and comparisons, its impact on the whole procedure is not significant. A modified minimal cut algorithm is also proposed in order to encourage balanced hierarchical decompositions (see Fig. 11). The proposed method is illustrated with the Support Vector Machine (SVM) as the binary classifier. Experimentally, it is shown to perform similarly to well-known multiclass extensions while having a learning complexity only slightly higher than OVO and a prediction complexity ranging from logarithmic to linear. This work has been accepted to the International Conference on Computer Vision Theory and Application (VISAPP 2014) [20].

../../../../projets/morpheme/IMG/svmtree.png

Figure 11. Type of tree that the proposed method builds during the learning stage (illustration with 5 classes). An example of classification of a new image signature is also illustrated by showing the visited nodes in boldface (read from root to leaf).

4.10. Classification of neurons to study Parkinson's disease

Participants: Alexis Zubiolo, Eric Debreuve, Xavier Descombes.

This work has been made in collaboration with Michèle Studer's team at iBV

In this project, the goal is to perform unsupervised classification of rat neurons in order to study the Parkinson's disease. The Institut de Biologie Valrose (iBV) provided us with 3-D images of rat cortices obtained by confocal microscopy. The discriminant features between normal and pathological neurons include the number of dendrites, the length and diameter of the apical dendrite, the shape and size of the soma$\cdots$ For each neuron, these features have to be computed automatically from the images. The specificity of this problem is that, for each rat cortex, we are given several images:

- one low resolution (LR) image which shows an overall view of the cortex and allows to compute the features related to the apical dendrite;
- some high resolution (HR) images (typically between 4 and 6) which provide close-ups of the somas of the neurons and allow to compute the other features.

This work consists in (1) extracting the neurons from the images (see Fig. 12), (2) matching the corresponding neurons in the HR and the LR images, (3) computing the features for each neuron, and (4) classifying the neurons, for example using a kernel Support Vector Machine (SVM).

../../../../projets/morpheme/IMG/initial_image.jpg

../../../../projets/morpheme/IMG/segmented_image

(a) (b)

Figure 12. Extraction of the neurons from the high-resolution image.

4.11. Curve and graph classification using a specific metric and kernel Support Vector Machines

Participants: Vladimir Gutov, Eric Debreuve, Xavier Descombes.

The analysis and comparison of trees are of great importance since many natural structures can be described using such models. In biology, lung airways, neurons, blood vessels$\cdots$ can be represented by trees (or, more generally, by graphs). Starting from a biological problem (automatically classifying neurons as wild or mutant), we studied the question of using Support Vector Machines (SVM) to classify continuous data such as curves, trees and graphs. Indeed, SVMs are designed for discrete data, looking for an optimal separation hyperplane or manifold in a discrete normed space. Manifolds are found when the original, linear SVM formulation is extended using the so-called kernel trick. The Gaussian kernel is the most popular one. By definition, the isotropic Gaussian kernel involves the two sample data to be compared through the distance between them. This opens the application of Gaussian-kernelized SVMs to any normed space. When dealing with (continuous) curves, the Fréchet distance can be used. We also tested a metric based on shape analysis [34]. Finally, a (meta-)distance between trees proved to be efficient in comparing axons [33] (see Fig. 13). The “meta” qualifier means that this distance builds upon a metric between curves and is valid for any such metric. It was tested using the shape-based metric [34]. We adapted an open-source SVM implementation to be able to use the three aforementioned metrics (two between curves, one between trees) and we validated the classification approach on synthetic data and on a small database of 20 wild-type neurons and 24 mutants provided by biologists.

4.12. Random forests for zooplankton classification

Participant: Eric Debreuve.

This work has been made in collaboration with Florent Baronian (Engineering student), Luc Deneire (I3S) and Marc Picheral (LOV)

../../../../projets/morpheme/IMG/axon.png

Figure 13. Left: Axonal tree imaged with a confocal microscope (maximum intensity projection of the 3-D acquisition). Right: Manual extraction of the axon.

An UVP embedded system (Underwater Vision Profiler) is a device mainly composed of a digital camera with a fixed focal distance, and a flash system designed to illuminate only the focal plane. The device is attached to a boat by a cable and it is let going deep to take pictures at various depths. The purpose of such acquisition campaigns is to analyze the population of zooplankton organisms in different oceanic regions (see Fig. 14). The fifth version of the UVP developed at the Laboratoire d'Océanographie de Villefranche (LOV) only stores the pictures. All the processing is done offline: zooplankton organisms are segmented, features are extracted and a classification into types of organisms is performed. With the upcoming sixth version, the goal is to make the device smaller, lighter, and autonomous (for some time) in order to be placed it in appliances designed to drift or navigate autonomously for weeks or months. This imposes to perform all the image processing tasks aboard, which limits the available processing power. Our work is to propose a classification method taking into account the constraints given by the teams in charge of the hardware design. We implemented a Random forest-based classifier which combines both good performances and low computational requirements. Since the images contain a lot of spurious objects called aggregates, we proposed a two-stage approach: the first stage is either a binary classifier or model checker tailored to eliminate the aggregates, while the second stage actually classifies the zooplankton organisms. We tested a combination of a one-class SVM (model checker) and a Random forest, and a combination of two Random forests, the first one being restricted to a binary classifier. Results were encouraging.

`../../../../projets/morpheme/IMG/uvp5.png`

`../../../../projets/morpheme/IMG/zooplankton.png`

Figure 14. Left: Version 5 of the UVP embedded system (Underwater Vision Profiler) used to take pictures of the zooplankton at different depths. Right: Some organisms composing the zooplankton (Lars Stemmann, LOV). Images acquired by the UVP have a much worse resolution.

4.13. Detection of Hedgehog protein using confocal imaging

Participants: Sylvain Prigent, Xavier Descombes.

This work was made in collaboration with P. Therond's group at iBV

P. Therond's lab is focusing on the understanding of how the secreted Hedgehog (Hh) morphogen, a dually lipidated highly hydrophobic molecule bound to membranes, is secreted, released and transported from the place of production in Drosophila. High resolution microscopy developed to identify and visualize such processes was successfully applied to address this question, but dynamics of membrane transport is still poorly understood mainly due to the lack of a reliable model or the need of fixation.

To statistically quantify the position of a population of Hedgehog proteins inside a cell, we need an automatic image processing method that detect each individual Hedgehog proteins in a 3D confocal image. The 3D images are obtained either by scanning Z slices (see figure 15), or by scanning XY slices (see figure 16). For both types of images, we used Marked Point Process (MPP) to detect 2D objects independently on each frame since proteins appear only in one slice. Using the Z slices, we observed that proteins appear as vertical rectangles whose size is close to the image resolution. We then design an MPP model to search for a configuration of vertically oriented rectangles that do not overlap one to each other. To define a data term for a given rectangle in the image, we calculate the Bhattacharyya distance between the population of pixels inside the rectangle and the population of pixels in the border of the rectangle. For the XY slices, we defined an MPP model to search for circles that do not overlap one to each other, and the data term is calculated with the Bhattacharyya distance as for the rectangle model. Examples of obtained results are shown in figures 15 and 16 .

`../../../../projets/morpheme/IMG/z_slice.png`

`../../../../projets/morpheme/IMG/z_slice_result`

(a)

(b)

Figure 15. Example of proteins detection on a single z image slice (a) Original z slice. (b) Obtained proteins detection.

../../../../projets/morpheme/IMG/xy_slice.png

../../../../projets/morpheme/IMG/xy_slice_resul

(a) (b)

Figure 16. Example of proteins detection on a single xy image slice (a) Original xy slice. (b) Obtained proteins detection.

SERPICO Project-Team

6. New Results

6.1. Lifetime estimation in photon counting-based fluorescence lifetime imaging microscopy

Participants: Philippe Roudot, Charles Kervrann.

In this study, we investigated a Maximum Likelihood (ML) framework for photon counting-based fluorescence lifetime estimation in Fluorescence Lifetime Imaging Microscopy (FLIM). Data collected at a given pixel consist of photon counts exponentially decreasing along the time and are assumed to follow Poisson statistics (see Fig. 6). A careful analysis of the biophysical phenomenon and instrument models are used to derive a proper ML framework for lifetime estimation. Unlike usual pointwise approaches, a neighborhood-wise approach is proposed to take explicitly into account the spatial correlation of data [15]. The application to real biological data allowed us to prove the spatial localisation of interactions, a new result which was not achievable with conventional methods. For future work, the main challenge is to extend the framework to deal with multi-exponential decay estimate and adaptive neighborhoods, a challenge we need to address for a large class of biological studies.
Reference: [15]
Partners: A. Chessel (University of Cambridge, UK), F. Waharte and J. Boulanger (UMR 144, PICT IBiSA, CNRS-Institut Curie)

6.2. Vesicle segmentation method with automatic scale selection in TIRF microscopy

Participants: Antoine Basset, Charles Kervrann, Patrick Bouthemy.

Accurately detecting cellular structures in fluorescence microscopy is of primary interest for further quantitative analysis such as counting, tracking or classification. We aimed at segmenting vesicles in Total Internal Reflection Fluorescence (TIRF) microscopy images.

In this study, we have proposed an original and efficient method – called SLT-LoG – for vesicle segmentation with fewer parameters than the state-of-the-art methods. It exploits the Laplacian of Gaussian (LoG) of the images at several scales. Since the vesicles size is almost constant in space and time, a prominent mode is expected in the empirical distribution of the scales at which the minima of LoG values are detected. It precisely corresponds to the optimal sought scale. The vesicle segmentation map is then derived by thresholding the LoG values obtained at this optimal scale. To set the threshold, we assume that the values of the LoG locally follow a normal distribution (see Fig. 7). For each point, we estimate the local mean and variance, and the threshold is deduced from a user-selected probability of false alarm.

We have evaluated our method on classical synthetic sequences for which the performances of many detection methods are available [52], [56]. The comparative results on the dataset demonstrated that our method outperforms well-known unsupervised methods. We have also obtained very satisfactory results on real complex TIRF sequences.
Partners: Jean Salamero, J. Boulanger (UMR 144, PICT IBiSA, CNRS-Institut Curie)

6.3. Conditional random fields for vesicle traffic analysis with background estimation

Participants: Thierry Pécot, Patrick Bouthemy, Charles Kervrann.

../../../../projets/serpico/IMG/FLIM-TCSPC.png

Figure 6. Example of typical Time-Correlated Single Photon Counting (TCSPC) FLIM data. Total fluorescence intensity is shown in the center and corresponds to the sum of fluorescence intensities along the time axis at each pixel. The four side graphs correspond to time dependent photon counts in four different regions with variable sizes. By considering large regions, we observe an exponential decreasing along the time of fluorescence lifetime (see D). A: one pixel region; B and C: 3×3 *patches at different locations; D:* 15×15 *patch and liftetime estimation by least-square fitting.*

../../../../projets/serpico/IMG/segment.png

Figure 7. Segmentation method applied to a real TIRF microscopy sequence showing the Rab11-mCherry protein. The SLT-LoG method is able to provide the entire spatial support of the vesicles.

../../../../projets/serpico/IMG/CRFS/gr-1a.png ../../../../projets/serpico/IMG/CRFS/gr-1b.png ../../../../projets/serpico/IMG/CRFS

Figure 8. Left: Real fluorescence microscopy image depicting GFP-Rab6 proteins. Center: estimated vesicular component. Right: estimated background component.

Image analysis applied to fluorescence live cell microscopy has become a key tool in molecular biology since it enables to characterize biological processes in space and time at the subcellular level. In fluorescence microscopy imaging, the moving tagged structures of interest, such as vesicles, appear as bright spots over a static or non-static background. In this work, we consider the problem of vesicle segmentation and time-varying background estimation at the cellular scale. The main idea is to formulate the joint segmentation-estimation problem in the general Conditional Random Field (CRF) framework. Furthermore, segmentation of vesicles and background estimation are alternatively performed by energy minimization using a min cut-max flow algorithm. The proposed approach relies on a detection measure computed from intensity contrasts between neighboring patches in fluorescence microscopy images. We have demonstrated the competitiveness of the proposed method through an experimental comparison with state-of-the-art methods in fluorescence videomicroscopy, for single cell studies. We have also characterized the density of Rab6 transport carriers spatially dispersed at the cell periphery, for two different specific adhesion geometries.

Partners: Jean Salamero, J. Boulanger (UMR 144, PICT IBiSA, CNRS-Institut Curie)

6.4. Exemplar-based occlusion handling and sparse continuous aggregation for optical flow computation

Participants: Denis Fortun, Patrick Bouthemy, Charles Kervrann.

Handling large displacements, motion details and occlusions all together remains an open issue for reliable computation of optical flow in a video sequence. Our recently investigated aggregation paradigm is an attractive approach supplying motion candidates at every pixel in a first step, and combining them in a second step to determine the global optical flow field [16]. We experimentally demonstrate that simple and purely local parametric estimations combined with patch correspondences are sufficient to produce highly accurate motion candidates. Nevertheless, the performances are limited by the presence of large occlusion areas. Therefore we have proposed an exemplar-based occlusion handling scheme integrated in the two steps of the aggregation process. At the first stage, local motion candidates sets are extended at the detected occluded pixels with candidates from non-occluded pixels, and specific occlusions due to camera motion are handled by estimating the dominant motion in the image. Local occlusion cues are extracted from this first step. Then, we define a global energy function which cooperatively selects the best motion candidates for each point while recovering the occlusion areas and ensuring smoothness properties. Results on small displacement sequences are competitive with state-of-the-art methods, and great improvements are observed in the case of large displacements and occlusions (Fig. 9).

Alternatively to the discrete aggregation based on graph cut optimization, a new continuous aggregation model has been designed. In accordance with the demonstrated evidence that the set of candidates always contains at least one accurate motion vector, the aggregation is formulated in a sparse framework restricting the number of non negligible weights associated to the candidates. The continuous framework is less dependent on the quality of the candidates and thus allows us to considerably reduce the computational cost of both aggregation and candidates estimation.
Reference: [16]

6.5. Correlation and variational approaches for motion and diffusion estimation

Participants: Denis Fortun, Charles Kervrann.

Diffusion coefficient estimation in live cell fluorescence imaging is usually achieved with correlation-based methods related to Image Correlation Spectroscopy (ICS) [42]. This approach requires a high computational cost and the spatial resolution of the resulting diffusion map is limited by the inherent block-based principle of the method. To overcome these drawbacks, we propose a novel diffusion estimation method in a variational framework providing dense and discontinuity-preserving diffusion fields. The diffusion equation is integrated in a global energy via a neighborhood-wise data term, positivity constraint and temporal integration. The

../../../../projets/serpico/IMG/cave_2_26.png ../../../../projets/serpico/IMG/cave_2_27.png [illegible]

Image I_1

Image I_2

Optical flow ground-truth

[illegible]

Our method

Weinzaepfel et al. [63]

Xu et al. [64]

../../../../projets/serpico/IMG/cave_2_26_occ.png [illegible]

Occlusion ground truth

Occlusions recovered by our method

Figure 9. Comparative evaluation of optical flow estimation for large displacements between our method, [63] and [64]. First row : two successive frames I_1 and I_2 and the ground truth motion field; second row: comparative estimation results; third row: evaluation of our occlusion map estimation.

performances of the variational and ICS approaches were compared on simulated sequences. We have demonstrated the accuracy of ICS in stationarity conditions, and we pointed out the advantages of dense variational estimation to accurately recover spatial and temporal discontinuities (Fig. 10).
Reference: [17]
Partners: Perrine Paul-Gilloteaux, Francois Waharte and Chen Chen (UMR 144, PICT IBiSA, CNRS-Institut Curie)

6.6. Classification of membrane dynamics in TIRF microscopy

Participants: Antoine Basset, Charles Kervrann, Patrick Bouthemy.

Recognizing dynamic protein behaviors in live cell fluorescence microscopy is of paramount importance to understand cell mechanisms. In the case of membrane traffic, cargo molecules are transferred from a donor to an acceptor compartments [49]. At each step, dedicated molecular platforms are acting to form, transport and address selected proteins. In microscopy imaging, this sequence of processes leads to a series of heterogeneous dynamics, which need to be untangled in order to understand the spatiotemporal coordination of the molecular actors. In this study, we aim at locating and recognizing temporal events in TIRF microscopy image sequences related to membrane dynamics. After segmenting the time-varying vesicles in the image, we exploit space-time information extracted from three successive images only to model, locate and recognize the two dynamic configurations of interest: translational motion or local fluorescence diffusion (see Fig. 11). A likelihood ratio test is defined to solve this issue. Results on synthetic sequences and real TIRF sequences demonstrated the accuracy and efficiency of the proposed method.
Partners: Jean Salamero, J. Boulanger (UMR 144, PICT IBiSA, CNRS-Institut Curie)

6.7. Crowd motion classification

Participants: Antoine Basset, Charles Kervrann, Patrick Bouthemy.

Important research efforts have been devoted to crowd analysis for several years [58], [65]. We are interested in this topic for two main reasons. First, views of crowded scenes are not that different of light microscopy intracellular images. Second, the addressed problem, i.e. motion understanding, is common, and we are investigating similar data-driven methodological approaches. This a way to cross-fertilize two domains.

We address the problem of classifying coherent crowd motions in videos recorded by a fixed camera. In contrast to most existing methods, which are based on trajectories or tracklets, our approach for crowd motion analysis provides a crowd motion classification on a frame-by-frame basis. Indeed, we only compute affine motion models from pairs of two consecutive video images. The classification itself relies on simple rules on the coefficients of the computed affine motion models, and therefore does not imply any prior learning stage. The overall method proceeds in three steps: we first compute a set of motion model candidates on a collection of windows of different sizes in the image, then we select the motion model at each point owing to a ML criterion, finally we determine the crowd motion class map with a hierarchical classification tree regularized by majority votes. The algorithm is almost parameter-free, and is extremely efficient in terms of memory and computation load. Experiments on computer-generated sequences [28] and real video sequences demonstrate that our method is accurate, and can successfully handle complex situations (see Fig. 12).
References:[14], [23]

6.8. Estimation of the flow of particles without tracking algorithm in fluorescence imaging

Participants: Thierry Pécot, Patrick Bouthemy, Charles Kervrann.

`../../../../projets/serpico/IMG/diff.png`

`../../../../projets/serpico/IMG/mask_diff`

(a) Frame 1

(b) Ground-truth diffusion field

`../../../../projets/serpico/IMG/profile2.png`

`../../../../projets/serpico/IMG/histogram`

(c) Estimated diffusion field D

(d) Histogram of D

`../../../../projets/serpico/IMG/profile.png`

`../../../../projets/serpico/IMG/plot9v2.p`

(e) Segmented diffusion field

(f) Profiles

Figure 10. Variational diffusion estimation on a simulated sequence with spatially variant diffusion. The curves of (f) are profiles of the dashed lines in (b),(c) and (e)

../../../../projets/serpico/IMG/VesicleMotionTypes.png

Figure 11. Classification results for a real TIRF sequence, whose estimated PSNR is 28.6. Results are displayed for a representative frame. The only classification error – framed in yellow – is a diffusion classified as translation. However, this vesicle has a very low intensity and changes its shape while diffusing. Two vesicles framed in green are detected as a single connected component. The vesicle framed in red corresponds to diffusing vesicles

../../../../projets/serpico/IMG/MarathonMotion.png

Figure 12. Two frames of the Marathon bend sequence. People run from upper left to upper right, describing a U. The movement is quite constant in the whole sequence and so is the classification: in the left branch, people go South (magenta), then turn counterclockwise (red) until the end of the bend. Some Eastward translation (yellow) is sometimes found here because of the large radius of curvature. Finally the North translation is recovered (blue). The points in the upper right corner of the image are classified as translations to the West (purple), but the translation direction is closer to North than to West (North-North-West): it is also due to the lateral presence of pedestrians walking to the left.

../../../../projets/serpico/IMG/EFP.png

Figure 13. Vesicle flows estimated with our method when considering a simple partition of 5 regions for an image sequence acquired in TIRF microscopy and showing the protein Clip170.

Automatic analysis of the dynamic content in fluorescence video-microscopy is crucial for understanding molecular mechanisms involved in cell functions. We have proposed an original approach for analyzing particle trafficking in these sequences. Instead of individually tracking every particle, we only need to locally count particles on regions over time and minimize a global energy function. We have specified three methods to determine the particle flow. We especially compared the NNLS algorithm [44] and the PPXA algorithm [33] known as well suited to non differentiable convex minimization problem [24]. We have conducted comparative experiments on synthetic and real fluorescence image sequences. We have shown that adding a sparsity constraint on the number of detected events allows us to reduce the number of false alarms. Compared to usual tracking methods, our approach is simpler and the results are very stable with respect to the only two parameters involved (see Fig. 13).
Reference: [24]

Partners: Jean Salamero, J. Boulanger (UMR 144, PICT IBiSA, CNRS-Institut Curie)

6.9. Probabilistic Tracking of fluorescent objects

Participants: Philippe Roudot, Charles Kervrann.

Image tracking of fluorescent objects, from labeled molecules to organelles and entire cells, is an essential task in the analysis of cellular functions. During the last decade, several algorithms have been tailored to cope with different types of cellular and subcellular motion down to Brownian single molecule behavior [8]. One of the remaining big challenges in this area of technology development has been the tracking of extremely heterogeneous movements of objects in crowded scenes. We tested several state-of-the-art algorithms [36], [40] to follow dense populations of diffusing particles, which suddenly change to directed motion. A frequent cellular scenario with this property is the jerky motion of vesicles and viruses switching between cytoplasmic diffusion and motor-mediated, fast displacements (see Fig. 14).

`../../../../projets/serpico/IMG/VimentinMotion.png`

Figure 14. Vimentin motility seems to present a large proportion of confined Brownian motion and rare, sudden, motor-mediated transport. Colored tracks have been computed with an advanced U-track parametrization (Unit length filament of Vimentin Y117L mutant fused to GFP and transfected into vimentin null epithelial cell (cell line SW13). Image acquired with a spinning disk confocal microscospe with a 100x objective zoom 1.5 (Numerical Aperture 1.4, pixel size 0.10905μm/pixel).

../../../../projets/serpico/IMG/TrackPerf.png

Figure 15. A) Example of tracks simulation presenting a density of 3 spots/µm^2. B) Correct linking percentage wrt density and motion type switching probability. Our method outperforms U-track by 15% in the hardest case. C) True positive and false positive ratio on the same simulation with a density of 3 spots/µm^2, comparing our method with U-track, U-track with an on-line process noise estimator and an IMM algorithm with forward-backward initialization.

../../../../projets/serpico/IMG/SpeedHet.png

Figure 16. Correct linking and false positive percentage wrt speed switching probability.

These switches are particularly challenging to detect because they occur rarely. The presence of numerous detected objects in the expected range of particle displacements makes the tracking ambiguous and induces wrong associations. Lowering the ambiguity by reducing the search range, on the other hand, is not an option, as this would increase the rate of false negatives.

We first explored the existing methods in the literature to analyze their strenghts and weakness for tracking objects with heterogeneous motion and high density. Based on the conclusion we draw, we proposed a new method build on the U-track platform [40]. More specifically, we propose an interacting multiple state model that exploits recursive tracking in multiple rounds in forward and backward temporal directions. As a result, it achieves convergence of the instantaneous speed estimate time-point-by-time-point. This allows us to predict and recover abrupt transitions from freely or confined diffusive to directed motion. To address the issue of a particle that disappears as a neighboring particle appears in the same image and thus to better detect track termination, we also exploit this recursive tracking by proposing a locally adaptive on-line estimation of the search window radius for assignment (a.k.a. gating), while most of state-of-the-art algorithms propose only a global search window radius or weak per-track search radius estimations. We have shown on simulated data that our method outperforms state-of-the-art algorithms that model motion heterogeneity on different scenarios, e.g. heterogeneous motion type (see Figure 15) and speed heterogeneity (see Fig. 16), while keeping the computational cost of a deterministic method (10% overhead with respect to U-track).

Partners: Gaudenz Danuser (Harvard Medical School, Boston, USA)

6.10. Microtubules modeling for variational assimilation analysis

Participants: Pierre Allain, Charles Kervrann.

Microtubules (MT) are highly dynamic tubulin polymers that are involved in many cellular processes such as mitosis, intracellular cell organization and vesicular transport. Nevertheless, the modeling of cytoskeleton and MT dynamics based on physical properties is difficult to achieve. We proposed to model microtubules as rigid and growing cylinders alike (Nedelec and Foethk 2007) [45] but including Newtonian dynamics. Using the Euler-Bernoulli beam theory, we have proposed then to model the rigidity of microtubules on a physical basis using forces, mass and acceleration. In addition, we linked microtubules growth and shrinkage to the presence of molecules (e.g. GTP-tubulin) in the cytosol. The overall model enables linking cytosol to microtubules dynamics in a constant state space, thus allowing usage of data assimilation techniques (see Fig. 17).

../../../../projets/serpico/IMG/MT-simu2D.png ../../../../projets/serpico/IMG/MT-simu3D.png

Figure 17. Left: Simulation of a 2D radial microtubule network. The results show growing and shrinking phases yielding inhomogeneous "pseudo-tubulin" concentration in the cytosol. MTs are bended according to fluid forces. Right: 3D simulation of MT nucleation and growth that mimics MT dynamics seeded onto a two vertical bar-shaped fibronectin pattern and observed in TIRF microscopy (courtesy of iRTSV/LPCV/PCM CEA-Grenoble).

6.11. Spot localization for TMA image analysis

Participants: Nam-Hoai Nguyen, Charles Kervrann.

A very first task of TMA (Tissue MicroArray) image analysis is to accurately localize spots (separate tissue core) representing arrays of 512 x 512 pixels each, in very large images of several thousands of pixels. For this purpose, we have investigated a three-stage methodological approach. First, since tissue cores are separately assembled in array (grid structure). We started to design a graphical model to eliminate image defects due to the presence of dusts or the imperfection of TMA blocks fabrication. In the second stage, a wavelet-like transform is currently used to recognize interested features (spots) given the size of spots *a prori*. Third, we started to investigate the superpixel-based image representation (SLIC) [27], [55] to handle very large images and biological details inside each spot.

Partners: V. Paveau (Innopys company)

VIRTUAL PLANTS Project-Team

5. New Results

5.1. Analysis of structures resulting from meristem activity

5.1.1. Acquisition and design of plant geometry

Participants: Chakkrit Preuksakarn [Univ Kasertsart, Thaïland], Mathilde Balduzzi, Frédéric Boudon, Christophe Pradal, Christophe Godin, Christian Fournier.

Virtual 3D model of plants are required in many areas of plant modeling. They can be used for instance to simulate physical interaction of real plant structures with their environment (ligth, rain, wind, pests, ...), to set up initial conditions of growth models or to assess their output against real data. In the past decade, methods have been developed to digitize plant architectures in 3D [59], [53]. These methods are based on direct measurements of position and shape of every plant organ in space. Although they provide accurate results, they are particularly time consuming. More rapid and automated methods are now required in order to collect plant architecture data of various types and sizes in a systematic way. In this aim, we explore the use of pictures, laser scanner, video and direct sketching.

- *Reconstruction of plant architecture from 3D laser scanner data.* (Chakkrit Preuksakarn, Mathilde Balduzzi, Frédéric Boudon, Christophe Godin, Pascal Ferraro [Labri, Bordeaux])

  We investigate the possibility to use 3D laser scanners to automate plant digitizing. We are developping algorithms to reconstruct branching systems without leaves or foliage from scanner data or from scan simulated on plant mock-up obtained using different digitizing method.

  For branching system, a number of automatic methods were proposed in the literature to reconstruct plausible branching structures from laser scanner data. The question of their assessment and accuracy is however critical for further exploitation in biological applications. For this, we developed an evaluation pipeline to assess reconstructions accuracy. A laser scan database on which experts built reference reconstructions is used as a basis of the evaluation. The pipeline is given two structures and compares both their elements and their organization. Similar elements are identified based on geometric criteria using an optimization algorithm. The organization of these elements is then compared and their similarity quantified. Two indices of geometrical and structural similarities are defined, and automatic reconstructions can thus be compared to reference structures to assess their accuracy. The method is successful at capturing the variation of similarities between two structures as different levels of noise are introduced. A first comparative evaluation of the different methods of the literature has been designed and conducted. This work has been presented at the FSPM conference and submitted at Annals of Botany for its special issue.

  We also investigated the reconstruction of tree foliage from 3D scans. Such elements are crucial to study the interaction of the plant with its environment. However, laser scans contain outliers on the silhouette of the scans that make the meshing of the pointset difficult. New generation of laser scanners provide intensity of the laser reflected on the surface of scanned objects. This intensity depends of the distance to the object, its optical property and the incidence angle. A first work on this topic showed that after correcting the distance effect, the incidence angle can be deduced from the intensity. From this result, we developed a reconstruction pipeline using the scan intensities and based on Shape-From-Shading approaches. This work has also been presented at the FSPM conference.

- *Reconstruction from video.* (Frédéric Boudon, Jerome Guenard [IRIT, Toulouse], Geraldine Morin [IRIT, Toulouse], Pierre Gurdjos [IRIT, Toulouse], Vincent Charvillat [IRIT, Toulouse])

`../../../../projets/virtual_plants/IMG/cherry_tree.png`

Figure 2. Reconstruction of a cherry tree. Left: photograph of the original tree. Right: 3D reconstruction from a laser scan rendered and integrated on the same background.

Even if mature computer vision techniques allow the reconstruction of challenging 3D objects from images, due to high complexity of plant topology, dedicated methods for generating 3D plant models must be devised. In collaboration with our colleagues from IRIT, Toulouse, we developed an analysis-by-synthesis method which generates 3D models of a plant from both images and a priori knowledge of the plant species.

Our method is based on a skeletonisation algorithm which allows to generate a possible skeleton from a foliage segmentation. Then, a 3D generative model, based on a parametric model of branching systems that takes into account botanical knowledge is built. This method extends previous works by constraining the resulting skeleton to follow hierarchical organization of natural branching structure. 3D models are then generated. A reprojection of the models can be compared with the original image to assess the visual accuracy of the reconstruction. We optimise the parameter values of the generative model based on the reprojection criterion. Realistic results are obtained on different species of plants, in particular vineyards. This work has been presented at the ISVC conference and published in LNCS [46].

- *Reconstruction of virtual fruits from pictures.* (Mik Cieslak, Nadia Bertin [Inra, Avignon], Frédéric Boudon, Christophe Godin, Michel Genard [Inra, Avignon], Christophe Goz-Bac [Université Montpellier 2])

  This research theme is supported by the Agropolis project Fruit3D.

  The aim of this work is to provide methods for generating fruit structure that can be integrated with models of fruit function and used to investigate such effects. To this end, we have developed a modeling pipeline in the OpenAlea platform that involves two steps: (1) generating a 3D volumetric mesh representation of the entire fruit, and (2) generating a complex network of vasculature that is embedded within this mesh. To create the 3D volumetric mesh, we use reconstruction algorithms from the 3D mesh generation package of the Computational Geometry Algorithms Library (CGAL). To generate the pattern of vasculature within this volumetric mesh, we use a Space Colonisation Algorithm that populates the volume of the fruit by simulating competition for space of the vasculature. We have applied our modeling pipeline to generate the internal and external geometry of a cherry tomato fruit using Magnetic Resonance Imaging data as input. These studies demonstrate the possibility to create species-specific models of fruit structure with relatively low effort [31]. These volumetric meshes are then combined with models of function to form integrative computational fruit models, which will help to investigate the effects of fruit structure on quality (see section 5.3.2).

- *Reconstruction of root structures.* (Julien Diener, Fredéric Boudon, Christophe Pradal, Christophe Godin, Philippe Nacry [BPMP, INRA], Christophe Périn [AGAP, CIRAD], Anne Dievart [AGAP, CIRAD], Xavier Draye [UCL, Belgium])

  Similarly to aerial part of plants, some needs for reconstruction procedure of root systems emerges. Most existing methods focus only on semi-automatic approaches. This does not support the high-throughput capabilities of acquisition systems. Within the Rhizopolis project (Agropolis foundation), we have designed an automatic analysis pipeline to extract root system architecture from images. This pipeline provides i) a model based segmentation method of the scanned image content (Petri plate, seeds, leaves and root pixels), ii) the extraction of a graph representation of the root system, and iii) a novel method to identifying the root axes organization.

5.1.2. Modeling the plant ontogenic programme

Participants: Christophe Godin, Yann Guédon, Evelyne Costes, Jean-Baptiste Durand, Anaëlle Ambreville, Pierre Fernique, Christophe Pradal, Jean Peyhardi, Catherine Trottier, Yassin Refahi, Etienne Farcot.

This research theme is supported by two PhD programmes.

The remarkable organization of plants at macroscopic scales may be used to infer particular aspects of meristem functioning. The fact that plants are made up of the repetition of many similar components at different scales, and the presence of morphological gradients, e.g. [48], [54], [55], [52], provides macroscopic evidence for the existence of regularities and identities in processes that drive meristem activity at microscopic scales. Different concepts have been proposed to explain these specific organizations such as "morphogenetic programme" [57], "age state" [51] or "physiological age" [50]. All these concepts state that meristem fate changes according to position within the plant structure and during its development. Even though these changes in meristem fate are specific to each species and lead to the differentiation of axes, general rules can be highlighted [51], [50]. Here we develop computational methods to decipher these rules.

- *Relating branching structure to the shoot properties* (Jean Peyhardi, Yann Guédon, Evelyne Coste, Catherine Trottier, Yves Caraglio [AMAP], Pierre-Eric Lauri [AGAP, AFEF team])

  Shoot branching structures often take the form of a succession of homogeneous branching zones and have been analyzed using segmentation models such as hidden semi-Markov chains. Axillary meristem fates are influenced by local properties of the parent shoot such as for instance its growth rate or local curvature. The objective of this work, which is part of the PhD subject of Jean Peyhardi, is to develop statistical models that generalize hidden semi-Markov chains with the capability to incorporate explanatory variables that vary along the parent shoot (e.g. maximum growth rate of the leaf, surface of the leaf, length of the internode, local curvature of the parent shoot). More precisely, the simple multinomial distributions that represent the axillary productions observed in the different branching zones are replaced by multinomial generalized linear models (GLMs). Since the two classical categories of multinomial GLMs that correspond either to nominal or ordinal categorical response variables were not appropriate, we chose to develop a new family of multinomial GLMs called multi-step multinomial GLMs that enable to tackle partially ordered categorical response variables. Typically, we need to distinguish different timing of branching (e.g. immediate shoot, one-year-delayed shoot and latent bud), different categories of offspring shoots (e.g. among one-year-delayed shoots, vegetative short shoot, vegetative long shoot and flowering shoot) and to specialize the explanatory variables for certain categories of offspring shoots (e.g. the growth of the parent shoot influence the immediate offspring shoots but not the one-year-delayed offspring shoots). The resulting integrative models are called semi-Markov switching multi-step multinomial GLMs and are applied to different data sets corresponding mainly to fruit tree branching structures.

- *Genetic determinisms of the alternation of flowering in apple tree progenies.* (Jean-Baptiste Durand, Jean Peyhardi, Baptiste Guitton [AGAP, AFEF team], Yan Holtz [AGAP, AFEF team] Catherine Trottier, Evelyne Costes, Yann Guédon)

  A first study was performed to characterize genetic determinisms of the alternation of flowering in apple tree progenies [34], [17]. Data were collected at two scales: at whole tree scale (with annual time step) and a local scale (annual shoot or AS, which is the portions of stem that were grown during the same year). Two replications of each genotype were available.

  Indices were proposed for early detection of alternation during the juvenile phase. They were based on a trend model and a quantification of the deviation amplitudes and dependency, with respect to the trend. This allows to quantify alternation from the yearly numbers of inflorescences at tree scale.

  However, phenotyping subsamples of AS sequences is more realistic in the framework of genotype selection. To model alternation of flowering at AS scale, a second-order Markov tree model was built. Its transition probabilities were modeled as generalized linear mixed models, to incorporate the effects of genotypes, year and memory of flowering for the Markovian part, with interactions between these components. Asynchronism of flowering at AS scale was also assessed using an entropy-based criterion.

  This work started during the PhD's work of Baptiste Guitton. It was then extended in 2012 by Yan Holtz during this Master 2 internship, supervised by Evelyne Costes and Jean-Baptiste Durand. New progenies were considered, as well as the performance of approximating the descriptors at whole tree scale with those at AS scale. These descriptors allowed the identification of QTL zones involved in the control of flowering in apple trees.

  As a perspective of this work, patterns in the production of children ASs (numbers of flowering and vegetative children) depending on the type of the parent AS must be analyzed using branching processes and different types of Markov trees, in the context of Pierre Fernique's PhD Thesis (see next item in Section 5.1.2).

- *Modeling branching patterns in fruit tree shoots through the characterization of their demographic properties* (Pierre Fernique, Jean-Baptiste Durand, Yann Guédon).

  To test the effect of some properties of a given parent shoot on the properties of its children shoots, statistical models based on multitype branching processes were developed. This kind of dependence between parent and children shoots is frequently at stake in fruit trees, for which the number of flowering or vegetative children of a parent shoot depends on its nature, with potential interactions with other factors. Thus, controlling demographic patterns of the shoots (through varietal selection or crop management strategies) is expected to bring substantial improvements in the quantity and quality of yields.

  Formally, the shoot properties are summed up using the notion of shoot state. The number of children shoots in each state is modeled through discrete multivariate distributions. Model selection procedures are necessary to specify parsimonious distributions. We developed an approach based on probabilistic graphical models to identify and exploit properties of conditional independence between numbers of children in different states, so as to simplify the specification of their joint distribution. The graph building stage was based on exploring the space of possible chain graph models, which required defining a notion of neighbourhood of these graphs. A parametric distribution was associated with each graph. It was obtained by combining families of univariate and multivariate distributions or regression models, and their mixtures. These families were chosen by selection model procedures among different parametric families [45], [32].

This work was carried out in the context of Pierre Fernique's first year of PhD (Montpellier 2 University and CIRAD). It was applied to model dependencies between short or long, vegetative or flowering shoots in apple trees. The results highlighted contrasted patterns related to the parent shoot state, with interpretation in terms of alternation of flowering (see previous item in Section 5.1.2). It was also applied to the analysis of the connections between cyclic growth and flowering of mango trees (see next item in Section 5.1.2). This work will be continued during Pierre Fernique's PhD thesis, with extensions to other fruit tree species and other parametric discrete multivariate families of distributions, including covariates and mixed effects.

- *Analyzing fruit tree phenology* (Anaëlle Dambreville, Jean-Baptiste Durand, Pierre Fernique, Yann Guédon, Christophe Pradal, Pierre-Eric Lauri [AFEF team, AGAP], Frédéric Normand, Catherine Trottier)

  Mango is a tropical tree characterized by strong asynchronisms within and between trees. Causation networks explaining the vegetative and reproductive growths within and between growing cycles were studied on the basis of generalized linear models [16]. We highlighted in this way marked interplays between structural and temporal components of tree structure development at three scales. At growth unit scale, a growth unit which appeared early in the growing cycle had higher rate of burst compared to late appeared growth units. At growing cycle scale, a flowering growth unit delayed its future vegetative growth compared to a vegetative growth unit. At tree scale, a fruiting tree delayed further vegetative growth and flowering compared to a non-fruiting tree. These results evidenced that tree phenology is strongly affected by structural components and not only by the environment.

  To study more precisely the interplays between all these structural components, we started during the internship of A. Jestin to build an integrative model to simulate the development of mango tree based on the L-system formalism and GLM to model the dependencies between events.

  We are also investigating jointly structure development and phenology of mango, and characterizing the specific spatio-temporal patterns leading to patches of vegetative or flowering growth units. Our approach is based on statistical models for trees; particularly hidden Markov tree models and multitype branching processes [32].

- *Integrative developmental growth stages of shoots* (Anaëlle Dambreville, Yann Guédon, Pierre-Eric Lauri [AFEF team, AGAP], Frédéric Normand)

  Growth and development are often studied as two separated processes. Our aim is to investigate the coordination between growth and development in mango shoots. We considered three types of organ, namely the shoot axis, its attached leaves and the inflorescence. Two types of data were collected during the shoot and inflorescence follow-up: developmental stages determined in an expert way and organs sizes determined from measurements. To give an integrative view of the shoot and inflorescence growth and development, we adopted the following strategy. For a given cultivar, we first built a multi-state model on the basis of absolute growth rate sequences deduced from the measurements. Using these models, we computed growth stages. These growth stages highlighted growth asynchronisms between two topologically-connected organs: the axis and its leaves. Then, we compared these growth stages with the developmental ones and we obtained strong matches between them. The integrated developmental growth stages emphasized that the developmental stages are markedly related to growth rates an can be interpreted in terms of physiological (hydraulics, carbohydrates partitioning) and developmental (organs preformation versus neoformation) processes.

- *Self-nested structure of plants.* (Christophe Godin, Romain Azais, Farah Ben Naoum, Jean-Baptiste Durand, Alain Jean-Marie)

  In a previous work [5], we designed a method to compress tree structures and to quantify their degree of self-nestedness. This method is based on the detection of isomorphic subtrees in a given tree and on the construction of a DAG, equivalent to the original tree, where a given subtree class is represented only once (compression is based on the suppression of structural redundancies in the original tree). In the compressed graph, every node representing a particular subtree in the original tree has exactly the same height as its corresponding node in the original tree. This method thus compresses a tree in width, but not in height. In this new work, we designed an extension of this compression method in which a tree is compressed in both width and height. The method is based on the detection of so-called *quasi-isomorphic paths* in a tree and on the compression of these paths in height. A paper describing the corresponding algorithms is being written.

We are currently studying how to generalize the previous approach using stochastic trees. The idea is to generalize the previously defined equivalence relation so that equivalent trees are identical in distribution now and are no longer strictly isomorphic. Algorithms to estimate the different distributions from tree samples rely on combinatorics arguments that are used to estimate the probability that an observed tree is produced by a particular stochastic tree model for which a conjecture has been proposed (still to be demonstrated). The asymptotic analysis of similar compression rates on ordered tree-graphs has been carried out by Flajolet et al. (1990) for different types of distributions (uniform, multi-type branching processes). The work is developed in the context of the Post-doc of Romain Azais.

5.1.3. Analyzing the influence of the environment on the plant ontogenic programme

Participants: Frédéric Boudon, Jean-Baptiste Durand, Christophe Godin, Yann Guédon, Jean Peyhardi, Pierre Fernique, Maryline Lièvre, Christine Granier, Evelyne Costes, Pascal Ferraro, Catherine Trottier.

This research theme is supported by three PhD programs.

The ontogenetic programme of a plant is actually sensitive to environmental changes. If, in particular cases, we can make the assumption that the environment is a fixed control variable (see section 5.1.2), in general the structure produced by meristem results from a tight interaction between the plant and its environment, throughout its lifetime. Based on observations, we thus aim to trace back to the different components of the growth (ontogenetic development and its modulation by the environment). This is made using two types of approaches. On the one hand, we develop a statistical approach in which stochastic models are augmented with additional time-varying explanatory variables that represent the environment variations. The design of estimation procedures for these models make it possible to separate the plant ontogenetic programme from its modulation by the environment. On the other hand, we build reactive models that make it possible to simulate in a mechanistic way the interaction between the plant development and its environment.

- *Influence of environment conditions and horticultural practices on the branching and axillary flowering structures of fruit tree shoots.* (Yann Guédon, Evelyne Costes [AFEF Team, AGAP], Ted DeJong [UC Davis], Claudia Negron [UC Davis]).

  In the context of a collaboration with Claudia Negron and Ted DeJong, we studied the influence of water availability and pruning practices on the branching and axillary flowering structures of different categories of almond shoots Stochastic models (hidden semi-Markov chains) were built for the branching and axillary flowering structures of different categories of almond shoots corresponding to different genetic backgrounds, levels of irrigation and pruning practices.

- *Analyzing growth components in trees.* (Yann Guédon, Yves Caraglio [AMAP], Olivier Taugourdeau [AMAP])

  In a forest ecology context, we identified robust indicators that summarize the balance between tree ontogeny and environmental constraints (mainly related to light environment). In this context, tree growth data typically correspond to the retrospective measurement of annual shoot characteristics (e.g. length, number of branches) along the main stem. We applied segmentation models (hidden Markov and semi-Markov chains) that enable to identify tree growth phases. This statistical modeling approach was applied to both deciduous (sessile oak and Persian walnut) and evergreen (Corsican pine and silver fir) tree species growing in contrasted conditions ranging from managed forest stands to unmanaged understoreys. The growth phase duration distributions estimated within these segmentation models characterize the balance between tree ontogeny and the environmental constraints in tree development at the population scale. These distributions had very contrasted characteristics in terms of shape and relative dispersion between ontogeny-driven and environment-driven tree development. The characteristics of growth phase duration distributions may change over tree life reflecting changes in tree competition.

- *Investigating how architectural development interfer with epidemics and epidemic control* (Christian Fournier, Corinne Robert [EGC], Guillaume Garin [ITK], Bruno Andrieu [EGC], Christophe Pradal)

  Recent considerations towards sustainable agriculture require identifying new natural strategies of crop protection. In this perspective, a better identification of major interactions inside pathosystems between the plants, the pathogens and their environment is crucial. These multiscale biological systems are complex: multiple relationships stand out with various dynamics and at various locations in the canopy, related to its architectural development. The purpose of this research is to provide a framework to study the influence of architectural development on pathosystems with modeling. A first generic framework was designed and implemented in the platform OpenAlea [36]. It allows implementing pathogens of different kind using the same concepts, and a re-use of plant models available in Openalea, thus simplifiying the development of pathosystem models based on 3D plants models. A second action was to develop a modular integrated model coupling architectural canopy development, disease dynamics, pesticide application, pesticide decay and effect of pesticide on disease dynamics [35]. This model is currently being assessed against data for validation, and aim at designing new strategies that reduce pesticide applications by increasing natural resistance linked to canopy architecture.

5.2. Meristem functioning and development

In axis 2 work focuses on the creation of a *virtual meristem*, at cell resolution, able to integrate the recent results in developmental biology and to simulate the feedback loops between physiology and growth. The approach is subdivided into several sub-areas of research.

5.2.1. Data acquisition and design of meristem models

Participants: Frédéric Boudon, Christophe Godin, Christophe Pradal, Léo Guignard, Vincent Mirabet [RDP, ENS], Jan Traas, Grégoire Malandain, Jean-Luc Verdeil [PHIV, AGAP].

This research theme is supported by the iSam and Morphogenetics projects.

- *Improvement of the MARS-ALT pipeline robustness* Meristem, laser microscopy, image reconstruction, cell segmentation, automatic lineaging
  Participants: Léo Guignard, Christophe Godin, Christophe Pradal, Grégoire Malandain, Guillaume Baty, Jan Traas, Patrick Lemaire, Pradeep Das [RDP, ENS], Yassin Refahi [RDP, ENS].

  The MARS-ALT (Multi-Angles Registration and Segmentation - Automatic Lineage Tracking) software pipeline automatically performs a segmentation at cell resolution from 3D or 2D voxel images where the membranes/walls are marked (by a die for example) and makes it possible to follow the lineage of these cells through time [4]. A new version of this pipeline is currently being developed. MARS-ALT Version 2 is based on the same algorithms and methods and is intended to improve the overall robustness of the pipeline (protocol, noise in the input image) and automate completely the process. To test the new pipeline, we use different acquisition protocols and different organisms (floral and apical meristems and the early stages of development of a marine animal *Phallusia mammillata*). The segmentation is corrected a posteriori to deal with imaging artifacts due to uncertainties of acquisition. The image data set on which we develop the methods consists of :

  - *Arabidopsis thaliana* shoot apical meristem and primordia with around 6000 cells. The organ is captured from three different angles every 4 hours during 2 or 3 days with a confocal microscope (Collaboration Sainsbury lab, Cambridge)
  - *Arabidopsis thaliana* flower meristems with around 2000 cells. The organ is also captured from three different angles with a confocal microscope (Collaboration RDP Lyon and Sainsbury lab)

  - *Phallusia mammillata* and *Ciona intestinalis* embryos with from 32 cells to around 1000 cells. The organism is captured from four different angles every minute during 2 to 3 hours with a SPIM (Single Plane Illumination Microscope) (Collaboration CRBM Montpellier / EMBL Heidelberg). This work is developed in the context of the PhD work of Léo Guignard.

The pipeline provides as an output segmented images on which metrics for each cells can be extracted such as volume, principal components, convex hull and more. A new non-linear registration algorithm developed by G. Malandain (MORPHEME team, Inria Sophia-Antipolis) is now available and will lead to an improvement of ALT algorithm. Redesign and improvement of the lineage tracking pipeline will be the next step.

`../../../../projets/virtual_plants/IMG/Arabes_fused_seg.png`

Figure 3. Superimposition of an automatic cell segmentation of an arabidopsis flower meristem using the new MARS pipeline with the original confocal image stack where the membranes are marked.

- *Design of 3D virtual atlases for specifying gene expression patterns* (Yassin Refahi, Christophe Godin, Jan Traas, Patrick Lemaire, Grégoire Malandain, Françoise Monéger [RDP, ENS])

  This research theme is supported the ANR GeneShape and iSam projects.

  To organize the various genetic, physiological, physical, temporal and positional informations, we build a spatialized and dynamic database. This database makes it possible to store all the collected information on a virtual 3D structure representing a typical organ. Each piece of information has to be located spatially and temporally in the database. Tools to visually retrieve and manipulate the information, quantitatively through space and time are being developed. For this, the 3D structure of a typical organ has been created at the different stages of development of the flower bud. This virtual structure contains spatial and temporal information on mean cell numbers, cell size, cell lineages, possible cell polarization (transporters, microtubules), and gene expression patterns. Such 3D virtual atlas is mainly descriptive. However, like for classical databases, specific tools make it possible to explore the virtual atlas according to main index keys, in particular spatial and temporal keys. Both a dedicated language and a 3D user interface are being designed to investigate and query the 3D virtual atlas.

A prototype version of the 3D virtual atlas was built 2 years ago [6]. Further developments of this tool will rely on the segmented images produced from microscopy, as presented in the previous section. In particular, a common underlying data structure has to be developed transversally to these two scientific developments. The definition of this data structure has been initiated last year through several team meetings, and should lead to a revised implementation next year.

5.2.2. Shape analysis of meristems

(Jonathan Legrand, Frédéric Boudon, Christophe Godin, Yann Guedon, Pradeep Das [ENS Lyon])
At cellular resolution, we studied the organization of cells in the meristems. The MARS-ALT pipeline provides rich spatio-temporal data sets for analyzing the development of meristems. A first step consisted of designing a dedicated graph for efficiently representing the spatial (adjacency between cells) and temporal (cell division) relationships between cells. Various variables can be attached either to the vertices (e.g. cell volume, inertia axes) or the edges (e.g. wall surface, distance between cell centroids). This graph may be augmented by new variables resulting from various spatial or temporal filtering (e.g. cell volumetric growth). We are now designing models and algorithms for finding patterns in time courses of meristems. In particular, we are investigating spectral clustering methods in order to define homogeneous regions in terms of cell identities in the context of the PhD Work of Jonathan Legrand.

5.2.3. Transport models

Participants: Michael Walker, Christophe Godin, Etienne Farcot, Jan Traas, Yuan Yuan [University of Newfoundland, Canada].

This research theme is supported by the ANR GeneShape and ERASysBio+ iSAM projects and Morphogenetics.

Active transport of the plant hormone auxin has been shown to play a key role in the initiation of organs at the shoot apex, and vein formation in both leaves and the shoot apical meristem. Polar localized membrane proteins of the PIN1 and AUX/LAX family facilitate this transport and observations and models suggest that the coherent organization of these proteins in the L1 layer is responsible for the creation of auxin maxima (surrounded by a depletion zone), which in turn triggers organ initiation close to the meristem center [58] [1]. Furthermore, canalized PIN allocations are thought to play a crucial role in vein formation in the leaf and in the L2. Previous studies have typically modeled the L1 and L2 with different models to explain different patterns of PIN allocations. In the last two years, we developed a unifying model showing that a unique flux-based model could be sufficient to explain PIN patterns in both L1 and L2 [27]. Contrary to our previous study [9], here no change in the model parameters is needed for this. Our approach is based on inherent topological and geometrical differences between the L1 and L2, specifically their dimensionality and the distribution of sources and sinks.

In a different perspective, another study on auxin transport models have been submitted this year. In this work, a generic, adimensional flux-based model of auxin transport was studied using a combination of analytic and numeric approach. The steady-states with uniform auxin distribution were characterised for arbitrary tissues, and some of their bifurcations (loss of stability and Hopf) were described [18]. This work, initiated during an "Explorateur" project funded by Inria during the period October 2012-January 2013, results from the collaboration between E. Farcot and Y. Yuan (Memorial University of Newfoundland, Canada).

5.2.4. Mechanical model

Participants: Olivier Ali, Christophe Godin, Benjamin Gilles, Frédéric Boudon, Jan Traas, Olivier Hamant [ENS-Lyon], Arezki Boudaoud [ENS-Lyon], Jérôme Chopard [University of Western Australia, Perth].

This research theme is supported by the ANR VirtualFlower and Geneshape projects together with the Inria project Morphogenetics and the ERC from Jan Traas.

The rigid cell walls that surround plant cells are responsible for their shape. These structures are under constraint due to turgor pressure inside the cell. To study the changes of shape in plant tissues during organogenesis, we need a mechanical model of tissue development at cellular resolution. We developed such a model, in which walls are characterized by their mechanical properties like the Young modulus which describes the elasticity of the material. Wall deformation results from forces due to turgor pressure. Growth results from cell wall synthesis that is triggered when wall deformation exceeds a particular threshold. The final shape of the tissue integrates mechanically all the local deformations of each cell.

To model this process, we used a tensorial approach to describe both tissue deformation and stresses. Deformations were decomposed into elementary transformations that can be related to underlying biological processes. However, we showed that the observed deformations does not map directly local growth instructions given by genes and physiology in each cell. Instead, the growth is a two-stage process where genes are specifying how cell walls should yield to mechanical stresses. In this way, different regions in the tissue with different cell identities can have different growth properties. The final shape of the tissue results from the integration of all these mechanical properties and stresses at organ level under the growth force due to turgor pressure at tissue scale.

A paper describing the mechanical model and its application to model primorium formation in the shoot apical meristem has been submitted to PNAS in December. Additionally, a redesign of our mechanical model using the SOFA framework is being finalized.

5.2.5. Gene regulatory networks

Modeling gene activities within cells is of primary importance since cell identities correspond to stable combination of gene expression [25].

- *Complex dynamics and spatial interactions in gene networks* (Yassin Refahi, Etienne Farcot, Christophe Godin)

  Complex computational and mathematical questions arise in the study of gene networks at two levels: (i) the single cell level, due to complex, nonlinear interactions, (ii) the tissue level, where multiple cells interact through molecular signals and growth, so that even simple local rules can challenge our intuition at higher scales.
  At the single cell level, new results were obtained in the framework of piecewise-linear models, it is in general very difficult to entirely characterize the attractors of a given system. In an attempt to improve our ability on this question, a probabilistic approach has been proposed in [14], in which it is shown that a Markov chain can built as an approximation of a given piecewise-linear system, and actually used to make predictions about its periodic attractors.

  At a higher scale, we have also continued the study of gene regulation in meristematic tissues. In the context of Y. Refahi's post-doc between Virtual Plants and the group of Henrik Jönsson in Cambridge (Sainsbury Laboratory), we have continued a work that was initiated in Y. Refahi's thesis. This work is motivated by recent biological results, indicating that gradient-like patterns originating from the external layers of meristems may play a decisive role in the specification of the pool of stem cells in a central position. Using the methods in [4], and their on-going improvements, we have acquired new 3D and 4D images that were then segmented. These structures will be used in the next few months to investigate generic patterning properties of gradient like morphogen patterns. This will require a thorough analysis of free diffusion in realistic geometries, as made possible by the newly acquired images. As a preliminary work, we are also currently investigating the formation of gradient patterns in idealised tissues, allowing for deeper analytic treatment than the complex structures obtained by microscopy.

5.2.6. Model integration

Participants: Frédéric Boudon, Christophe Godin, Eugenio Azpeitia, Laurent Laplaze, Jan Traas, François Parcy.

This research theme is supported by the ANR/BBSRC project iSam.

Our approach consists of building a programmable tissue which is able to accept different modeling components. This includes a central data structure representing the tissue in either 2-D or 3-D, which is able to grow in time, models of gene activity and regulation, models of signal exchange (physical and chemical) between cells and models of cell cycle (which includes cell division). For each modeling component, one or several approaches are investigated in depth, possibly at different temporal and spatial scales, using the data available from the partners (imaging, gene networks, and expression patterns). Approaches are compared and assessed on the same data. The objective of each submodel component will be to provide plugin components, corresponding to simplified versions of their models if necessary, that can be injected in the programmable tissue platform. This work is developed in collaboration with the RDP group at ENS-Lyon [56] and the CPIB group in Nottingham, UK [49].

- *Development of a computer platform for the 'programmable tissue'.* (Frédéric Boudon, Christophe Godin)

  One key aspect of our approach is the development of a computer platform dedicated to programming virtual tissue development. This platform will be used to carry out integration of the different models developed in this research axis. The platform is based on *OpenAlea*. Partner models can be integrated in the platform in a non-intrusive way (the code of their model need not be rewritten). In this context, model integration will i) consist of designing adequate data-structures at different levels that will be exchanged and reused among the different plug-in models and ii) defining control flows at adequate levels to avoid the burden of excessive interaction between components. In the past year, progress has been made in defining a generic tissue data structure that could be used in this platform, through several group meetings along the year. A redesign of the structure is in progress.

- *Design of a genetic model of inflorescence development.* (Etienne Farcot, Eugenio Azpeitia, Christophe Godin, François Parcy)

  We studied the regulatory network that control flower development during morphogenesis. To overcome the network complexity and integrate this regulation during ontogenesis, we have developed a first model of the control of floral initiation by genes, and in particular the situation of cauliflower mutants, in which the repeatedly meristem fails in making a complete transition to the flower. This work couples models at different scales, since gene regulation is described by a minimal gene network, which is used as a decision module in an L-system model of the inflorescence architecture. This mixed model has led us to make different hypotheses about gene interactions and hormonal regulation. First predictions about gene actors controling the passage to flower could be verified. A first integrated picture of flower development could be reached in the context of the internship of Eugenio Azpeitia (PhD Student).

5.3. Multi-scale models and analysis: from cells to plant architecture (and back)

5.3.1. Transport model in roots

Participants: Mikaël Lucas [IRD], Christophe Pradal, Christophe Godin, Christophe Maurel [BPMP].

This research theme is supported by the ANR project HydroRoot.

A model of Arabidopsis thaliana root hydraulics at the cellular level was developped in the OpenAlea modeling platform. The model relies on the integration throughout root architecture of elementary hydraulic components. Each component integrates local radial and axial water flows. Axial hydraulic conductivity is calculated according to Poiseuille's law, based on local size of xylem vessels. Radial hydraulic conductivity is determined in part by aquaporin activity and was set constant throughout root architecture in the first model versions. In its current state, the model is parameterized using architectural, tissular and physiological data that were experimentally determined in the Aquaporin group at BPMP. The architectural reconstruction of the root system is based on a tridimensional multi-scale tree graph (MTG). The current model is capable of predicting the water flow that is transported by a root system in the standard experimental conditions used in the Aquaporin group. This model was used to perform sensitivity analyses and determine the respective contributions to root hydraulic dynamics of various biological parameters (axial and radial hydraulic conductivites, root architecture). One major finding is that the root hydraulic conductivity (Lpr) computed from the model is highly dependent on root architecture. This is due to the limiting role of axial (xylem) conductance, one feature that had been neglected in previous representations of root water transport. The radial hydraulic conductivity may primarily be limiting in conditions of Lpr inhibition, since its increase from values in control roots has marginal effects on Lpr. A new set of experimental data including root diameter repartitions in wild-type plants, and xylem vessel diameters in mutants with altered xylem morphology (irx3, esk1) will be used to implement the model. Root cell hydraulic conductivities will also be measured in these and aquaporin mutant phenotypes. Our aim is to check whether, based on anatomical and morphological data, the model can properly predict the radial hydraulic conductivity of these genotypes.

5.3.2. Transport in fruits

Participants: Mik Cieslak, Nadia Bertin [Inra, Avignon], Frédéric Boudon, Christophe Godin, Michel Genard [Inra, Avignon], Christophe Goz-Bac [Université Montpellier 2].

This research theme is supported by the Agropolis project Fruit3D.

Understanding the controlling factors of fruit quality development is challenging, because fruit quality results from the interplay between physical and physiological processes that are under the control of genes and the environment. Although process-based models have been used to make significant progress in understanding these factors, they ignored to a large extent the shape and internal structure of the fruit.

To help characterizing effects of fruit shape and internal structure on quality, the creation of a 3D virtual fruit model that integrates fruit structure and function with growth governed by environmental inputs has been investigated. For this, a modeling pipeline has been developed that includes the following steps: creation of a 3D volumetric mesh of the internal fruit structure, including vasculature (see section 3). Based on previous compartment models of fruit physiology developed at Avignon, we have then developed models of water and carbon transport that have been coupled with the 3D model of fruit. In the 3D model, different equations are describing the transport between adjacent regions of the fruit represented as a 3D mesh. The integration through space and time is carried out using a standard integration scheme (Runge-Kutta of order 4).

This approach has been applied to study tomato fruit (Solanum lycopersicum) by constructing 3D volumetric meshes from different sources (images of perpendicular fruit slices and MRI data), and integrating water and carbon transport processes into these meshes. To illustrate the tomato model, a simulation of one season of the fruit's growth has been performed and its results compared with an already published process-based tomato fruit model. We first showed that our spatialized model is compliant with classical results of the abstract process-based models but also provides additional information on the internal heterogeneity of the fruit, such as a gradient in sugar concentration. Once the model is calibrated and evaluated, our approach will be suitable for studying the effects of internal fruit heterogeneity and overall shape on fruit quality development.

5.3.3. Analyzing root growth and branching

Participants: Beatriz Moreno Ortega, Sixtine Passot, Yann Guédon, Laurent Laplaze [IRD, DIADE], Mikaël Lucas [IRD, DIADE], Bertrand Muller [INRA, LEPSE].

This research theme is supported by two PhD programmes.

../../../../projets/virtual_plants/IMG/virtualfruit.png

Figure 4. Virtual models of peaches reconstructed from images with simulated vasculatur to simulate carbon and water transport in the fruit

New 2D and 3D root phenotyping plateforms are emerging with associated image analysis toolbox (e.g. SmartRoot). The analysis of complex root phenotyping data is thus a new challenge in developmental biology.

We aim at developing a pipeline of methods for analyzing root systems at three scales:

1. tissular scale to identify and characterize the meristem, elongation and mature zones along a root using piecewise heteroscedastic linear models.
2. individual root scale to analyze the dynamics of root elongation
3. root system scale to analyze the branching structure.

This pipeline of analysis methods will be applied to different species (maize, millet and *arabidopsis*) and for different biological objectives (study of genetic diversity for millet and of metabolic and hormonal controls of morphogenesis for maize).

5.3.4. Analyzing shoot and leaf elongation

Participants: Maryline Lièvre, Yann Guédon, Christine Granier [INRA, LEPSE].

This research theme is supported by one PhD programme.

The analysis of phenotyping data coming from automated platforms such as PHENOPSIS often focuses on the growth of a leaf at a given rank along the stem. We aim at developing a pipeline of methods for analyzing the growth of *arabidopsis* shoot at three scales:

1. tissular scale using a probabilistic model of endoreduplication for modeling the distribution of the leaf epidermis cell surfaces. Endoreduplication, which is a replication of the nuclear genome in the absence of cell division that leads to elevated nuclear gene content, strongly affects the leaf epidermis cells of *arabidopsis*.
2. organ scale using nonlinear regression model for analyzing the growth of each successive leaf.

3. shoot scale: The outputs of the analyses at the tissular and organ scales will be summarized as multivariate sequences along the shoots characterizing each successive leaf. These sequences will be augmented by supplementary morphological variables characterizing leaf shape and properties (e.g. presence/absence of trichomes). These sequences will be globally analyzed in order to take into account plant ontogeny and in particular the successive developmental stages before the floral transition for the wild type and selected mutants of *arabidopsis*.

5.3.5. Analyzing perturbations in Arabidopsis thaliana phyllotaxis

Participants: Christophe Godin, Yann Guédon, Yassin Refahi, Etienne Farcot, Teva Vernoux, Fabrice Besnard [RDP, ENS].

This research theme is supported by iSAM.

The geometric arrangement of lateral organs along plant stems, named phyllotaxis, shows a variety of striking patterns with remarkable regularities and symmetries. This has interested biologists, physicists, mathematicians and computer scientists for decades. These studies have lead to a commonly accepted standard interpretation of phyllotaxis that postulates that organs inhibit the formation of new organs in their vicinity. At a molecular scale, these inhibitory fields have been shown to result from the spatio-temporal distribution of the plant hormone auxin. This model theoretically explains a large part of the diversity of phyllotactic patterns observed in plants.

The cytokinin hormones are known to play a significant role in the regulation of phyllotaxis. Fabrice Besnard and Teva Vernoux realized that *Arabidopsis thalianaahp6* mutants, which are perturbed in the cytokinin signaling pathway, showed unusual chaotic perturbations of the phyllotaxis at macroscopic level.

In order to characterize these perturbations, we designed a pipeline of models and methods which relies of combinatorial and statistical techniques. Using this pipeline of methods, we have shown that the perturbation patterns in both wild-type and mutant plants can be explained by permutations in the order of insertion along the stem of 2 or 3 consecutive organs. The number of successive synchronized organs between two permutations reveals unexpected patterns that depend on the nature of the preceding permutation (2- or 3-permutation). We identified significant individual deviations of the level of baseline segments with reference to 137.5°, which confirms theoretical model predictions. Finally, we highlighted a marked relationship between permutation of organs and defects in the elongation of the internodes in between these organs.

We then looked at the origin of these permutations using confocal microscopy and realized that organs were in fact frequently co-initiated in the mutant, leading after development randomly in half of the cases to permutations. We concluded that the mutant is actually perturbed in the time between consecutive organ initiation (i.e. the plastochrone), while relative angular positions are not affected. After closer inspection, we realized that the mutated gene encode a protein diffusing from the organs and creating a field around the organs that regulates the plastochrone. We could demonstrate that in the mutant, the absence of this field lead to co-initiations and subsequently to the observed permutations.

Altogether, this study sheds a new light on our interpretation of phyllotaxis, revisiting the standard model and suggesting that several fields based on auxin and cytokinin with different properties are required to provide robustness to phyllotaxis. An overview of this work has been published in the journal Nature in December online [13]. Methodological developments were published more extensively in [20])

CORTEX Team

6. New Results

6.1. Understanding embodied neural systems

Participants: Dominique Martinez, Carlos Carvajal-Gallardo, Georgios Detorakis.

6.1.1. Bio-physical modeling and embodied olfaction

Our understanding of the computations that take place in the human brain is limited by the extreme complexity of the cortex, and by the difficulty of experimentally recording neural activities, for practical and ethical reasons. The Human Genome Project was preceded by the sequencing of smaller but complete genomes. Similarly, it is likely that future breakthroughs in neuroscience will result from the study of smaller but complete nervous systems, such as the insect brain or the rat olfactory bulb. These relatively small nervous systems exhibit general properties that are also present in humans, such as neural synchronization and network oscillations. Our goal has been therefore to understand the role of these phenomena by combining biophysical modelling and experimental recordings, before applying this knowledge to humans. In the last year, we have extended our neuronal model of the insect olfactory system. This model is capable of reproducing and explaining the stereotyped multiphasic firing pattern observed in pheromone sensitive antennal lobe neurons [10].

Using this model in robotic experiments and insect antennae as olfactory sensors, we related these multiphasic responses to action selection. The efficiency of the model for olfactory searches was demonstrated in driving the robot toward a source of pheromones. Two different classes of strategies are possible for olfactory searches, those based on a spatial map, e.g. Infotaxis, and those where the casting-and-zigzagging behaviour observed in insects is purely reactive, without any need for an internal memory, representation of the environment, or inference [15]. Our goal was to investigate this question by implementing infotactic and reactive search strategies in a robot and test them in real environmental conditions. We previously showed that robot Infotaxis produces trajectories that feature zigzagging and casting behaviours similar to those of moths, is robust and allows for rapid and reliable search processes. We have implemented infotactic and reactive search strategies in a cyborg using the antennae of a tethered moth as sensors, since no articial sensor for pheromone molecules is presently known [10].

6.1.2. Somato-sensory cortex

In a joint work with the Mnemosyne team, we have investigated the formation and maintenance of ordered topographic maps in the primary somatosensory cortex as well as the reorganization of representations after sensory deprivation or cortical lesion. We consider both the critical period (postnatal) where representations are shaped and the post-critical period where representations are maintained and possibly reorganized. We hypothesize that feed-forward thalamocortical connections are an adequate site of plasticity while cortico-cortical connections are believed to drive a competitive mechanism that is critical for learning. We model a small skin patch located on the distal phalangeal surface of a digit as a set of 256 Merkel ending complexes (MEC) that feed a computational model of the primary somatosensory cortex (area 3b). This model is a two-dimensional neural field where spatially localized solutions (a.k.a. bumps) drive cortical plasticity through a Hebbian-like learning rule. Simulations explain the initial formation of ordered representations following repetitive and random stimulations of the skin patch. Skin lesions as well as cortical lesions are also studied and results confirm the possibility to reorganize representations using the same learning rule and depending on the type of the lesion. For severe lesions, the model suggests that cortico-cortical connections may play an important role in complete recovery [11], [19], [7].

6.1.3. K-cells in visuomotor tasks

In another joint work with the Mnemosyne team, we have explored the role of the thalamus in visuomotor tasks implicating non-standard ganglion cells. Such cells in the retina have specific loci of projection in the visuomotor systems and particularly in the thalamus and the superior colliculus. In the thalamus, they feed the konio pathway of the LGN. Exploring the specificities of that pathway, we discovered it could be associated to the matrix system of thalamo-cortical projections, known to allow for diffuse patterns of connectivity and to play a major role in the synchronization of cortical regions by the thalamus. An early model led to the design of the corresponding information flows in the thalamo-cortical system, that we expanded, in the framework of the Keops project, to be applied to real visuomotor tasks [13].

We proposed to implement the computational principles raised by the study on the K-cells of the retina using a variational specification of the visual front-end, with an important consequence. In such a framework, the GC are not to be considered individually, but as a network, yielding a mesoscopic view of the retinal process. Given natural image sequences, fast event-detection properties appear to be exhibited by the mesoscopic collective non-standard behavior of a subclass of the so-called dorsal and ventral konio-cells (K-cells) that correspond to specific retinal output. We considered this visual event detection mechanism to be based on image segmentation and specific natural statistical recognition, including temporal pattern recognition, yielding fast region categorization. We discussed how such sophisticated functionalities could be implemented in the biological tissues as a unique generic two-layered non-linear filtering mechanism with feedback. We used computer vision methods to propose an effective link between the observed functions and their possible implementation in the retinal network. The available computational architecture is a two-layers network with non-separable local spatio-temporal convolution as input, and recurrent connections performing non-linear diffusion before prototype based visual event detection [17].

6.2. Neuro-inspired computational models

Participants: Yann Boniface, Benoît Chappet de Vangel, Bernard Girau, Patrick Hénaff.

6.2.1. Motion detection

We develop bio-inspired neural architectures to extract and segment the direction and speed components of the optical flow from sequences of images. Following this line, we have built additional models to code and distinguish different visual sequences. The structure of these models takes inspiration from the course of visual movement processing in the human brain, such as in area MT (middle temporal) that detects patterns of movement, or area FBA where neurons have been found to be sensitive to single spatio-temporal patterns. This work has been extended to complex movements: to fight, to wave, to clap, using real-world video databases [9].

6.2.2. Multimodal learning through joint dynamic neural fields

We have developed a coherent multimodal learning for a system with multiple sensory inputs. To this aim, we modified the BCM synaptic rule, a local learning rule, to obtain the self organization of our neuronal inputs maps and we used a CNFT based competition to drive the BCM rule. In practice, we introduced a feedback modulation of the learning rule, representing multimodal constraints of the environment. We also introduced an unlearning term in the BCM equation to solve the problem of the different temporalities between the raise of the activity within modal maps and the multimodal learning of the organization of the maps [12].

6.2.3. Adaptive sensori-motor loop

We develop bio-inspired neural controllers to control humanoids robot when they interact physically (or socially) with the human. We focus on the role of rythmicity in the interaction: how the phenomena of coupling, synchrony or others are involved in the interaction between humans? what models of neural structures can incorporate rythmicity intrinsically, and can include learning or adaptive mechanisms of the rythmicity.

6.2.4. Randomly spiking dynamic neural fields

We have defined a new kind of spiking neural field that is able to use only local links while transmitting spikes through the map by succesive random propagations. Such a model is able to be mapped onto FPGAs, while maintaining most properties of neural fields. This model has been validated from a behavioral point of view, and a fully scalable hardware implementation has been designed with several thousands of neurons on-chip. These first results are the object of an article that is currently reviewed after requested revisions.

ARAMIS Team

6. New Results

6.1. Spatial and anatomical regularization of SVM

Participants: Rémi Cuingnet, Joan Glaunès, Marie Chupin, Habib Benali, Olivier Colliot [Correspondant].

We developed a general framework to introduce spatial and anatomical priors in SVM for brain image analysis based on regularization operators. A notion of proximity based on prior anatomical knowledge between the image points is defined by a graph (e.g. brain connectivity graph) or a metric (e.g. Fisher metric on statistical manifolds). A regularization operator is then defined from the graph Laplacian, in the discrete case, or from the Laplace-Beltrami operator, in the continuous case. The regularization operator is then introduced into the SVM, which exponentially penalizes high frequency components with respect to the graph or to the metric and thus constrains the classification function to be smooth with respect to the prior. It yields a new SVM optimization problem whose kernel is a heat kernel on graphs or on manifolds. We then present different types of priors and provide efficient computations of the Gram matrix. The proposed framework is finally applied to the classification of brain magnetic resonance (MR) images (based on gray matter concentration maps and cortical thickness measures) from 137 patients with Alzheimer's disease and 162 elderly controls. The results demonstrate that the proposed classifier generates less-noisy and consequently more interpretable feature maps (Figure 1) with high classification performances.

More details in [4].

6.2. Segmentation of the hippocampus in neurodegenerative dementias

Participants: Leonardo Cruz de Souza, Marie Chupin, Maxime Bertoux, Stéphane Lehéricy, Bruno Dubois, Foudil Lamari, Isabelle Le Ber, Michel Bottlaender, Olivier Colliot [Correspondant], Marie Sarazin.

Our team develops various applications of our automatic segmentation method SACHA to neurological disorders, in particular in neurodegenerative dementias. This research is done in close collaboration with IM2A (Institut de la Mémoire et de la Maladie d'Alzheimer, Bruno Dubois and Marie Sarazin) at Pitié-Salpêtrière hospital.

We previously showed that automatic hippocampal segmentation can discriminate patients with Alzheimer's disease (AD) from elderly control subjects, with high sensitivity and specificity. In patients with Alzheimer's disease, we further studied the relationship between hippocampal atrophy and memory deficits. We also showed that hippocampal volume loss is correlated to tau and hyperphosphorylated tau levels measured in the cerebro-spinal fluid (CSF) but not with $A\beta_4 2$ levels.

Here, our objective was to study the ability of hippocampal volumetry (HV) to differentiate between two neurodegenerative dementias: Alzheimer's disease (AD) and fronto-temporal dementia (FTD). Seventy-two participants were included: 31 AD patients with predominant and progressive episodic memory deficits associated with typical AD cerebrospinal fluid (CSF) profile and/or positive amyloid imaging (PET with 11C-labeled Pittsburgh Compound B [PiB]), 26 patients with behavioral variant FTD (bvFTD) diagnosed according to consensual clinical criteria and with no AD CSF profile, and 15 healthy controls without amyloid retention on PiB-PET exam. HV were segmented with our automated method and were normalized to total intracranial volume (nHV). Significant reductions in HV were found in both AD and bvFTD patients compared with controls, but there were no significant difference between AD and bvFTD patients. Mean nHV distinguished normal controls from either AD or bvFTD with high sensitivity (80.6% and 76.9%, respectively) and specificity (93.3% for both), but it was inefficient in differentiating AD from bvFTD (9.7% specificity). There was no difference in the clinical and neuropsychological profiles according to HV in bvFTD and AD patients. In conclusion, when considered alone, measures of HV are not good markers to differentiate AD from bvFTD. Hippocampal sclerosis associated with FTD may explain the high degree of overlap in nHV between both groups.

`../../../../projets/aramis/IMG/fig20eps.png`

Figure 1. Anatomical regularization of support vector machines for automatic classification of patients with Alzheimer's disease. The figure displays the normalized vector orthogonal to the optimal margin hyperplane, for increasing levels of regularization.

More details in [5].

6.3. Diffeomorphic Iterative Centroids for Template Estimation on Large Datasets

Participants: Claire Cury [Correspondant], Joan Glaunès, Olivier Colliot.

A common approach for analysis of anatomical variability relies on the estimation of a template representative of the population. The Large Deformation Diffeomorphic Metric Mapping is an attractive framework for that purpose. However, template estimation using LDDMM is computationally expensive, which is a limitation for the study of large datasets. We proposed an iterative method which quickly provides a centroid of the population in the shape space. This centroid can be used as a rough template estimate or as initialization of a template estimation method. The approach was evaluated on datasets of real and synthetic hippocampi segmented from brain MRI. The results showed that the centroid is correctly centered within the population and is stable for different orderings of subjects. When used as an initialization, the approach allows to substantially reduce the computation time of template estimation.

More details in [30].

6.4. Sparse Adaptive Parameterization of Variability in Image Ensembles

Participants: Stanley Durrleman [Correspondant], Sarang Joshi, Stéphanie Allassonnière.

We introduce a new parameterization of diffeomorphic deformations for the characterization of the variability in image ensembles. Dense diffeomorphic deformations are built by interpolating the motion of a finite set of control points that forms a Hamiltonian flow of self-interacting particles. The proposed approach estimates a template image representative of a given image set, an optimal set of control points that focuses on the most variable parts of the image, and template-to-image registrations that quantify the variability within the image set. The method automatically selects the most relevant control points for the characterization of the image variability and estimates their optimal positions in the template domain. The optimization in position is done during the estimation of the deformations without adding any computational cost at each step of the gradient descent. The selection of the control points is done by adding a L^1 prior to the objective function, which is optimized using the FISTA algorithm.

Related publication: [12]

6.5. Toward a comprehensive framework for the spatiotemporal statistical analysis of longitudinal shape data

Participants: Stanley Durrleman [Correspondant], Xavier Pennec, Alain Trouvé, José Braga, Guido Gerig, Nicholas Ayache.

We introduce a comprehensive framework for the statistical analysis of longitudinal shape data. The proposed method allows the characterization of typical growth patterns and subject-specific shape changes in repeated time-series observations of several subjects. This can be seen as the extension of usual longitudinal statistics of scalar measurements to high-dimensional shape or image data.

The method is based on the estimation of continuous subject-specific growth trajectories and the comparison of such temporal shape changes across subjects. Differences between growth trajectories are decomposed into morphological deformations, which account for shape changes independent of time, and time warps, which account for different rates of shape changes over time.

Given a longitudinal shape data set, we estimate a mean growth scenario representative of the population, and the variations of this scenario both in terms of shape changes and in terms of change in growth speed. Then, intrinsic statistics are derived in the space of spatiotemporal deformations, which characterize the typical variations in shape and in growth speed within the studied population. They can be used to detect systematic developmental delays across subjects.

../../../../projets/aramis/IMG/IJCV_Digit2.jpg

Figure 2. Left: template image estimated from 20 images of the US postal database. Momentum vectors are placed at the most variable places and paramterize mappings from the template to each image in the data set. Right: sample images from the data set (top) and template image deformed to match the corresponding sample image (bottom)

In the context of neuroscience, we apply this method to analyze the differences in the growth of the hippocampus in children diagnosed with autism, developmental delays and in controls. Result suggest that group differences may be better characterized by a different speed of maturation rather than shape differences at a given age. In the context of anthropology, we assess the differences in the typical growth of the endocranium between chimpanzees and bonobos. We take advantage of this study to show the robustness of the method with respect to change of parameters and perturbation of the age estimates.

Related publication: [13]

6.6. Bayesian Atlas Estimation for the Variability Analysis of Shape Complexes

Participants: Pietro Gori [Correspondant], Olivier Colliot, Yulia Worbe, Linda Marrakchi-Kacem, Sophie Lecomte, Cyril Poupon, Andreas Hartmann, Nicholas Ayache, Stanley Durrleman.

We propose a Bayesian framework for multi-object atlas estimation based on the metric of currents which permits to deal with both curves and surfaces without relying on point correspondence. This approach aims to study brain morphometry as a whole and not as a set of different components, focusing mainly on the shape and relative position of different anatomical structures which is fundamental in neuro-anatomical studies. We propose a generic algorithm to estimate templates of sets of curves (fiber bundles) and closed surfaces (sub-cortical structures) which have the same “form” (topology) of the shapes present in the population. This atlas construction method is based on a Bayesian framework which brings to two main improvements with respect to previous shape based methods. First, it allows to estimate from the data set a parameter specific to each object which was previously fixed by the user: the trade-off between data-term and regularity of deformations. In a multi-object analysis these parameters balance the contributions of the different objects and the need for an automatic estimation is even more crucial. Second, the covariance matrix of the deformation parameters is estimated during the atlas construction in a way which is less sensitive to the outliers of the population.

Related publication: [33]

6.7. Geodesic regression of shape and image data

Participants: James Fishbaugh [Correspondant], Marcel Prastawa, Guido Gerig, Stanley Durrleman.

Shape regression is emerging as an important tool for the statistical analysis of time dependent shapes. We develop a new generative model which describes shape change over time, by extending simple linear regression to the space of shapes represented as currents in the large deformation diffeomorphic metric mapping (LDDMM) framework. By analogy with linear regression, we estimate a baseline shape (intercept) and initial momenta (slope) which fully parameterize the geodesic shape evolution. This is in contrast to previous shape regression methods which assume the baseline shape is fixed. We further leverage a control point formulation, which provides a discrete and low dimensional parameterization of large diffeomorphic transformations. This flexible system decouples the parameterization of deformations from the specific shape representation, allowing the user to define the dimensionality of the deformation parameters. We present an optimization scheme that estimates the baseline shape, location of the control points, and initial momenta simultaneously via a single gradient descent algorithm.

Shapes can be given as 3D meshes (as in [32]) or as 3D images (as in [31]).

Related publications: [32], [31].

6.8. Discriminating brain microbleeds using phase contrast MRI in a multicentre clinical dataset

Participants: Takoua Kaaouana [Correspondant], Marie Chupin, Didier Dormont, Ludovic de Rochefort, Thomas Samaille.

../../../../projets/aramis/IMG/IJCV_4D_Regression.jpg

a- Temporal regression of endocasts of bonobos (top) and chimpanzees (bottom)

../../../../projets/aramis/IMG/IJCV_4D_Registration.jpg

b - morphological deformation

../../../../projets/aramis/IMG/Pietro_Final_template.png

Figure 4. Atlas construction from a data set of left caudate nucleus and its associated fiber bundle that were segmented in images of patients with Gilles de la Tourette syndrome and controls. An initial template complex determines the topology of the model. Its shape is optimized given the patients data or the controls data only, thus resulting in two atlases showing different distributions of the fibers on the surface of the nucleus.

../../../../projets/aramis/IMG/IPMI_James.png

Figure 5. Average development of genu fiber tract from 2 to 24 months. Top row shows observed data for all subjects, which is clustered around 2, 12, and 24 months. Bottom row shows genu fiber tracts estimated from geodesic regression at several time points with velocity of fiber development displayed on the estimated fibers.

Brain microbleeds (BMBs) have emerged as a new imaging marker of small vessel diseases and they may play a crucial role in degenerative pathology such as Alzheimer's disease. Composed of hemosiderin, BMBs can be efficiently detected with MRI sequences sensitive to magnetic susceptibility (e.g. gradient recalled echo T2*W images). Nevertheless, that identification remains challenging because of confounding structures and lesions. Most T2*-weighted hyposignals result from local magnetic field inhomogeneity and can be identified either as BMBs, veins or brain micro-calcifications (BMCs). Differential diagnosis of BMBs and BMCs usually requires an additional CT scan. Quantitative susceptibility mapping techniques were proposed to discriminate between diamagnetic and paramagnetic structures, but they require a full 3D dataset and complex post-processing. We introduced a fast 2D phase processing technique including unwrapping and harmonic filtering thus yielding the internal field map, namely the field map generated only by sources within the volume of interest. We demonstrate its applicability and robustness on multicenter data acquired in standardized clinical settingand and its ability to discriminate between paramagnetic BMBs and diamagnetic BMCs through the use of the orientation of the dipolar pattern.

Related publications: [36].

6.9. Network symmetries and functional modules in the brain

Participants: Vincenzo Nicosia, Miguel Valencia, Mario Chavez [Correspondant], Albert Diaz-Guilera, Vito Latora.

We study the classical Kuramoto model in which the oscillators are associated to the nodes of a network and the interactions include a phase frustration, thus preventing full synchronization. The system organizes into a regime of remote synchronization where pairs of nodes with the same network symmetry are fully synchronized, despite their distance on the graph. We provide analytical arguments to explain this result and we show how the frustration parameter affects the distribution of phases. An application to brain networks suggests that anatomical symmetry plays a role in neural synchronization by determining correlated functional modules across distant locations.

Related publication: [19]

6.10. Accessibility of cortical networks during motor tasks

Participants: Mario Chavez [Correspondant], Fabrizio de Vico Fallani, Miguel Valencia, Mario Chavez, Julio Artieda, Vito Latora, Donatella Mattia, Fabio Babiloni.

Recent findings suggest that the preparation and execution of voluntary self-paced movements are accompanied by the coordination of the oscillatory activities of distributed brain regions. We used electroencephalographic source imaging methods to estimate the cortical movement-related oscillatory activity during finger ex- tension movements. We applied network theory to investigate changes (expressed as differences from the baseline) in the connectivity structure of cortical networks related to the preparation and execution of the movement. We computed the topological accessibility of different cortical areas, measuring how well an area can be reached by the rest of the network. Analysis of cortical networks revealed specific agglomerates of cortical sources that become less accessible during the preparation and the execution of the finger movements. The observed changes neither could be explained by other measures based on geodesics or on multiple paths, nor by power changes in the cortical oscillations.

Related publication: [3]

6.11. Abnormal functional connectivity between motor cortex and pedunculopontine nucleus following chronic dopamine depletion

Participants: Miguel Valencia, Mario Chavez [Correspondant], Julio Artieda, J. Paul Bolam, Juan Mena-Segovia.

../../../../projets/aramis/IMG/ImageTakoua.png

Figure 6. Philips (a,b,c,d) and Siemens (e,f,g,h) sample cases. Magnitude image (a,e), native phase image (b,f), and internal field map; axial (c,g) and sagittal (d,h). Zoom in white rectangle showing a dipolar pattern BMB (white arrow) and a physiologic calcification of the choroid plexus (black arrow).

../../../../projets/aramis/IMG/synchSymmetry.png

Figure 7. a) Brain areas with similar and dissimilar phases of the frustrated Kuramoto model are colored and superimposed onto an anatomical image. b) Examples of functional data from one subject recorded at the brain areas indicated in panel a). Colors are the same as those used in the anatomical image. c) Functional correlation (normalised values) Z *between pairs of nodes as a function of their phase differences* $\Delta\theta$ *according to the simulated Kuramoto dynamics. The black solid curve corresponds to the average value over all the subjects, while the gray area covers the* 5^{th} *and the* 95^{th} *percentiles of the distribution. The dashed horizontal line indicates the threshold for statistical significant correlations (*$p < 0.05$*, corrected for multiple comparisons).*

../../../../projets/aramis/IMG/generalScheme4eps.png

Figure 8. a) Averaged EMG and EEG (recorded at the postcentral region) signals of a subject during the execution of finger movements. Boxes define the three temporal epochs of EEG activity studied here: baseline (BASE), preparation (PRE) and execution period (EXE). Vertical dotted line indicates the movement onset. Examples of scalp and source-level networks obtained from one subject, at the frequency band Beta1, during the epoch EXE are shown in panels b) and c), respectively. Color map codes the number of connections.

The activity of the basal ganglia is altered in Parkinson's disease (PD) as a consequence of the degeneration of dopamine neurons in the substantia nigra pars compacta. This results in aberrant discharge patterns and expression of exaggerated oscillatory activity across the basal ganglia circuit. Altered activity has also been reported in some of the targets of the basal ganglia, including the pedunculopontine nucleus (PPN), possibly due to its close interconnectivity with most regions of the basal ganglia. However, the nature of the involvement of the PPN in the pathophysiology of PD has not been fully elucidated. We recorded local field potentials in the motor cortex and the PPN in the 6-hydroxydopamine (6-OHDA)-lesioned rat model of PD under urethane anesthesia. By means of linear and nonlinear statistics, we analyzed the synchrony between the motor cortex and the PPN, and the delay in the interaction between these two structures. We observed the presence of coherent activity between the cortex and the PPN in low- (5-15 Hz) and high-frequency bands (25-35 Hz) during episodes of cortical activation. In each case the cortex led the PPN. Dopamine depletion strengthened the interaction of the low-frequency activities by increasing the coherence specifically in the theta and alpha ranges and reduced the delay of the interaction in the gamma band. Our data show that cortical inputs play a determinant role in leading the coherent activity with the PPN, and support the involvement of the PPN in the pathophysiology of PD.

Related publication: [25]

6.12. Subthalamic Nucleus High-Frequency Stimulation Restores Altered Electrophysiological Properties of Cortical Neurons in Parkinsonian Rat

Participants: Bertrand Degos, Jean Michel Deniau, Mario Chavez [Correspondant], Nicolas Maurice.

Electrophysiological recordings performed in parkinsonian patients and animal models have confirmed the occurrence of alterations in firing rate and pattern of basal ganglia neurons, but the outcome of these changes in thalamo-cortical networks remains unclear. Using rats rendered parkinsonian, we investigated, at a cellular level in vivo, the electrophysiological changes induced in the pyramidal cells of the motor cortex by the dopaminergic transmission interruption and further characterized the impact of high-frequency electrical stimulation of the subthalamic nucleus, a procedure alleviating parkinsonian symptoms. We provided evidence that a lesion restricted to the substantia nigra pars compacta resulted in a marked increase in the mean firing rate and bursting pattern of pyramidal neurons of the motor cortex. These alterations were underlain by changes of the electrical membranes properties of pyramidal cells including depolarized resting membrane potential and increased input resistance. The modifications induced by the dopaminergic loss were more pronounced in cortico-striatal than in cortico-subthalamic neurons. Furthermore, subthalamic nucleus high- frequency stimulation applied at parameters alleviating parkinsonian signs regularized the firing pattern of pyramidal cells and restored their electrical membrane properties.

Related publication: [7]

6.13. Non-parametric resampling of random walks for spectral networks clustering

Participants: Fabrizio de Vico Fallani [Correspondant], Vincenzo Nicosia, Vito Latora, Mario Chavez.

Parametric resampling schemes have been recently introduced in complex network analysis with the aim of assessing the statistical significance of graph clustering and the robustness of community partitions. We proposed a method to replicate structural features of complex networks based on the non-parametric resampling of the transition matrix associated with an unbiased random walk on the graph. We tested this bootstrapping technique on synthetic and real-world modular networks and we showed that the ensemble of replicates obtained through resampling can be used to improve the performance of standard spectral algorithms for spectral clustering of graphs.

Related publication: [43]

6.14. Multiscale topological properties of functional brain networks during motor imagery after stroke

Participants: Fabrizio de Vico Fallani [Correspondant], Floriana Pichiorri, Giovani Morone, Marco Molinari, Fabio Babiloni, Febo Cincotti, Donatella Mattia.

In recent years, network analyses have been used to evaluate brain reorganization following stroke. However, many studies have often focused on single topological scales, leading to an incomplete model of how focal brain lesions affect multiple network properties simultaneously and how changes on smaller scales influence those on larger scales. In an EEG-based experiment on the performance of hand motor imagery (MI) in 20 patients with unilateral stroke, we observed that the anatomic lesion affects the functional brain network on multiple levels. In the beta (13–30 Hz) frequency band, the MI of the affected hand (Ahand) elicited a significantly lower smallworldness and local efficiency (Eloc) versus the unaffected hand (Uhand). Notably, the abnormal reduction in Eloc significantly depended on the increase in interhemispheric connectivity, which was in turn determined primarily by the rise of regional connectivity in the parieto-occipital sites of the affected hemisphere. Further, in contrast to the Uhand MI, in which significantly high connectivity was observed for the contralateral sensorimotor regions of the unaffected hemisphere, the regions with increased connectivity during the Ahand MI lay in the frontal and parietal regions of the contralaterally affected hemisphere. Finally, the overall sensorimotor function of our patients, as measured by Fugl–Meyer Assessment (FMA) index, was significantly predicted by the connectivity of their affected hemisphere. These results improve on our understanding of stroke-induced alterations in functional brain networks.

Related publication: [6]

6.15. Wavelet analysis in ecology and epidemiology: impact of statistical tests

Participants: Bernard Cazelles, Kevin Cazelles, Mario Chavez [Correspondant].

Wavelet analysis is now frequently used to extract information from ecological and epidemiological time series. Statistical hypothesis tests are conducted on associated wavelet quantities to assess the likelihood that they are due to a random process. Such random processes represent null models and are generally based on synthetic data that share some statistical characteristics with the original time series. This allows the comparison of null statistics with those obtained from original time series. When creating synthetic datasets, different techniques of resampling result in different characteristics shared by the synthetic time series. Therefore, it becomes crucial to consider the impact of the resampling method on the results. We have addressed this point by comparing seven different statistical testing methods applied with different real and simulated data. Our results showed that statistical assessment of periodic patterns is strongly affected by the choice of the resampling method, so two different resampling techniques could lead to two different conclusions about the same time series. Moreover, we showed the inadequacy of resampling series generated by white noise and red noise that are nevertheless the methods currently used in the wide majority of wavelets applications in epidemiology. Our results highlight that the characteristics of a time series, namely its Fourier spectrum and autocorrelation, are important to consider when choosing the resampling technique. Results suggest that data-driven resampling methods should be used such as the hidden Markov model algorithm and the ‘beta-surrogate’ method.

Related publication: [2]

../../../../projets/aramis/IMG/figureFabrizio.jpg

Figure 9. Grand average of brain networks in the Beta band during the MI of the unaffected Uhand and affected Ahand hand. Top plots: Scalp representation relative to Uhand (panel A) and Ahand (panel B) condition. Nodes are positioned according the actual EEG montage scheme. Blue and red lines denote the links within the unaffected (Uhemi) and the affected (Ahemi) hemisphere, respectively. Gray lines denote the inter-hemispheric links. The intensity of the color and the thickness of the lines vary as function of the number of patients exhibiting that significant link. Bottom part: graph representation of the brain networks relative to Uhand (panel A) and Ahand (panel B) condition. In this representation nodes are spatially repositioned through a force-based algorithm so that all the links are approximately of equal length with as few crossing edges as possible. Only links that were in common to more than 4 patients (20% of the sample) are illustrated here. Blue and red nodes indicate scalp electrodes placed over the undamaged (Uhemi) and damaged (Ahemi) hemisphere, respectively. The midline scalp electrodes (from Fpz to Oz) are illustrated as white nodes

ASCLEPIOS Project-Team

5. New Results

5.1. Medical Image Analysis

5.1.1. Segmentation of cardiac images from magnetic resonance

Participants: Jan Margeta [Correspondant], Kristin Mcleod, Antonio Criminisi [MSRC], Nicholas Ayache.

This work has been partly supported by Microsoft Research through its PhD Scholarship Programme and the European Research Council through the ERC Advanced Grant MedYMA (on Biophysical Modeling and Analysis of Dynamic Medical Images).

Cardiac imaging, Magnetic resonance, Image segmentation, Maching learning

- We contributed our previous method to build left ventricle myocardium segmentation consensus based on the STAPLE algorithm [26]
- We enhanced our segmentation method with extra features based on the distance transform and image vesselness measures in order to segment left atria (see Fig. 1) from 3d MRI images [49]. We participated with this method in the left atrium segmentation challenge at MICCAI 2013.

Figure 1. Segmented atria meshes from the validation dataset.

5.1.2. Brain tumor image processing and modeling

Participants: Bjoern Menze [Correspondant], Hervé Delingette, Nicholas Ayache, Nicolas Cordier, Erin Stretton, Jan Unkelbach.

We developed a new non-parametric lesion growth model for the analysis of longitudinal image sequences [59], evaluated the parametric tumor growth model of Konukoglu on longitudinal data, focusing on the relevance of DTI [40], [57], and addressed the question of how to detect tumor growth from longitudinal sequences of patients treated with angiogenesis inhibitors using registration techniques [47]. We also completed work for the 2012 MICCAI Challenge on Brain Tumor Image Segmentation (MICCAI-BRATS 2012) [79], where we also tested some of our own brain tumor image segmentation models based on random forests [42] and patch regression [38], we also participated in MICCAI-BRATS 2013 in Nagoya, Japan [67].

5.1.3. Further developing the random forest framework for medical computer vision tasks

Participants: Bjoern Menze [Correspondant], Matthias Schneider, Ezequiel Geremia, Rene Donner, Georg Langs, Gabor Szekely.

Methodological contributions include the further development of the random forest framework. We introduced the “spatially adaptive” random forest (SARF) classifier [42], and evaluated Hough regression forests for interest point detection in whole body CT image analysis, as well as for vessel detection and tracking [54]. We also evaluated alternative patch-based methods for whole body image registration [41]. As a related community effort, we organized the MICCAI-MCV workshop, also in conjunction with the MICCAI conference in Nagoya, Japan [65].

5.1.4. Statistical Analysis of Diffusion Tensor Images of the Brain

Participants: Vikash Gupta [Correspondent], Nicholas Ayache, Xavier Pennec.

Diffusion Tensor Imaging of the Brain, Tractography, Super-resolution, Statistical analysis

Diffusion tensor imaging (DTI) is gaining interest as a clinical tool for studying a number of brain diseases pertaining to white matter tracts and also as an aid in neuro-surgical planning. Unfortunately, in a clinical environment, diffusion imaging is hampered by the long acquisition times, low signal to noise ratio and a prominent partial volume effect due to thick slices. We are developing a framework for increasing the resolution of the low-resolution clinical CTI images. The method uses a maximum likelihood strategy to account for the noise and an anisotropic regularization prior to promote smoothness in homogeneous areas while respecting edges. The technique is called Higher Resolution Tensor Estimation and it uses a single clinical acquisition to produce high resolution tensor images. We aim to replace resampling techniques used for tensor normalization in population based studies, with the present method. The method itself along with quantitative results on tractography 2 were presented in MICCAI 2013 [45].

5.1.5. 3D/2D coronary arteries registration

Participants: Thomas Benseghir [Correspondant], Grégoire Malandain, Régis Vaillant [GE-Healthcare], Nicholas Ayache.

This work is done in collaboration with GE-Healthcare (Buc).

3D/2D registration; computed tomography angiography; CTA; X-ray fluoroscopy; coronary arteries

Endovascular treatment of coronary arteries involves catheter navigation through patient vasculature. Projective angiography guidance is limited in the case of chronic total occlusion where the occluded vessel can not be seen. Integrating standard preoperative CT angiography information with live fluoroscopic images addresses this limitation but requires alignment of both modalities.

We published the Iterative Closest Curve (ICC) algorithm [36] in the MICCAI 2013 conference :

- The ICC-algorithm mimics the ICP-algorithm [6], curves being considered instead of points.
- Contrary to closest point pairing, the resulting pairings assure a topological and geometrical coherence since a curve is paired to another one (cf Figure 3).
- The developed method can deal with differences in both datasets by considering outlier rejection at the level of curve and the level of point.

5.1.6. Automatic Registration of Endoscopic Images

Participants: Anant Vemuri [Correspondant], Stéphane Nicolau [IHU Strasbourg], Luc Soler [IHU Strasbourg], Nicholas Ayache.

This work is performed in collaboration with IHU Strasbourg.

Image registration; Endoscopic imaging; Biopsy Relocalization

---

[6] P.J. Besl and N.D. McKay. A method for registration of 3-D shapes

../../../../projets/asclepios/IMG/DTI.jpg

Figure 2. Middle column shows a comparatively dense fiber bundle in the fornix region for the Higher Resolution Tensor Estimation method (superior-inferior view) compared to tensor resampling (left column). Right column shows a quantitative comparison of fiber lengths.

Figure 3. Left : Pairing obtained with the ICP algorithm, Middle : Pairing obtained with the ICC algorithm, Right : ICC (green) and ICP (red) registration results.

The screening of cancer lesions in the oesophagus involves obtaining biopsies at different regions along the oesophagus. The localization and tracking of these biopsy sites inter-operatively poses a significant challenge for providing targeted treatments.

Our work [61] introduces a novel framework for accurate re-positioning of the endoscope at previously targeted sites:

- it includes an electromagnetic tracking system in the loop and provides a framework for utilizing it for re-localization inter-operatively.
- We have shown on three in-vivo porcine interventions that our system can provide accurate guidance information, which was qualitatively evaluated by five experts.

5.2. Biological Image Analysis

5.2.1. Pre-clinical molecular imaging: motionless 3D image reconstruction in micro-SPECT

Participants: Marine Breuilly [Correspondant, Inria], Grégoire Malandain [Inria], Nicholas Ayache [Inria], Jacques Darcourt [UNS-CAL], Philippe Franken [UNS-CAL], Thierry Pourcher [CEA].

This work is jointly conducted with the Transporter in Imagery and Oncologic Radiotherapy team (TIRO, CEA-CAL-UNSA) located in Nice.

SPECT/CT, small animal, respiratory motion, respiratory gating, 4D images, stomach, 99mTc-pertechnetate biodistribution, compartmental analysis

This work has been conducted on SPECT images acquired with a small animal device. Dynamic SPECT images provide functional information targeted by a specific radiotracer (99mTc-pertechnetate) that permit the tracking and quantifying of evolving phenomena.

- Respiratory motion induces an artificial enlargement of the moving structures (tumours, organs) in SPECT images, and biases the quantification.
- A full ad-hoc method was presented that allows the reconstruction of a single 3D SPECT image without motion artefacts [37], [6], [1].

5.2.2. Pre-clinical molecular dynamic imaging: ^{99}m*Tc-pertechnetate biodistribution model of murine stomach with micro-SPECT*

Participants: Marine Breuilly [Correspondant, Inria], Grégoire Malandain [Inria], Nicholas Ayache [Inria], Jacques Darcourt [UNS-CAL], Philippe Franken [UNS-CAL], Thierry Pourcher [CEA].

This work is jointly conducted with the Transporter in Imagery and Oncologic Radiotherapy team (TIRO, CEA-CAL-UNSA) located in Nice.

SPECT/CT, small animal, 4D images, stomach, ^{99}mTc-pertechnetate biodistribution, compartmental analysis

Using the coupled SPECT and CT device dedicated to small animals, functional information targeted by a specific radiotracer (^{99}mTc-pertechnetate) can be imaged dynamically.

- ^{99}mTc-pertechnetate is an iodide analog related to the NIS gene. Thus iodide uptake kinetics can be studied through the study of ^{99}mTc- pertechnetate biodistribution.
- Dynamic SPECT images exhibit a progressive accumulation of ^{99}mTc-pertechnetate in the stomach wall and diffusion in the stomach cavity.
- A first simplified model for stomach ^{99}mTc-pertechnetate biodistribution was proposed and studied with a compartmental analysis approach using a simplified two-compartment (stomach wall and cavity) model with one input (blood) (see Figure 4) [1].
- Time activity curves of each compartment were obtained from dynamic images thanks to an original layer-based decomposition of the stomach [1].
- The first estimation of the model transfer parameters K_{ij} was performed by numerically solving the inverse problem [1].

../../../../projets/asclepios/IMG/breuilly.png

Figure 4. Simplified two-compartment model of ^{99}m*Tc-pertechnetate biodistribution in murine stomach with time-activity curves for each compartment.*

5.3. Computational Anatomy

5.3.1. Longitudinal brain morphometry: statistical analysis and robust quantification of anatomical changes

Participants: Marco Lorenzi [Correspondant], Xavier Pennec, Nicholas Ayache.

Longitudinal analysis, Alzheimer's Disease, non-linear registration, brain morphometry

This project is based on the PhD thesis defended in 2012 by Marco Lorenzi, and aims at developing robust and effective instruments for the analysis of longitudinal brain changes, with special focus on the study of brain atrophy in Alzheimer's disease. The project relies on the analysis of follow-up magnetic resonance images of the brain by means of non-linear registration. During 2013 the main scientific achievements were the following:

- We developed and distributed the LCC-logDemons, an accurate and robust diffeomorphic non-linear registration algorithm [14], [16]. The algorithm implements the symmetric Local Correlation Coefficient (LCC) and is suited for both inter and intra-subject registration. The software is freely available for research purposes here.
- We investigated the problem of comparing the trajectories of longitudinal morphological changes estimated in different patients. Based on our previous work on parallel transport in diffeomorphic registration, we proposed the "pole ladder" for the efficient normalization of longitudinal trajectories in a common reference space [15], [48].
- We defined an effective framework for the statistical analysis of longitudinal brain changes in clinical groups. The proposed framework enabled the characterization of abnormal morphological changes in healthy subjects at risk for Alzheimer's disease [46].
- We addressed the multi-scale analysis of longitudinal volume changes encoded by deformation fields. We provided a probabilistic framework for the consistent definition of anatomical regions of longitudinal brain atrophy across spatial scales, in order to robustly quantify regional volume changes in populations or in single patients. The framework was applied to the longitudinal analysis of group-wise atrophy in Alzheimer's disease (Figure 5), and to the tracking and quantification of treatment efficacy on brain tumors [47].

5.3.2. Longitudinal Analysis and Modeling of Brain Development during Adolescence

Participants: Mehdi Hadj-Hamou [Correspondant], Xavier Pennec, Nicholas Ayache.

This work is partly funded through the ERC Advanced Grant MedYMA 2011-291080 (on Biophysical Modeling and Analysis of Dynamic Medical Images).

Brain development, adolescence, longitudinal analysis, non-rigid registration algorithm

Due to the lack of tools to capture the subtle changes in the brain, little is known about its development during adolescence. The aim of this project is then to provide quantification and models of brain development during adolescence based on non-rigid registration of longitudinal MRIs (enabling us to capture these changes). The analysis pipeline is the following (Figure 6) :

- Register each patient's pair of images in order to get access to the longitudinal changes defined by a transformation field (parameterized by a Stationary Velocity Field).
- Transport every deformation field in a common space (template) to obtain the mean scenario and quantify the changes.
- Propose simplified models of the anatomical changes occurring during adolescence abstracting the results of the analysis.

5.3.3. Reduced-Order Statistical Models of Cardiac Growth, Motion and Blood Flow

Participants: Kristin Mcleod [Correspondant], Maxime Sermesant, Xavier Pennec.

This work was partially funded by the EU projects Care4me (ITEA2) and MD-Paedigree (FP7).

../../../../projets/asclepios/IMG/FigureMarcoLorenzi.png

Figure 5. Group-wise scale-space analysis for the 1-year brain atrophy in 30 AD patients.

`../../../../projets/asclepios/IMG/Pipeline_all_wt.png`

Figure 6. Pipeline for the longitudinal analysis of brain development during adolescence.

Statistical analysis, image registration, Demons algorithm, reduced models, CFD, Polyaffine, cardiac motion tracking

This work involves developing reduced models of cardiac growth, motion and blood flow, with application to the Tetralogy of Fallot heart [28].

- Extending the 2012 reduced order model of cardiac motion based on a polyaffine log-demons registration proposed at the 2012 STACOM MICCAI workshop, an additional cardiac-specific prior was added to the model to give more physiologically meaningful weight functions. Using this method, the trace of the affine matrix per region was plotted over time to establish differences between healthy subjects and asynchronous heart failure patients. The method and results were presented at the 2013 FIMH conference [52].
- Going further in analysing the affine parameters per region, statistical methods were applied to the registration parameters of the method proposed at the 2012 STACOM MICCAI workshop [50]. By applying principal component analysis to the transformation parameters stacked either column-wise or row-wise, population-based descriptors of motion in terms of the temporal or spatial components were obtained. The method was applied to 15 healthy subjects and 2 heart failure patients and presented at the 2013 MICCAI conference [51].
- The analysis of a statistical model for reduced blood flow simulations in the pulmonary artery proposed in the 2010 STACOM workshop was extended to a journal version [10], [64]. The previous work was extended to re-solve the obtained pressure and velocity bases for the subject-specific geometry by solving the Navier Stokes equations on the reduced bases. The method was applied to a data-set of 17 Tetralogy of Fallot patients.

5.3.4. Geometric Statistics

Participants: Xavier Pennec [Correspondant], Nina Miolane, Christof Seiler [Stanford], Susan Holmes [Stanford].

This work is partly funded through a France Stanford collaborative project grant (2013-2014).

Statistics, manifolds, Lie groups

The study of bi-invariant means on Lie groups [53] was further pushed by looking for the conditions of existence of bi-invariant semi-Riemannian metrics, thus relaxing the positivity constraint of Riemannian metrics [4]. This idea was based on the fact that such a bi-invariant semi-Riemannian metric exists of SE(3). Unfortunately, this does not generalize to higher dimensions. Other results on geometric statistics on regions for in the context of group-valued trees for deformation analysis were presented in [55].

5.4. Computational Physiology

5.4.1. Modeling and Simulation of Longitudinal Brain MRIs with Atrophy in Alzheimer's Disease

Participants: Bishesh Khanal [Correspondant], Nicholas Ayache, Xavier Pennec.

Alzheimer's Disease (AD), modeling atrophy, bio-physical model, simulation

We have implemented a 3D bio-physical model for the deformation of the brain with Alzheimer's Disease (AD).The model produces a deformation field of the brain when a known distribution of local volume change (atrophy) is given as the input. The obtained deformation is then used to warp the original 3D MR image. The major contribution of this work corresponds to the block “Brain Deformation” in Figure 7 .

5.4.2. Registration of time series of cardiac images

Participants: Loic Le Folgoc [Correspondant], Hervé Delingette, Antonio Criminisi, Nicholas Ayache.

This work has been partly supported by Microsoft Research through its PhD Scholarship Programme and the European Research Council through the ERC Advanced Grant MedYMA (on Biophysical Modeling and Analysis of Dynamic Medical Images).

`../../../../projets/asclepios/IMG/adModel.png`

Figure 7. Modeling and simulation of longitudinal brain MRIs with atrophy in Alzheimer's Disease

Registration, Automatic Relevance Determination, Magnetic resonance, 3D-US

- We developed a generic approach to registration building on the framework of Automatic Relevance Determination. We applied this framework to the tracking of heart motion throughout time series of images from cine-MR, tagged-MR and 3D-US modalities.
- Our approach allows for the joint determination of model parameters, such as noise and regularization parameters, decreasing the need for manual tuning and preprocessing. Moreover, it is suitable for further analysis of uncertainty in the output of the registration.

`../../../../projets/asclepios/IMG/mesh.png` `../../../../projets/asclepios/IMG/field`

Figure 8. An instance of motion tracking on a cine MR frame. (Left) Mesh contour propagated to end systole via the registration output (Right) Computed displacement field.

5.4.3. Real-Time Cardiac Electrophysiology Computing for Training Simulator

Participants: Hugo Talbot [Correspondant], Hervé Delingette, Stephane Cotin, Maxime Sermesant, Christian Duriez.

This work was performed in collaboration with the SHACRA team in Lille.

Cardiac electrophysiology simulation, Cryoablation simulation, SOFA framework, GPU computing, patient-specific study

- Cardiac arrhythmia is a very frequent pathology that comes from abnormal electrical activity in the myocardium. This work aims at developing a training simulator for cardiologists in the context of catheterization and thermo-ablation of these arrhythmias. After tackling the issue of fast electrophysiology computation [27], a first version of our training simulator was proposed which combines virtual catheterization and interactive GPU electrophysiology modeling [70]. This year, the simulator has been improved by tackling the issue of interactive catheter navigation inside a moving venous system and a beating heart [70]. The simulator was demonstrated during the VRIPHYS 2013 workshop in Lille and the Inria-industry meeting in Paris. Personalization of the electrophysiological model using the data assimilation library Verdandi has been initiated.
- Cryotherapy simulation in collaboration with the IHU Strasbourg has been performed. This technique consists in inserting needles that freeze the surrounding tissues, thus immediately leading to cellular death of the tissues. We built a simulator able to place the cryoprobes and run a simulation representing the evolution of iceballs in living tissues [58].

5.4.4. Personalized model of the heart for cardiac therapy planning

Participants: Stéphanie Marchesseau [Correspondant], Maxime Sermesant, Hervé Delingette, Nicholas Ayache.

`../../../../projets/asclepios/IMG/TrainingSystem.jpg`

Figure 9. Setup of our simulator dedicated to thermo-ablation for cardiac electrophysiology.

This work is performed in the context of the PhD of Stéphanie Marchesseau in collaboration with St Thomas Hospital in London and was partially funded by ERC MedYMA.

- Personalization of the mechanical function of the heart from time series of cardiac images has been achieved by combining global calibration of a few global parameters [18] with estimation of regional contractility parameters [17] using data assimilation techniques.
- Personalized cardiac models were used to create synthetic images [22] of cardiac motion thus allowing the benchmarking of motion tracking algorithms [8], [39].

5.4.5. Cardiac Arrhythmia Radio-frequency Ablation Planning

Participants: Rocio Cabrera Lozoya [Correspondant], Maxime Sermesant, Hervé Delingette, Nicholas Ayache.

This work is performed in the context of the PhD of Rocío Cabrera Lozoya in collaboration with the CHU LIRYC Bordeaux and is funded by ERC MedYMA.

- Biophysical model development for the prediction of radio frequency ablation sites for ventricular tachycardias.
- Target site map generation for ablation therapy guidance
- Structural and functional characterization of target sites using 3D imaging and EP measurements through machine learning algorithms.
- Prediction validation with acquired clinical data

5.4.6. Computational modeling of radiofrequency ablation for the planning and guidance of abdominal tumor treatment

Participants: Chloé Audigier [Correspondant], Hervé Delingette, Tommaso Mansi, Nicholas Ayache.

This PhD is carried out jointly between the Asclepios research group, Inria Sophia Antipolis, France and the Image Analytics and Informatics global field, Siemens Corporate Research, Princeton, USA.

Radio Frequency Abation, Patient-Specific Simulation, Lattice Boltzmann Method, Computational Fluid Dynamics, Heat Transfer, Therapy Planning, Liver

- In order to obtain a computational framework for patient-specific planning of radiofrequency ablation, a patient-specific detailed anatomical model of the liver has been extracted from a standard CT image and then meshed with tetrahedra. The structures of interest include : parenchyma, lesion, hepatic vein and vena cava.
- A computational fluid dynamic model is used to estimate the patient-specific blood flow in the hepatic circulatory system. It was combined with a porous media model to compute the patient-specific blood flow distribution inside the parenchyma using the results of the CFD solver (pressures).
- Bio-heat equations and a cell death model to account for cellular necrosis have been implemented with FEM using SOFA and a Lattice Boltmann Model to model heat propagation in biological tissues [35] leading to improved accuracy and computational efficiency.

5.4.7. Tumor growth assessment based on biophysical modeling

Participants: Erin Stretton [Correspondant], Bjoern Menze, Nicholas Ayache, Hervé Delingette.

This work was carried out during the Phd of Erin Stretton and was funded by the Care4Me project. It was performed in collaboration with Pr Mandonnet, Lariboisière hospital in Paris, and the German Cancer Research Center (DKFZ)

Glioma simulation, tumor growth.

We aim at developing image analysis methods [23] using tumor growth models in order to guide the planning of therapies (surgical removal and chemotherapy) for brain cancer (glioma) patients. Our work is focused on these objectives :

- Predicting the location of glioma recurrence after a resection surgery;
- Determining the description the of tumor cell diffusion tensor in white matter (patient-based, atlas based or isotropic) which leads to the most accurate results for predicting future tumor growth [57];
- Comparing tumor growth speeds between 8 patient cases based on biophysical modeling and various manual methods.

5.4.8. Brain tumor growth modeling : Application to radiation therapy

Participants: Matthieu Lê [Correspondant], Jan Unkelbach, Nicholas Ayache, Hervé Delingette.

This work is carried out between Asclepios research group, and the Department of Radiation Oncology of the Massachusetts General Hospital, Boston, USA. Part of this work was funded by the European Research Council through the ERC Advanced Grant MedYMA.

Glioma simulations, radiation therapy, target delineation, vasogenic edema

- We developed a tumor growth model for high grade gliomas, based on different types of cell and the vascularization of the brain.
- We studied multimodal brain tumor images to evaluate tumor infiltration.
- We used a Fisher-Kolmogorov model to improve target volume delineation for radiation therapy (see Figure 11)

5.4.9. Multimodal patch-based glioma segmentation

Participants: Nicolas Cordier [Correspondant], Bjoern Menze, Hervé Delingette, Nicholas Ayache.

Part of this work was funded by the European Research Council through the ERC Advanced Grant MedYMA (on Biophysical Modeling and Analysis of Dynamic Medical Images).

../../../../projets/asclepios/IMG/pipelinebis.png

Figure 10. Steps of the proposed method (blue: input, green: processes, purple: output).

`../../../../projets/asclepios/IMG/dose_distribution.png`

Figure 11. Comparison of the dose distribution (in Grey) clinically delivered and based on the Fisher-Kolmogorov model.

Brain, MRI, Glioma, Patch-based Segmentation, Tumor Simulation

- A patch-based approach was developed for glioma segmentation based on multi-channel 3D MRI. The method is fully automatic and does not require any learning step.
- Features: multi-channel MR intensities in local neighborhoods.
- A heuristic label fusion strategy was introduced and showed promising results, as shown in Figure 12 .
- The algorithm was ranked 5th in the Brain Tumor Segmentation Challenge (BraTS) at MICCAI 2013 [67].
- Large unlabeled glioma MRI databases are being incorporated in the framework.

`../../../../projets/asclepios/IMG/CordierFig3.png`

Figure 12. Real high-grade case. From left to right: Vote maps for background, necrosis and non-enhancing tumor (merged), edema, enhancing tumor; Segmentation map. From top to bottom: saggital, axial, and coronal views.

ATHENA Project-Team

6. New Results

6.1. Improving Diffusion MRI Signal and Acquisition

6.1.1. Design of multishell sampling schemes with uniform coverage in diffusion MRI

Participants: Emmanuel Caruyer [SBIA, University of Pennsylvania Medical School,USA], Christophe Lenglet [CMRR, Department of Radiology, University of Minnesota,USA], Guillermo Sapiro [Electrical & Computer Engineering Dept, Duke University,USA], Rachid Deriche.

In diffusion MRI, a technique known as diffusion spectrum imaging reconstructs the propagator with a discrete Fourier transform, from a Cartesian sampling of the diffusion signal. Alternatively, it is possible to directly reconstruct the orientation distribution function in q-ball imaging, providing so-called high angular resolution diffusion imaging. In between these two techniques, acquisitions on several spheres in q-space offer an interesting trade-off between the angular resolution and the radial information gathered in diffusion MRI. A careful design is central in the success of multishell acquisition and reconstruction techniques.

The design of acquisition in multishell is still an open and active field of research, however. In this work, we provide a general method to design multishell acquisition with uniform angular coverage. This method is based on a generalization of electrostatic repulsion to multishell.

The impact of our method on the angular resolution in one and two bundles of fiber configurations is evaluated using simulations. Compared to more commonly used radial sampling, we show that our method improves the angular resolution, as well as fiber crossing discrimination.

This work has been published in [14].

6.1.2. Motion detection in diffusion MRI via online ODF estimation

Participants: Emmanuel Caruyer [SBIA, University of Pennsylvania Medical School,USA], Iman Aganj [Martinos Center for Biomedical Imaging, MGH, Harvard Medical School,USA], Christophe Lenglet [CMRR, Department of Radiology, University of Minnesota,USA], Guillermo Sapiro [Electrical & Computer Engineering Dept, Duke University,USA], Rachid Deriche.

The acquisition of high angular resolution diffusion MRI is particularly long and subject motion can become an issue. The orientation distribution function (ODF) can be reconstructed online incrementally from diffusion-weighted MRI with a Kalman filtering framework. This online reconstruction provides real-time feedback throughout the acquisition process. In this work, the Kalman filter is first adapted to the reconstruction of the ODF in constant solid angle. Then, a method called STAR (STatistical Analysis of Residuals) is presented and applied to the online detection of motion in high angular resolution diffusion images. Compared to existing techniques, this method is image based and is built on top of a Kalman filter. Therefore, it introduces no additional scan time and does not require additional hardware. The performance of STAR is tested on simulated and real data and compared to the classical generalized likelihood ratio test. Successful detection of small motion is reported (rotation under 2 degrees) with no delay and robustness to noise.

This work has been published in [13].

6.1.3. A Robust variational approach for simultaneous smoothing and estimation of DTI

Participants: Rachid Deriche, Meizhu Liu [Department of CISE, University of Florida, Gainesville, USA], Baba C. Vemuri [Department of CISE, University of Florida, Gainesville, USA].

Estimating diffusion tensors is an essential step in many applications — such as diffusion tensor image (DTI) registration, segmentation and fiber tractography. Most of the methods proposed in the literature for this task are not simultaneously statistically robust and feature preserving techniques. In this work, we propose a novel and robust variational framework for simultaneous smoothing and estimation of diffusion tensors from diffusion MRI. Our variational principle makes use of a recently introduced total Kullback–Leibler (tKL) divergence for DTI regularization. tKL is a statistically robust dissimilarity measure for diffusion tensors, and regularization by using tKL ensures the symmetric positive definiteness of tensors automatically. Further, the regularization is weighted by a non-local factor adapted from the conventional non-local means filters. Finally, for the data fidelity, we use the nonlinear least-squares term derived from the Stejskal–Tanner model. We present experimental results depicting the positive performance of our method in comparison to competing methods on synthetic and real data examples.

This work has been published in [20].

6.1.4. Tensor estimation and visualization using dMRI

Participants: Dalila Cherifi [University of Boumerdes, Algeria], Ali Chellouche [University of Boumerdes, Algeria], Amazigh Ait-Ouakli [University of Boumerdes, Algeria], Youcef Benamara [University of Boumerdes, Algeria], Rachid Deriche.

Diffusion tensor imaging in a non-invasive in vivo image modality that allows us to measure molecular diffusion of water in tissues. We characterize diffusion transport of water by an effective diffusion tensor D. The practical importance of the effective diffusion tensor is that it contains new and useful structural and physiological informations about tissues that were previously unobtainable. In this work, we present a software implementation of the estimation of these tensors and their visualization in order to extract these informations.

This work has been published in [28]

6.2. Modeling in Diffusion MRI

6.2.1. A computational diffusion MRI and parametric dictionary learning framework for modeling the diffusion signal and its features

Participants: Sylvain Merlet, Emmanuel Caruyer [SBIA, University of Pennsylvania Medical School,USA], Aurobrata Ghosh, Rachid Deriche.

In this work, we first propose an original and efficient computational framework to model continuous diffusion MRI (dMRI) signals and analytically recover important diffusion features such as the Ensemble Average Propagator (EAP) and the Orientation Distribution Function (ODF). Then, we develop an efficient parametric dictionary learning algorithm and exploit the sparse property of a well-designed dictionary to recover the diffusion signal and its features with a reduced number of measurements. The properties and potentials of the technique are demonstrated using various simulations on synthetic data and on human brain data acquired from 7T and 3T scanners. It is shown that the technique can clearly recover the dMRI signal and its features with a much better accuracy compared to state-of-the-art approaches, even with a small and reduced number of measurements. In particular, we can accurately recover the ODF in regions of multiple fiber crossing, which could open new perspectives for some dMRI applications such as fiber tractography.

This work has been published in Medical Image Analysis [21]. It is part of Merlet's PhD thesis defended on Sept. 11th, 2013 [11].

6.2.2. Continuous diffusion signal, EAP and ODF estimation via compressive sensing in diffusion MRI

Participants: Sylvain Merlet, Rachid Deriche.

In this work, we exploit the ability of Compressed Sensing (CS) to recover the whole 3D Diffusion MRI (dMRI) signal from a limited number of samples while efficiently recovering important diffusion features such as the Ensemble Average Propagator (EAP) and the Orientation Distribution Function (ODF). Some attempts to use CS in estimating diffusion signals have been done recently. However, this was mainly an experimental insight of CS capabilities in dMRI and the CS theory has not been fully exploited. In this work, we also propose to study the impact of the sparsity, the incoherence and the RIP property on the reconstruction of diffusion signals. We show that an efficient use of the CS theory enables to drastically reduce the number of measurements commonly used in dMRI acquisitions. Only 20–30 measurements, optimally spread on several b-value shells, are shown to be necessary, which is less than previous attempts to recover the diffusion signal using CS. This opens an attractive perspective to measure the diffusion signals in white matter within a reduced acquisition time and shows that CS holds great promise and opens new and exciting perspectives in diffusion MRI (dMRI).

This work has been published in Medical Image Analysis [22]. It is part of Merlet's PhD thesis defended on Sept. 11th, 2013 [11].

6.2.3. Constrained diffusion kurtosis imaging using ternary quartics & MLE

Participants: Aurobrata Ghosh, Tristan Milne, Rachid Deriche.

Diffusion kurtosis imaging (DKI) is a recent improvement over diffusion tensor imaging that characterizes tissue by quantifying non-gaussian diffusion using a 3D fourth-order kurtosis tensor. DKI needs to consider three constraints to be physically relevant. Further, it can be improved by considering the Rician signal noise model. A DKI estimation method is proposed that considers all three constraints correctly, accounts for the signal noise and incorporates efficient gradient-based optimization to improve over existing methods.

In this work, the ternary quartic parameterization is utilized to elegantly impose the positivity of the kurtosis tensor implicitly. Sequential quadratic programming with analytical gradients is employed to solve nonlinear constrained optimization efficiently. Finally, a maximum likelihood estimator based on Rician distribution is considered to account for signal noise.

Extensive experiments conducted on synthetic data verify a MATLAB implementation by showing dramatically improved performance in terms of estimation time and quality. Experiments on in vivo cerebral data confirm that in practice the proposed method can obtain improved results.

This work has been published in [18].

6.2.4. Compressive Sensing DSI

Participants: Sylvain Merlet, Michael Paquette [Sherbrooke Connectivity Imaging Laboratory, Computer Science Departement, Université de Sherbrooke, Québec, Canada], Maxime Descoteaux [Sherbrooke Connectivity Imaging Laboratory, Computer Science Departement, Université de Sherbrooke, Québec, Canada], Rachid Deriche.

Compressive Sensing (CS) offers an efficient way to decrease the number of measurements required in Diffusion Spectrum Imaging (DSI). This method aims to reconstruct the Ensemble Average Propagator (EAP) and, for the purpose of this contest, we compute the numerical Orientation Distribution Function (ODF) by integrating the EAP over a solid angle. In this abstract, we briefly describe three important points underlying the CS technique in order to accelerate DSI, namely the sparsity, the Restricted Isometry Property (RIP) and the L1 reconstruction scheme. Due to the high b-values required in the sampling protocol, our approach enters the heavyweight sampling category. Nevertheless, only 64 measurements are used for the reconstruction.

This work has been published in [31]. It is part of Merlet's PhD thesis defended on Sept. 11th, 2013 [11].

6.2.5. 4th Order symmetric tensors and positive ADC modelling

Participants: Aurobrata Ghosh, Rachid Deriche.

High Order Cartesian Tensors (HOTs) were introduced in Generalized DTI (GDTI) to overcome the limitations of DTI. HOTs can model the apparent diffusion coefficient (ADC) with greater accuracy than DTI in regions with fiber heterogeneity. Although GDTI HOTs were designed to model positive diffusion, the straightforward least square (LS) estimation of HOTs doesn't guarantee positivity. In this work, we address the problem of estimating 4th order tensors with positive diffusion profiles. Two known methods exist that broach this problem, namely a Riemannian approach based on the algebra of 4th order tensors, and a polynomial approach based on Hilbert's theorem on non-negative ternary quartics. In this work, we review the technicalities of these two approaches, compare them theoretically to show their pros and cons, and compare them against the Euclidean LS estimation on synthetic, phantom and real data to motivate the relevance of the positive diffusion profile constraint.

This work has been published in [37].

6.2.6. Higher-Order tensors in diffusion imaging: A survey

Participants: Thomas Schultz [MPI for Intelligent Systems, Tubingen, Germany], Andrea Fuster [Eindhoven University of Technology, The Netherlands], Aurobrata Ghosh, Luc Florack [Eindhoven University of Technology, The Netherlands], Rachid Deriche, Lek-Heng Lim [University of Chicago, USA].

Diffusion imaging is a noninvasive tool for probing the microstructure of fibrous nerve and muscle tissue. Higher-order tensors provide a powerful mathematical language to model and analyze the large and complex data that is generated by its modern variants such as High Angular Resolution Diffusion Imaging (HARDI) or Diffusional Kurtosis Imaging. This survey gives a careful introduction to the foundations of higher-order tensor algebra, and explains how some concepts from linear algebra generalize to the higher-order case. From the application side, it reviews a variety of distinct higher-order tensor models that arise in the context of diffusion imaging, such as higher-order diffusion tensors, q-ball or fiber Orientation Distribution Functions (ODFs), and fourth-order covariance and kurtosis tensors. By bridging the gap between mathematical foundations and application, it provides an introduction that is suitable for practitioners and applied mathematicians alike, and propels the field by stimulating further exchange between the two.

This work has been published in [39].

6.2.7. Regularized spherical polar fourier diffusion MRI with optimal dictionary learning

Participants: Jian Cheng [University of North Carolina at Chapel Hill,USA], Tianzi Jiang [LIAMA, China], Rachid Deriche, Shen Dinggang [University of North Carolina at Chapel Hill,USA], Yap Pew-Thian [University of North Carolina at Chapel Hill,USA].

One important problem in diffusion MRI (dMRI) is to recover the diffusion weighted signal from only a limited number of samples in q-space. An ideal framework for solving this problem is Compressed Sensing (CS), which takes advantage of the signal's sparseness or compressibility, allowing the entire signal to be reconstructed from relatively few measurements. CS theory requires a suitable dictionary that sparsely represents the signal. To date in dMRI there are two kinds of Dictionary Learning (DL) methods: 1) discrete representation based DL (DR-DL), and 2) continuous representation based DL (CR-DL). Due to the discretization in q-space, DR-DL suffers from the numerical errors in interpolation and regridding. By considering a continuous representation using Spherical Polar Fourier (SPF) basis, this work proposes a novel CR-DL based Spherical Polar Fourier Imaging, called DL-SPFI, to recover the diffusion signal as well as the Ensemble Average Propagator (EAP) in continuous 3D space with closed form. DL-SPFI learns an optimal dictionary from the space of Gaussian diffusion signals. Then the learned dictionary is adaptively applied for different voxels in a weighted LASSO framework to robustly recover the diffusion signal and the EAP. Compared with the start-of-the-art CR-DL method by Merlet et al. and DRDL by Bilgic et al., DL-SPFI has several advantages. First, the learned dictionary, which is proved to be optimal in the space of Gaussian diffusion signal, can be applied adaptively for different voxels. To our knowledge, this is the first work to learn a voxel-adaptive dictionary. The importance of this will be shown theoretically and empirically in the context of EAP estimation. Second, based on the theoretical analysis of SPF basis, we devise an efficient learning process in a small subspace of SPF coefficients, not directly in q-space as done by Merlet et al.. Third, DL-SPFI also devises different regularization for different atoms in the learned dictionary for robust estimation, by considering the structural prior in the space of signal exemplars. We evaluate DL-SPFI in comparison to L1-norm regularized SPFI (L1-SPFI) with fixed SPF basis, and the DR-DL by Bilgic et al. The experiments on synthetic data and real data demonstrate that the learned dictionary is sparser than SPF basis and yields lower reconstruction error than Bilgic's method, even though only simple synthetic Gaussian signals were used for training in DL-SPFI in contrast to real data used by Bilgic et al.

This work has been published in [27].

6.2.8. Fiber orientation distribution from non-negative sparse recovery

Participants: Aurobrata Ghosh, Thinhinane Megherbi [USTHB, Algeria], Linda Oulebsir-Boumghar [USTHB, Algeria], Rachid Deriche.

We revisit the theory of spherical deconvolution and propose a new fiber orientation distribution (FOD) model that can efficiently reconstruct extremely narrow fiber-crossings from limited number of acquisitions. First, we show how to physically model fiber-orientations as rank-1 tensors. Then, we parameterize the FODs with tensors that are decomposable into non-negative sums of rank-1 tensors and finally, we propose a non-negative sparse recovery scheme to estimate FODs of any tensor order from limited acquisitions. Our method features three important advantages: (1) it estimates non-negative FODs, (2) it estimates the number of fiber-compartments, which need not be predefined and (3) it computes the fiber-directions directly, rendering maxima detection superfluous. We test for various SNRs on synthetic, phantom and real data and find our method accurate and robust to signal-noise: fibers crossing up to 23° are recovered from just 21 acquisitions. This opens new and exciting perspectives in diffusion MRI (dMRI), where our improved characterization of the FOD can be of great help for applications such as tractography.

This work has been published in [29].

6.2.9. A polynomial approach for extracting the extrema of a spherical function and its application in diffusion MRI

Participants: Aurobrata Ghosh, Elias Tsigaridas [PolSys Project-Team, Inria, Paris Rocquencourt, France], Bernard Mourrain [Galaad Project-Team, Inria, Sophia Antipolis, Méditerranée, France], Rachid Deriche.

This work has been partially supported by the ANR project NucleiPark and the France-Parkinson Association.

Antipodally symmetric spherical functions play a pivotal role in diffusion MRI in representing sub-voxel-resolution microstructural information of the underlying tissue. This information is described by the geometry of the spherical function. In this work we propose a method to automatically compute all the extrema of a spherical function. We then classify the extrema as maxima, minima and saddle-points to identify the maxima. We take advantage of the fact that a spherical function can be described equivalently in the spherical harmonic (SH) basis, in the symmetric tensor (ST) basis constrained to the sphere, and in the homogeneous polynomial (HP) basis constrained to the sphere. We extract the extrema of the spherical function by computing the stationary points of its constrained HP representation. Instead of using traditional optimization approaches, which are inherently local and require exhaustive search or re-initializations to locate multiple extrema, we use a novel polynomial system solver which analytically brackets all the extrema and refines them numerically, thus missing none and achieving high precision.

To illustrate our approach we consider the Orientation Distribution Function (ODF). In diffusion MRI the ODF is a spherical function which represents a state-of-the-art reconstruction algorithm whose maxima are aligned with the dominant fiber bundles. It is, therefore, vital to correctly compute these maxima to detect the fiber bundle directions. To demonstrate the potential of the proposed polynomial approach we compute the extrema of the ODF to extract all its maxima. This polynomial approach is, however, not dependent on the ODF and the framework presented in this work can be applied to any spherical function described in either the SH basis, ST basis or the HP basis.

This work has been published in [19].

6.2.10. *ODF maxima computation using hill climbing algorithm*

Participants: Thinhinane Megherbi [USTHB, Algeria], Makhlouf Laouchedi [EMP, Algeria], Houssem Khabatti [EMP, Algeria], Linda Oulebsir-Boumghar [USTHB, Algeria], Ishak Serrat [EMP, Algeria], Vincent Perlbarg [LIF, UMRS 678, INSERM, UPMC - Paris 6], Rachid Deriche.

Diffusion MRI (dMRI) is the only technique to probe in-vivo and non-invasively fiber structure of white matter. Diffusion was first modeled using the classical Second Order Diffusion Tensor model. However, this model is limited in regions of multiple fiber crossings and this has motivated the development of many approaches to extract crossing fibers. Methods like Diffusion Spectrum Imaging (DSI), High Angular Resolution Diffusion Imaging (HARDI) and the High Order Tensor techniques have been proposed to reconstruct specific functions like the Orientation Distribution Function (ODF) whose maxima do correspond to the directions of the multiple fibers.

In this work, we are interested to extract all the crossing fibers characterized as the maxima of the Orientation Distribution Function (ODF). A Hill Climbing algorithm based approach has been developed and implemented to efficiently and accurately extract all the fibers. Promising experimental results obtained with synthetic and real data illustrate the potential of the technique.

This work has been submitted to ISBI'2014 and accepted for presentation and publication.

6.2.11. *On SHORE and SPF bases*

Participants: Elodie Pozzi, Gonzalo Sanguinetti, Rachid Deriche.

The 3D Simple Harmonic Oscillation Reconstruction and Estimation (SHORE) basis and the Spherical Polar Fourier (SPF) basis were introduced recently to represent the dMRI signal in the full 3D Q-space. SPF presents some continuity problems at the origin which led to our development of the modified SPF basis we introduced to overcome this issue. These bases can be written with radial and angular functions. The radial part of the decomposition is a family of orthogonal functions (the Gauss-Laguerre functions) and the angular component are the spherical harmonic functions. Even though they look similar, they have different properties. The first objective of this work has been to analyse and clarify the differences between those bases. This has been accomplished by describing the spanned spaces. The second goal has been to classify the bases according to their continuity and differentiability and thus draw a more focused comparison between.

This on-going work will be submitted to a journal.

6.3. From DW-MRI to Fiber Pathways and Microstructures Recovery

6.3.1. *Mapping Average axon diameters under long diffusion time*

Participants: Gonzalo Sanguinetti, Rachid Deriche.

This work proposes an original method to recover axon diameter distribution (ADD) parameters using nuclear magnetic resonance. White matter (WM) is modelled as a bi-compartmental medium composed of an intra axonal space where the diffusion is restricted and an extra axonal space where diffusion is hindered. Under the assumption of long diffusion time, we provide a novel and efficient model for the component of the signal due to the restricted part. This technique might be interpreted as an interesting simplification of the AxCaliber framework, which leads to a simpler model and an extremely faster acquisition protocol. To test and validate our method, we use the open-source toolkit Camino for computing Monte-Carlo simulations of NMR data and model the WM as 3D cubic environments, formed by parallel cylinders with gamma distributed radii. Promising experimental results illustrate the potential of the proposed method.

This work has been submitted to ISBI'2014 and accepted for presentation and publication.

6.3.2. *NMR characterization of cylinder radii distribution using a SHORE-based regularization method.*

Participants: Gonzalo Sanguinetti, Daniel Alexander [Centre for Medical Image Computing, Dept. Computer Science, UCL], Matt Hall [Centre for Medical Image Computing, Dept. Computer Science, UCL], Rachid Deriche.

In this work, we extend the framework presented by Ozarslan et al [79] by adding a regularization term for better measuring the moments of a cylinder radii distribution by means of NMR acquisitions. The added value of the regularization term is tested and validated using Monte Carlo simulations of NMR signals from complex white matter-like environment. The open source toolkit CAMINO [50] is used for computing the simulations and an excellent agreement is obtained between the ground truth and the estimated moments.

This work has been submitted to ISMRM'2014.

6.3.3. Quantitative comparison of reconstruction methods for intra-voxel fiber recovery from diffusion MRI

Participants: Emmanuel Caruyer [SBIA, University of Pennsylvania Medical School, USA], Sylvain Merlet, Rachid Deriche.

In diffusion MRI, a technique known as diffusion spectrum imaging reconstructs the propagator with a discrete Fourier transform, from a Cartesian sampling of the diffusion signal. Alternatively, it is possible to directly reconstruct the orientation distribution function in q-ball imaging, providing so-called high angular resolution diffusion imaging. In between these two techniques, acquisitions on several spheres in q-space offer an interesting trade-off between the angular resolution and the radial information gathered in diffusion MRI. A careful design is central in the success of multishell acquisition and reconstruction techniques and the design of acquisition in multishell is still an open and active field of research.

In this work, we propose a novel method to design sampling schemes with optimal angular coverage and show the positive impact on angular resolution in diffusion MRI. Our method is based on a generalization of electrostatic repulsion to multishell and allows to design multishell acquisition with uniform angular coverage.

We evaluated the impact of our method using simulations, on the angular resolution in one and two bundles of fiber configurations. Compared to more commonly used radial sampling, we show that our method improves the angular resolution, as well as fiber crossing discrimination.

This work has been published in [16].

6.3.4. Choosing tractography parameters to improve connectivity mapping

Participants: Gabriel Girard [SCIL Lab., Sherbrooke University], Kevin Whittingstall [SCIL Lab., Sherbrooke University], Kevin Whittingstall [SCIL Lab., Sherbrooke University], Rachid Deriche.

Diffusion-weighted imaging (DWI) is often used as a starting point for in vivo white matter (WM) connectivity to reconstruct potential WM pathways between brain areas. Tractography algorithms have many parameters which can influence reconstruction and connectivity. Various choices of parameters have been proposed. But how does one choose the best set of parameters ? In this study, we varied three critical parameters while monitoring connectivity score using the Tractometer evaluation system on the International Symposium on Biomedical Imaging (ISBI) Challenge synthetic dataset. The three parameters were: The maximum deviation angle between two consecutive tractography steps (this addresses the hypothesis of smoothness of the WM pathways), the spherical function (SF) threshold (this aims at removing noisy propagation directions during the tractography process) and the initial SF threshold (this aims at removing initial noise at the seeds and to start tractography in a good tangent direction to the WM bundle).

This work has been submitted to ISMRM'2014.

6.3.5. Improved tractography using structural priors

Participants: Gabriel Girard [SCIL Lab., Sherbrooke University], Maxime Descoteaux [SCIL Lab., Sherbrooke University], Kevin Whittingstall [SCIL Lab., Sherbrooke University], Rachid Deriche.

In this work, we propose better tractography parameters in term of global connectivity and a novel tractography stopping criterion based on partial volume estimation maps, calculated from a T1-weighted image. We also propose a particle filtering method using anatomical information as prior for tractography to enforce streamlines connecting gray matter regions and reducing the proportion of erroneous streamlines. Results show streamlines more uniformly distributed among long and short, and small and large white matter bundles. This provides connectivity estimation not underestimated for bundles having higher complexity. Quantitative analysis is done on synthetic datasets and qualitative results are shown on real data. The proposed method takes advantage of prior information on the brain to change the dMRI-based tracking direction and help providing streamlines that can quantify the brain structure.

This on-going work will be submitted to NeuroImage.

6.3.6. From diffusion MRI to brain connectomics

Participants: Aurobrata Ghosh, Rachid Deriche.

Diffusion MRI (dMRI) is a unique modality of MRI which allows one to indirectly examine the microstructure and integrity of the cerebral white matter in vivo and non-invasively. Its success lies in its capacity to reconstruct the axonal connectivity of the neurons, albeit at a coarser resolution, without having to operate on the patient, which can cause radical alterations to the patient's cognition. Thus dMRI is beginning to assume a central role in studying and diagnosing important pathologies of the cerebral white matter, such as Alzheimer's and Parkinson's diseases, as well as in studying its physical structure in vivo. In this work, we present an overview of the mathematical tools that form the framework of dMRI – from modelling the MRI signal and measuring diffusion properties, to reconstructing the axonal connectivity of the cerebral white matter, i.e., from Diffusion Weighted Images (DWIs) to the human connectome.

This work has been published in [38].

6.4. Forward and Inverse Problems in MEEG

6.4.1. Source localization using rational approximation on plane sections

Participants: Maureen Clerc [Athena Project-Team, Inria, Sophia Antipolis, Méditerranée, France], Théodore Papadopoulo [Athena Project-Team, Inria, Sophia Antipolis, Méditerranée, France], Juliette Leblond [Apics Project-Team, Inria, Sophia Antipolis, Méditerranée, France], Jean-Paul Marmorat [CMA, Ecole des Mines Paristech, Sophia Antipolis, France].

In functional neuroimaging, a crucial problem is to localize active sources within the brain non-invasively, from knowledge of electromagnetic measurements outside the head. Identification of point sources from boundary measurements is an ill-posed inverse problem. In the case of electroencephalography (EEG), measurements are only available at electrode positions, the number of sources is not known in advance and the medium within the head is inhomogeneous. We have presented [49] a new method for EEG source localization, based on rational approximation techniques in the complex plane. The method is used in the context of a nested sphere head model, in combination with a cortical mapping procedure. Results on simulated data prove the applicability of the method in the context of realistic measurement configurations. In the continuation of this work, we are in discussion with an industrial partner (BESA, Munich) for a scientific partnership.

6.4.2. Dictionary learning for multitrial datasets

Participants: Maureen Clerc, Sebastian Hitziger, Théodore Papadopoulo.

Following the path opened with the Consensus matching Pursuit method (CMP) [46], we continue our endeavour to avoid signal averaging using directly the raw signal with the assumption that events of interest are those that repeat in each trial [36]. Towards such a goal, and to improve the simple dictionary used in CMP, we have adapted dictionary learning methods to multitrial bio-electric signals, by explicitly implementing jitter invariance [30]. This allows for a much more detailed data-driven description of events. For example, using local field potential signals of chemically induced spikes (in a rat model), we have been able to distinguish several spike shapes which show some coherence in time. The method has been recently extended to detect spike events in continuous signals (i.e. not organized in epochs). While it requires a good signal to noise ratio, the method is very general and has also been used for various other signal types (see section 6.5).

6.5. Coupling functional and structural models

6.5.1. Cortex parcellation via diffusion data as prior knowledge for the MEG inverse problem

Participants: Anne-Charlotte Philippe, Maureen Clerc, Théodore Papadopoulo, Rachid Deriche.

In this work, a new approach is presented for the recovery of dipole magnitudes in a distributed source model for magnetoencephalographic (MEG) imaging. This method consists in introducing prior knowledge regarding the anatomical connectivity in the brain to this ill-posed inverse problem. Towards this goal, a cortex parcellation is performed using structural information coming from diffusion MRI (dMRI), the only non-invasive modality allowing to have access to the structure of the WM tissues. Then, sources in the same diffusion parcel are constrained, in the MEG inverse problem, to have close magnitude values. Results of our method on MEG simulations are presented and favorably compared with classical source reconstruction methods.

This work has been published in [32], and is part of A.C. Philippe's Ph.D thesis [12].

6.5.2. Diffusion-Weighted Imaging tractography-based parcellation of the human cortex as regularization term for the MEG inverse problem

Participants: Anne-Charlotte Philippe, Maureen Clerc, Théodore Papadopoulo, Rachid Deriche.

The purpose of this work is to advocate the use of structural connectivity information to regularize the ill-posed MEG inverse problem. Diffusion MRI being the only non invasive modality allowing to have access to the connectivity profile of cortical sources, the proposed method called Diff-MNE consists in the introduction of a cortex parcellation based on diffusion data regularization term to the MEG inverse problem. Our method is tested on synthetic and real human brain data and compared to the classical minimum-norm method. Results show that a diffusion-based cortex parcellation as a regularization term for the MEG inversion process improves the source reconstruction. This proves the interest of merging diffusion MRI and MEG data.

This work is under submission to a Neuroimage and is part of A.C. Philippe's Ph.D thesis [12]

6.5.3. Propagation of epileptic spikes revealed by diffusion-based constrained MEG source reconstruction

Participants: Anne-Charlotte Philippe, Maureen Clerc, Théodore Papadopoulo, Rachid Deriche.

In this work, we study the propagation of an epileptic spike (from single event data). As in the two previous sections, a cortex parcellation is performed using structural information coming from diffusion MRI Then, a MEG inverse problem is defined on a parcellated source space which imposes constant activity on each parcel. This inverse problem is applied separately for measurements obtained in a given time range. The most active parcels over the time range are located and their time course are displayed. This allowed the study of the propagation of an epileptic spike via those active parcels. Results on real data shows varying spatial propagations of an epileptic spike for the same subject.

This work has been published in [40], and is part of A.C. Philippe's Ph.D thesis [12].

6.6. Brain Computer Interfaces

6.6.1. Combining ERD and ERS features to create a system-paced BCI

Participants: Maureen Clerc, Joan Fruitet, Théodore Papadopoulo, Eoin Thomas.

An important factor in the usability of a brain computer interface (BCI) is the setup and calibration time required for the interface to function accurately. Recently, brain-switches based on the rebound following motor imagery of a single limb effector have been investigated as basic BCIs due to their good performance with limited electrodes, and brief training session requirements. In this work, a BCI is proposed which expands the methodology of brain-switches to design an interface composed of multiple brain-buttons. The algorithm is designed as a system paced interface which can recognise 2 intentional-control tasks and a no-control state based on the activity during and following motor imagery in only 3 electroencephalogram channels. An online experiment was performed over 6 subjects to validate the algorithm, and the results show that a working BCI can be trained from a single calibration session and that the post motor imagery features are both informative and robust over multiple sessions.

This work has been published in [24].

6.6.2. An analysis of performance evaluation for motor-imagery based BCI

Participants: Maureen Clerc, Matthew Dyson [Laboratoire de Neurosciences Cognitives, Marseille], Eoin Thomas.

In recent years, numerous brain–computer interfaces (BCIs) based on motor-imagery have been proposed which incorporate features such as adaptive classification, error detection and correction, fusion with auxiliary signals and shared control capabilities. Due to the added complexity of such algorithms, the evaluation strategy and metrics used for analysis must be carefully chosen to accurately represent the performance of the BCI. In this work, metrics are reviewed and contrasted using both simulated examples and experimental data. Furthermore, a review of the recent literature is presented to determine how BCIs are evaluated, in particular, focusing on the relationship between how the data are used relative to the BCI subcomponent under investigation. From the analysis performed in this study, valuable guidelines are presented regarding the choice of metrics and evaluation strategy dependent upon any chosen BCI paradigm.

This work has been published in [23].

6.6.3. Bandit algorithms for faster task selection in BCI

Participants: Maureen Clerc, Aurélien Emmanuel, Joan Fruitet [former Athena PhD student], Alexandra Carpentier [Sequel Project-Team, Inria Lille], Rémi Munos [Sequel Project-Team, Inria Lille].

Brain–computer interfaces (BCIs) based on sensorimotor rhythms use a variety of motor tasks, such as imagining moving the right or left hand, the feet or the tongue. Finding the tasks that yield best performance, specifically to each user, is a time-consuming preliminary phase to a BCI experiment. This study presents a new adaptive procedure to automatically select (online) the most promising motor task for an asynchronous brain-controlled button. We have developed for this purpose an adaptive algorithm UCB-classif based on the stochastic bandit theory and design an EEG experiment to test our method. We compare (offline) the adaptive algorithm to a naïve selection strategy which uses uniformly distributed samples from each task. We also run the adaptive algorithm online to fully validate the approach. By not wasting time on inefficient tasks, and focusing on the most promising ones, this algorithm results in a faster task selection and a more efficient use of the BCI training session. More precisely, the offline analysis reveals that the use of this algorithm can reduce the time needed to select the most appropriate task by almost half without loss in precision, or alternatively, allow us to investigate twice the number of tasks within a similar time span. Online tests confirm that the method leads to an optimal task selection. This study is the first one to optimize the task selection phase by an adaptive procedure. By increasing the number of tasks that can be tested in a given time span, the proposed method could contribute to reducing ’BCI illiteracy’.

This work has been published in [17].

6.6.4. Enhancing visuospatial attention performance with brain-computer interfaces

Participants: Thomas Brochier [Institut des Neurosciences de La Timone, Marseille], Maureen Clerc, Romain Trachel.

Brain-Computer Interfaces (BCI) can provide innovative solutions beyond the medical domain. In human research, visuospatial attention is often assessed from shifts in head or gaze orientation. However in some critical situations, these behavioral features can be dissociated from covert attention processes and brain features may indicate more reliably the spatial focus of attention. In this context, we investigate whether EEG signals could be used to enhance the behavioral performance of human subjects in a visuospatial attention task. Our results demonstrate that a BCI protocol based on adaptive or warning displays can be developed to shorten the reaction time and improve the accuracy of responses to complex visual targets. We performed offline and online tests demonstrating the validity of this type of approach.

This work was presented at conferences in the HCI community [35] and in the Neural Engineering community [34].

6.6.5. Verbal communication through brain computer interfaces

Participants: Maureen Clerc, Dieter Devlaminck, Claude Desnuelle [CHU de Nice l'Archet], Violaine Guy [CHU de Nice l'Archet], Manu Maby [Centre de Recherche Neurologique de Lyon], Jérémie Mattout [Centre de Recherche Neurologique de Lyon], Théodore Papadopoulo.

Brain Computer Interfaces (BCI) provide a way of communicating directly from brain activity, bypassing muscular control. We report some recent advances in a BCI communication system called the P300 speller, which is a virtual brain-operated keyboard. This system relies on electroencephalographic activity time-locked to the flashing of the desired letters. It requires calibration of the system, but very little training from the user. Clinical tests are being conducted on a target population of patients suffering from Amyotrophic Lateral Sclerosis, in order to confirm the usability of the P300 speller for reliable communication.

This work has been published in [26]. It is also the object of an intensive clinical study on 20 patients which we are currently conducting at Nice University Hospital.

DEMAR Project-Team

6. New Results

6.1. Modelling and Identification

6.1.1. Emergence of Motor Synergy in Reaching Task via Tacit Learning -computational motor control

Participants: Mitsuhiro Hayashibe, Shingo Shimoda [RIKEN,Japan].

The dynamics of multijoint limbs often causes complex dynamic interaction torques which are the inertial effect of other joints motion. It is known that Cerebellum takes important role in a motor learning by developing the internal model. We propose a novel computational control paradigm in vertical reaching task which involves the management of interaction torques and gravitational effect. The obtained results demonstrate that the proposed method is valid for acquiring motor synergy in the system with actuation redundancy and resulted in the energy efficient solutions. It is highlighted that the tacit learning in vertical reaching task can bring computational adaptability and optimality with model-free and cost-function-free approach differently from previous studies.

6.1.2. Anatomy Transfer

Participants: Dicko Ali-Hamadi, Tiantian Liu, Benjamin Gilles, Ladislav Kavan, Francois Faure, Olivier Palombi, Marie-Paule Cani.

Characters with precise internal anatomy are important in film and visual effects, as well as in medical applications. We propose the first semi-automatic method for creating anatomical structures, such as bones, muscles, viscera and fat tissues. This is done by transferring a reference anatomical model from an input template to an arbitrary target character, only defined by its boundary representation (skin). The fat distribution of the target character needs to be specified. We can either infer this information from MRI data, or allow the users to express their creative intent through a new editing tool. The rest of our method runs automatically: it first transfers the bones to the target character, while maintaining their structure as much as possible. The bone layer, along with the target skin eroded using the fat thickness information, are then used to define a volume where we map the internal anatomy of the source model using harmonic (Laplacian) deformation. This way, we are able to quickly generate anatomical models for a large range of target characters, while maintaining anatomical constraints.

6.1.3. Center of Mass Estimation in Multicontact Situations: Simulation

Participants: Alejandro González, Mitsuhiro Hayashibe, Emel Demircan [Stanford Univ.], Philippe Fraisse.

Center of mass (CoM) estimation can be used to evaluate human stability during rehabilitation. A personalized estimation can be obtained using the serial equivalent static chain (SESC) method, calibrated using a series of static postures. The estimation accuracy is dependent on the number and quality of poses used during calibration. Currently, this limits the method's application to unimpaired individuals. We present a preliminary study of a SESC identified in a multi-contact scenario during a Sit-to-Stand task. Stanford's SAI (Simulation and Active Interface) platform was used to emulate human motion and predict relevant reaction forces. The CoM estimation obtained is valid for motions similar to those used during identification. Since the SAI's human model is fully defined, in terms of mass and limb lenths, its exact center of mass is known. Using a 3-dimensional model, the estimated mean error was less than 26 mm for a Sit-to-Stand task involving displacements along all axes. As such, personalized CoM estimation can be available for patients with a limited range of whole body motion.

../../../../projets/demar/IMG/VerticalReachingInria.png

Figure 1. (left)Schematic representation of vertical reaching task. (right)End point transition. (a) only with PD feedback control (b) with tacit learning in addition to the PD control.

`../../../../projets/demar/IMG/anatomytransferTeaser.png`

Figure 2. A reference anatomy (left) is automatically transferred to arbitrary humanoid characters. This is achieved by combining interpolated skin correspondences with anatomical rules.

`../../../../projets/demar/IMG/StanBot.png`

Figure 3. After identifying the SESC parameter of the humanoid (Stanbot), it is possible to estimate the position of its CoM. SESC identification was performed with a reduced number of postures, to mimic a patient in need of additional support to maintain a standing pose.

Fig. 3 shows the identification results [27]. Even with a limited motion, it was possible to estimate the position of the simulated robot's CoM projected to onto the ground. This estimation errors are likely due to the lack of an exciting trajectory for identification. Nonetheless, if the postures used during identification describe a patient's range of motion, the CoM estimation can still be valuable.

6.1.4. Interface for identification of the Statically Equivalent Serial Chain's parameters and Balance Assessment

Participants: Alejandro González, Mitsuhiro Hayashibe, Philippe Fraisse.

CoM trajectory can be used to improve the current rehabilitation standards. After an identification phase, a personalized CoM estimate can be obtained using a SESC. Furthermore, using low-cost sensors (Kinect and Wii balance board), make the personalized estimate feasible inside a patient's home. This work focuses on the effect that a visual adaptive interface can have on the SESC identification phase. Specifically on improving its speed and quality. A study conducted on 6 subjects showed a faster convergence and a lower root mean square error (rmse) when the adaptive interface is applied. We find that for the same error (30 mm), the identification with the interface was performed in half the time (86 s) than the one without it (163 s). Similarly, for the same session length (120 s), rmse was of 24.5 mm using the interface and of 34.5 mm without it.

Additionally CoM dynamics may be used to determine stability. Fig. 4 shows an example of this during a squat task. The zero rate of angular momentum (ZRAM) can be used to determine the dynamic stability of a humanoid robot. It can be used to determine the position foot placement to avoid falls.

6.1.5. Forward Estimation of Joint Torque from EMG Signal through Muscle Synergy Combinations

Participants: Zhan Li, Mitsuhiro Hayashibe, David Guiraud.

We investigate the approaches of estimating the ankle joint torque from EMG/activations of associated muscle groups. The approaches discussed fall into two main categories: i) full utilization of both of extension and flexion EMG/activations for estimating the joint torque; ii) exploitation of muscle synergy extraction of EMG/actvations and consequent usage of extracted components in reduced space for estimating the joint torque. Comparison is made between the two methods with experimental data of five able-bodied subjects. From the results we conclude that, method ii) with muscle synergy extraction may not degrade the performance of method i) but meanwhile show the muscle synergic ratios for generating the joint torque, and involvement of joint position and velocity information can improve the estimation for both methods.

6.1.6. Prediction of hand tremor through EMG-based fatigue tracking

Participants: Sourav Chandra, Mitsuhiro Hayashibe, Thondiyath Asokan [IITMadras, India].

Laparoscopic surgical procedure is a very tiring procedure for a surgeon due to the specialized prolonged arm movement with a modular tool. Prolonged activity of arm muscle in such condition induces muscle fatigue which induces hand tremor. Hand tremor not only drastically affects positional accuracy; it also increases the collateral tissue damage. Nullification of this tremor has been the area of active research topic in surgical robotics for last few decades. Though Surface ElectroMyoGram (SEMG) has been used for modeling hand tremor of microsugeons, a single model for predicting amplitude and frequency of such tremor has not been investigated for laparoscopy so far. A model of muscle fatigue induced hand tremor in laparoscopic activity is necessary in order to nullify this hand tremor effect and increase positional accuracy. SEMG is a crucial biopotential in order to get the estimation of the muscle fatigue state. A positive correlation was found among sEMG and hand tremor in frequency domain as shown in Fig. 6 below. In this work, a model based prediction of fatigue induced hand tremor will be investigated with the vicinity of SEMG and other wearable inertial sensor data. The model is intended to have a dynamic structure, which can capture the complexity of the muscle fatigue state to some extent and its effect on the hand tremor amplitude and frequency.

6.1.7. Mobile gait analysis

Participants: Vincent Bonnet [M2H, UM1, Montpellier], Christine Azevedo Coste, Christian Geny [CHU Montpellier], Lionel Lapierre [LIRMM, Montpellier], René Zapata [LIRMM, Montpellier].

`../../../../projets/demar/IMG/ZRAM.png`

Figure 4. We show the trajectory of the zero rate of angular momentum (ZRAM) for a sqat task. When the ZRAM is found inside the support polygon, the movement can be considered stable. Unstablle movements (C-D) do not determine a fall. The subject may still recover using a balancing strategy; or by taking a step.

../../../../projets/demar/IMG/synergy1.png

Figure 5. Normalized muscle synergy ratios of the five subjects, under isotonic situation with 10Nm plantar load and 5Nm dorsi load

The Video-Kinect-Bot, an affordable mobile platform for pathological gait analysis was developed to assess pathological spatio-temporal parameters. The system, drove by a Kinect sensor, is able to follow a patient at constant distance on his own defined path, and to estimate gait spatio-temporal parameters. Robust Tracking-Learning-Detection algorithm estimates the positions of targets attached to the trunk and heels of the patient. Real-condition experimental validation including corridor, occlusion cases, and illumination change was performed. A gold standard stereophotogrammetric system was also used and shown a good tracking of patient and an accuracy in stride length estimate of 2%.

The Empirical Mode Decomposition (EMD) method was evaluated to estimate the 3D orientation of the lower trunk during walking using the angular velocity signals generated by a wearable inertial measurement unit (IMU) and notably flawed by drift. The IMU was mounted on the lower trunk (L4-L5) with its active axes aligned with the relevant anatomical axes. The proposed method performs an offline analysis but has the advantage of not requiring any parameter tuning. The method was validated in two groups of 15 subjects, one during overground walking, with 180° turnings, and the other during treadmill walking, both for steady-state and transient speeds, using stereophotogrammetric data. Comparative analysis of the results showed that the IMU/EMD method is able to successfully detrend the integrated angular velocities and estimate lateral bending, flexion-extension as well as axial rotations of the lower trunk during walking with RMS errors of 1 deg for straight walking and lower than 2.5 deg for walking with turnings. This work was accepted for publication in Sensors journal for a special issue concerning wearable-sensor for gait analysis in 2014 with the following collaborators V. Bonnet, S. Ramdani, C. Azevedo-Coste, P. Fraisse, C. Mazzà and A. Cappozzo. Data relative to the pitch, roll and yaw angles obtained for one randomly selected treadmill walking trial. The integrated angular velocities (grey line) and the resulting trends (black line) are estimated using EMD (a) during all the trial; zoom over 20 s on the corresponding detrended angles are thereafter estimated (black line) and compared with those obtained using stereophotogrammetry (grey line).

This work is supported by a NOVARTIS funding (see Partnerships and Cooperations section).

../../../../projets/demar/IMG/semg.png

Figure 6. Frequency domain correlation of sEMG and hand tremor during prolonged arm movement

../../../../projets/demar/IMG/VikiBot.png

../../../../projets/demar/IMG/Tracking_snapshot.

*Figure 7. (**Left:**) Detail of the Video-Kinect-Bot system and experimental setup (**Right:**) example of subject tracking in real-environement.*

6.2. Function control and synthesis

6.2.1. Analysis of infection risk in surgery block

Participants: Christine Azevedo Coste, Roger Pissard Gibollet [SED Inria Grenoble Rhône-Alpes], Gabriel Birgand [Bichat Hospital, Paris], Jean-Christophe Lucet [Bichat Hospital, Paris], Gaelle Toupet [Bichat Hospital, Paris].

Despite the increasing implementation of preventive measures, surgical-site infection still induces a substantial burden. Inappropriate staff behaviors can lead to environmental contamination in the operating room and subsequent surgical site infection. The present study focuses on the continued assessment of operating room staff behavior using a motion tracking system, and the evaluation of the impact of this behavior on the surgical-site infection risk during surgical procedures.

A multicenter observational study has been done in 2013, including 10 operating rooms of cardiac and orthopedic surgery in 12 healthcare facilities. A motion tracking system including 8 optical cameras (VICON-Bonita®) recorded movements of reflective markers placed on the surgical caps/hoods of each person entering the room. Different configurations of markers positioning were used to distinguish between staff category. Doors opening were observed as well by means of wireless inertial sensors fixed on the doors and synchronized with the motion tracking system. We have collected information on the operating room staff, surgical procedures and surgical environment characteristics ([2]).

Recorded data will be analyzed and staff behaviors will be assessed by the quantification of displacements within the operating room. Results will aim at bringing a rational to the prevention of airborne microorganism transmission by the description of best behaviors rules in the operating room.

This protocol was approved by the Institutional Review Board of the (IRB) of Paris North Hospitals, Paris 7 University, AP-HP (n° 11-113, April 6 2012). The work is supported by Inria SENSBIO ADT and ARIBO Preqhos project.

6.2.2. Drop-foot correction in post-stroke hemiplegic patients

Participants: Christine Azevedo Coste, Roger Pissard-Gibollet [SED Inria Grenoble Rhône-Alpes], Jérôme Froger [Nîmes Hospital, Le Grau du Roi], Claire Delablachelerie [Nîmes Hospital, Le Grau du Roi].

Electrical stimulation has been proven to have orthotic and carryover effects on individuals with post-stroke hemiplegia with a foot drop syndrome. One of the drawbacks of the technique is the lack of adaptability to changes in gait (speed, type of floor, stairs, dorsiflexion quality etc). But, real-time modification of stimulation patterns is not feasible using gait event detection like proposed in all available stimulators. In the present study we investigate two questions: 1) is it possible to validate on individuals with foot drop an algorithm able to estimate online the continuous gait cycle phase from a unique wireless sensor placed on lower limbs and 2) is it possible to trig a drop foot stimulator based on events extracted from this phase information.

Methods : 20 subjects with post-stroke hemiplegia participated to the study. A wireless inertial measurement unit was placed on the unaffected leg of the subjects and was used to estimate the gait phase on a distant labtop. The subjects performed 3 trials in each of the 3 following conditions: **C1** no stimulation aid, **C2** electrical stimulation assistance triggered by heel switch **C3** electrical stimulation assistance wirelessly triggered based on the proposed algorithm.

Results : 1) the proposed algorithm was able to estimate online the continuous gait cycle phase, 2) events could be extracted from this phase information in order to trig an electrical stimulator using this algorithm instead of heel switch.

Conclusion : the online estimation of continuous gait cycle phase on individuals with stroke is possible. Events can be extracted from the phase information in order to trig a stimulator **C3** instead of using heel switch detection **C2**. The robustness of the proposed solution to gait modifications is intrinsically guaranteed by the use of automatic control theory. These results open promising applications using programmable stimulators which parameters could be modified online based on gait phase observation.

This protocol was approved by Nîmes Ethical Committee, AFSSAPS and CNIL. The work is supported by Inria SENSBIO ADT.

6.2.3. Freezing of Gait detection in Parkinsonian individuals

Participants: Christine Azevedo Coste, Christian Geny [CHU Montpellier], Maud Pasquier [Inria Grenoble Rhône-Alpes], Benoît Sijobert.

Parkinson's disease (PD) is the second most common neurodegenerative disorder. This chronic disease can lead to gait disturbances and falls inducing important reduction of the quality of life. One common symptom is the Freezing of Gait (FOG), an episodic inability to generate effective stepping in the absence of any known cause other than Parkinsonism. It can occur during initiation of the first step, turning, dual task, walking through narrow spaces, reaching destinations or passing through doorways. It is an episodic absence or marked reduction of forward progression of the feet despite the intention to walk. FOG are reported by the patient as a subjective feeling of "the feet being glued to the ground". Clinical evaluation of video recordings of patients by one to three observers is the gold standard to identify FOG events. The evaluation of clinical effects of the treatments would benefit from objective, standardized FOG measures. Moore et al. (2008;2013) have proposed a technique to identify FOG episodes based on the frequency properties of leg vertical accelerations (fig.9). The approach is based on the hypothesis that FOG occurrences are associated to trembling motion, which affect limb acceleration signal. They have introduced the so-called freeze index (FI): the ratio between the signal power in the trembling band (3 Hz - 8 Hz) and the signal power in the locomotor band (0.5 Hz-3 Hz). The FI method was validated using one to 7 accelerometers mounted on patients with satisfactory detection results. If many FOG episodes can be associated to festination (trampling) it is not the case for all of them. Therefore, we claim that all the FOG episodes cannot be detected by the FI method. In the present paper we propose a complementary index in order to take into account not only festination but also other freezing characteristics. Furthermore we intend to propose a solution based on a minimal number of embedded sensors and detection algorithms for future real-time applications.

`../../../../projets/demar/IMG/archi.jpg`

Figure 8. ***System architecture****. Description of the system architecture used in the study. A sensor node (inertial measurement unit (IMU)) is placed on the unaffected side shank. Data is sent to the sink node of the laptop. Data is processed on the laptop and a gait phase is estimated. Depending on the phase value the stimulator is switched ON through its trigger node. An extra sensor node is also sending data to the sink node and data is saved for offline processing.*

`../../../../projets/demar/IMG/freezing.png`

Figure 9. Freezing of Gait observation using inertial sensors.

This work is supported by SENSBIO Inria ADT and DEMARPARK AOI (see Partnerships and Cooperations section).

6.2.4. Effects of direct electrical stimulation of the brain during awake surgeries: towards improvements of the functional mapping

Participants: François Bonnetblanc, David Guiraud, Marion Vincent, Mitsuhiro Hayashibe, Hugues Duffau [Neurosurgery department, CHU-Gui de Chauliac], Guillaume Herbet [Neurosurgery department, CHU-Gui de Chauliac], Benedicte Poulin-Charronnat [LEAD, Univ. Dijon].

« Awake surgery » consists in removing some infiltrative and slow-growing brain tumoral tissu in an awake patient. The neurosurgeon performs an anatomo-functional mapping of the brain by electrically stimulating brain areas near the tumor to discriminate functional vs. non functional areas. This stimulation is both made cortically and sub-cortically to preserve the functional connectivity. During the surgery itself, the patients are also involved by performing some tasks. Their recovery remains impressive with respect to the lesional volume. Despite the slow-growth of the lesion is invoked, these observations question our understanding of brain plasticity phenomena. Our multi-disciplinary approach aims to (i) better understand the effects of direct electrical stimulation of the brain to improve the functional mapping and also (ii) to build new functional assessments performed by the patient and based on new technologies applied to Health. By systematically performing these precise assessments before, after and during the surgery we hope to better understand brain functions, plasticity and dynamics in order to improve the surgical planning, functional mapping, rehabilitation procedures and quality of life of the patients.

6.2.5. Translational research and stroke

Participants: Anirban Dutta, David Guiraud.

Stroke is caused when an artery carrying blood from heart to an area in the brain bursts or a clot obstructs the blood flow thereby preventing delivery of oxygen and nutrients. About half of the stroke survivors are left with some degree of disability where the impairment of walking has been mentioned most frequently as the most important disability. There is, therefore, a pressing need to leverage insights from animal and human studies to address the complexity in clinical translation of rational multi-level electrotherapy protocols where the ability to customize such novel electrotherapy protocols has only recently become possible with advanced computational tools. Therefore the challenge is to develop advanced computational modeling tools at Inria, France, to design and customize innovative electrotherapy protocols to patient-specific needs, and then closely integrate them to drive (perhaps the first) individualized non-invasive electrotherapy program for clinical validation. The ongoing steps are i) Develop computational methods to identifying neural circuits related to the recovery from stroke [26], ii) Develop a computational method for online targeting of neural circuits and related pathways with non-invasive electrotherapy [21],[25], iii) Validate individualized multi-level non-invasive electrotherapy program with NIBS as an adjuvant treatment to NMES-assisted gait rehabilitation following stroke.

6.2.6. Projet PERIMED

Participants: Thomas Guiho, Christine Azevedo, Luc Bauchet, Charles Fattal, David Guiraud, Jean-Rodolphe Vignes.

Born in the 70's, Spinal cord stimulation is a general term including both peridural and intradural stimulation. Encouraged by Harkema's clinical result (in one paraplegic patient with step-like EMG activity) [43], several recent studies in rodents elicited locomotor synergies, bladder/bowel improvements and, in certain circumstances, restoration of supraspinal control after spinal cord injury [45]. Based on this previous work, our approach, mainly focused on bladder and bowel functions, aims both at asserting these discoveries in an intermediate model (pigs weighing between 50 and 60 kgs) and at providing further insight in spinal cord circuitries.

6.2.7. Investigation of strategies for selective small nerve fiber stimulation in an animal model

Participants: Paweł Maciejasz, Olivier Rossel, Christine Azevedo Coste, David Andreu, David Guiraud, Hubert Taillades [Institute of Biology, Montpellier].

`../../../../projets/demar/IMG/fb.png`

Figure 10. Awake brain surgery and functional mapping.

`../../../../projets/demar/IMG/pig.png`

Figure 11. Supraspinal cord stimulation .

The electrical stimulation of nerve fibers may allow to restore or augment some body functions lost due to disease or injury. However, in typical peripheral nerves there are thousands of nerve fibers with various diameters and functions. When standard rectangular pulses are used for nerve fiber stimulation, the big fibers are activated before smaller ones. However, for many clinical applications it would be beneficial if small fibers could be activated without activation of the big ones.

Already many stimulation techniques have been proposed for fiber type and diameter selective stimulation, e.g. analog block, slowly rising pulses, high frequency block. However, due to limited efficiency of those techniques, they are still not used in clinical practice. Based on the results of the computer simulations and the experiments performed previously in the earthworm model by our team, we have proposed some modification to the existing techniques, that may increase their efficiency.

In order to verify if the proposed modifications allow for increased selectivity of stimulation as compared to the techniques already proposed, a series of experiments in rabbit model has been scheduled. The experiment consists of two phases. The objective of the first one (3-6 rabbits), the preliminary one, is to determine an adequate method to evaluate the effects of stimulation, i.e. to find out a reliable method that would allow for discrimination between various types of fibers being activated by the stimulation. The objective of the second phase (5-15 rabbits), the exploratory one, is to compare and quantify the performance of various strategies for fiber type selective stimulation.

The experiments have been performed by the DEMAR team in rabbits in the Institute of Biology in Montpellier. During this experiment the sciatic nerve of the rabbit has been stimulated using tripolar nerve cuff electrode, whereas ENG and EMG signals, as well as ankle torque have been recorded. The experiments have been started in December 2012 and so far only the first phase of the study has been completed.

The experiment was authorized through the local ethics committee for animal experiment (authorization N° CEEA-LR-12084). The work is supported by INTENSE Project.

6.3. Neuroprostheses and technology

6.3.1. Abstraction and composition for formal design of neuroprotheses

Participants: Hélène Leroux, David Andreu, Karen Godary [LIRMM].

In the framework of specification and implementation of complex digital systems on FPGA, we have developped an approach based on components whose behavior and composition are specified by generalized interpreted T-time Petri nets. One of the inherent difficulties for designer is, on the behavioral part, to account for exceptions. This often leads to a complex modeling and is a source of human errors. Indeed, it is intricate to express all the possible situations (i.e. current state of model). We have defined a way to model exception handling by integrating the well-know concept of macroplace into the formalism. The analysability of the model and the efficiency of the implementation on FPGA (reactivity and surface, ie number of logic blocks) have been preserved. An example of macroplace is given in figure 13; it contains a sub-net (set of places of its refinement) from which exception handling is simply described by a dedicated output transition (transition te on fig. 13), whatever is the current state of the sub-net.

The new formalism has been defined, as well as the model transformation based equivalent PNML generation for using existing analysis tools.

Ongoing work deals with solving state evolution conflicts introducing priorities between transitions, to avoid reaching inconsistent global state while synchronously executing the model.

6.3.2. New FES dedicated digital processor for neurostimulator

Participants: David Andreu, David Guiraud.

We designed (patent pending) and prototyped a new neural FES dedicated processor and its associated (more compact and efficient) set of instructions, as well as an embedded sequencer for accurate timing in sequencing stimulations to be performed (by the stimulator). The new neural stimulator is based on a dedicated ASIC (Application Specific Integrated Circuit), that is able to drive up to 24 channels of stimulation in absolute synchronization, and with a programmable and controlled current level distribution (patent pending). This ASIC also allows for impedance measurement. The functions of the stimulator are currently implemented in two separate chips: an analog stimulation front-end (ASIC) and a field-programmable gate array (FPGA) embedding the logic control. The FPGA embeds the new FES dedicated processor setting the output stage configuration (poles configuration and current ratio between them) and running potentially complex stimulation profiles (with a 1 μs time step and 5 μA current step); example of generated stimulations are shown in Figure 12 . It also embeds the protocol stack allowing for remote programming and online control. Online control relies on advanced and efficient modulation mechanisms, e.g. coefficient based modulation preserving balanced stimulation (Figure 12). And last but not least, it also embeds a monitoring module ensuring the respect of safety constraints stemming both from target tissue protection and electrode integrity preservation; this reference model based monitoring module ensures (configurable) current and quantity of injected charges limits and thus safe stimulation whatever are electrodes to be used (particularly for thin-film micro-electrodes). Safety limits must be defined by users (partners) according to the target and electrodes to be used.

6.3.3. Attitude reconstruction from IMU signals

Participants: Jonathan Peguet [IFMA], Daniel Simon, Christine Azevedo Coste, Roger Pissard-Gibollet [SED Inria Grenoble Rhône-Alpes].

Inertial Measurement Units (IMUs) are currently used by the team for real-time estimation of limbs attitude, e.g. as in section 6.2.2 where the attitude of a leg while walking feeds a agit phase estimator. The IMUs embedded in the FOX nodes (manufactured by HiKob) include 3 gyrometers, 3 accelerometers and 3 magnometers, from which the attitude (e.g. Euler angles) of the node can be computed. The raw measurement signals can be either processed locally in the nodes, or sent on wireless links to be processed on a remote computer.

The raw signals issued from sensors are subject to noise and bias. Additionally, the raw data flow can be corrupted by timing disturbancies induced by communication and computation. Hence, the attitude reconstruction filters must be robust against disturbancies such as noise, bias, jitter and data loss. To evaluate the robustness of attitude reconstruction filters, a software simulation package dedicated to IMUs design and analysis has been customized from the Imusim package (initially developed in Python under GPL at Univ. of Edinburgh, U.K. [44]).

The Imusin modeling features include realistic IMUs models with noise and bias, calibration procedures, radio channels deficiencies and computing timing parameters. Several versions of Extended Kalman Filters and Non-Linear Observers, in particular those previously developed at Inria Grenoble Rhône-Alpes, have been integrated and succesfully tested against measuring noise. The work is supported by Inria SENSBIO ADT.

6.3.4. Fast simulation of hybrid dynamical systems

Participants: Abir Ben Khaled [IFPEN], Daniel Simon, Mongi Ben Gaid [IFPEN].

When dealing with the design of complex systems, simulation is an indisputable step between concept design and prototype validation. Realistic simulations allow for the preliminary evaluation, tuning and possibly redesign of proposed solutions ahead of implementation, thus lowering the risks. However, the simulation of high-fidelity models is time consuming, and reaching real-time constraints is out of the capabilities of monotithic simulations running on single cores.

The aim of the on-going work is to speed up the numerical integration of hybrid dynamical systems, eventually until reaching a real-time execution, while keeping the integration errors inside controlled bounds. The basic approach consists in splitting the system into sub-models, which are integrated in parallel. I has been shown that an efficient partition must minimize the interactions between sub-models, in particular by confining

../../../../projets/demar/IMG/PastedGraphic-1.jpg

Figure 12. Examples of stimulations (observed at the output of the stimulator)

discontinuities processing inside each component. Automatic partitioning, based on some particular incidence matrices of the original system, has been investigated [17]. The method was tested with an automotive engine model, but it is generic and can be applied to other systems of hybrid ODEs/DAEs, as are large sets of muscular fibers.

6.3.5. ENG amplifier front-end

Participants: Mariam Abdallah, Fabien Soulier, Serge Bernard, Guy Cathébras.

Electroneurogram acquisition systems are usually based on tripolar cuff electrodes that are known to decrease noise from external sources, such as muscular fibers (EMG) or stimulation artifacts. Thus, we studied a preamplifier associated with this kind of electrode in a true-tripole configuration. It is designed at the transistor level to lower the number of transistors while still rejecting parasitical signals. This kind of integration reduces the size, power consumption and noise of the preamplifier compared to classical true-tripolar structures.

`../../../../projets/demar/IMG/eng_true_tripole.png`

Figure 13. Classical structure of a true-tripolar ENG preamplifier.

The true-tripole configuration consists of linear combination of signals coming from the three poles

$$V_{out} = A\left(V_{in1} - \frac{V_{in2} + V_{in3}}{2}\right) \tag{7}$$

This combination is usually realized thanks to several differential amplifiers as shown in the figure 13 , whereas the proposed preamplifier is designes as a differential pair whose negative input transistor is split into only two smaller ones. (fig. 14). The circuit is based on a modified ASIC in AMS CMOS $0.35\,\mu$m technology, with 3.3 V supply. The preamplifier provides three functions which are:

- to combine the input signals as shown in the equation (1),
- to barely amplify the neural signal to an acceptable SNR,
- and to present a differential output to a variable-gain amplifier (not presented here, but integrated into the ASIC).

../../../../projets/demar/IMG/eng_preamplifier.png

Figure 14. Three input preamplifier schematic.

To characterize the three-input preamplifier, we have to define three orthogonal modes, starting with the main mode expected to be amplified

$$V_{in1} = -2V_{in2} = -2V_{in3}, \tag{8}$$

the common mode, and the differential parasitic mode expected to be as low as possible to achieve a good EMG rejection:

$$V_{in1} = V_{in2} = V_{in3}, \tag{9}$$

$$V_{in1} = 0, \qquad V_{in2} = -V_{in3}. \tag{10}$$

DC and AC simulations were performed for these three modes. The results are presented in the figure 15 . The main results of these simulations are:

- more than 150 dB rejection ratio for the common and differential modes compared to main one,
- dynamic range of about 5 mV,
- 200 kHz bandwidth (that is far above the needs for ENG acquisition),
- the estimated flicker noise due to input and load transistors is below the μV on the required bandwidth.

This work has been presented to the 18th IFESS Annual Conference [16].

6.3.6. Characterization of the CAFE12 chip

Participants: Jérémie Salles, Fabien Soulier, Serge Bernard, Guy Cathébras.

../../../../projets/demar/IMG/eng_dc_charact.png

../../../../projets/demar/IMG/eng_ac_charact.p

Figure 15. DC (on the left) and AC (on the right) simulation results for main (b), common (a) and differential modes (c). AC differential mode is too low to be simulated using typical values.

The circuit CAFE12 (Cool Analog Front End, 12 poles) used in StimND, which was designed in 2006, exploit the bases of a DEMAR patent. A characterization of a circuit (1st version, manufactured in 2006-2007) showed limitations to its capabilities. Thus a 2nd version was designed and manufactured in 2012 to improve the circuit linearity and consumption. CAFE12 is an ASIC generating 12 current outputs. This ASIC was developed in high voltage CMOS technology (H35, Austria Mikro Systems). Each output is able to deliver/absorb a current as high as 5mA.

The measurements presented below were carried on 3 CAFE12_V2 prototypes (C8002 & C8003). Some comparisons with 2 CAFE12_V1 (T1201 & nD09) prototypes are also shown.

- Integral non linearity (INL) Figure 16 highlights the INL improvement of the anodic generators. This significant decrease is due to a wiring modification on a specific operational amplifier (OPA). No improvement on the cathodic side was expected.

- Differential non linearity (DNL) The noticed improvement on the INL is also noticeable on the differential non linearity on both anodic and cathodic sides (Figure 17 and Figure 18).

The following table 1 sums up the characterization results. The main achievements are better anodic generators (linearity and gain) and a reduced static consumption.

Table 1.

	INL (LSB)		DNL (LSB)		Gain accuracy $(\%)$		Power (mW)
	Anode	Cathode	Anode	Cathode	Anode	Cathode	Total
CAFE12_V1	-26 to 4	±5	±0.02	±0.015	±13.43	±6.19	50
CAFE12_V2	-6 to 2	-4 to 6	±0.01	±0.01	±6.53	±5.90	38

../../../../projets/demar/IMG/INL_anodes_Anciens_Nouveaux.png

Figure 16. INL comparison, anodic generators (CAFE12_V1 & CAFE12_V2)

`../../../../projets/demar/IMG/DNL_anodes_Anciens_Nouveaux.png`

Figure 17. DNL comparison, anodic side (CAFE12_V1 & CAFE12_V2)

`../../../../projets/demar/IMG/DNL_cathodes_Anciens_Nouveaux.png`

Figure 18. DNL comparison, cathodic side (CAFE12_V1 & CAFE12_V2)

GALEN Project-Team

6. New Results

6.1. Shape, Grouping and Recognition

6.1.1. Descriptors

Participants: Eduard Trulls, Iasonas Kokkinos.

In [30] we have extended our prior work on dense scale- and rotation- invariant image descriptors to take into account soft segmentation information. This allows us to discard measurements stemming from background structures, and as such renders our descriptors invariant to background changes and occlusions. This has allowed us to obtain state-of-the-art results on tasks such as large-displacement optical flow and wide-baseline stereo. We have made the implementation of these descriptors publicly available.

6.1.2. 3D structure detection

Participants: Haithem Boussaid, Iasonas Kokkinos.

In [22] we have started exploring the potential of combinatorial optimization in the medical imaging realm. We cast the problem of finding a 3D structure (a brain tumor) as that of finding the mode of a nonparametric distribution, constructed through Kernel Density Estimation. Current techniques for doing this (e.g. Mean Shift mode-seeking, Fast Gauss Transforms, etc.) are either iterative, or linear in the number of pixels, with a typically large constant. Instead, we develop a scheme that involves a very low-constant linear-time preprocessing step, and then uses Branch-and-Bound for fast mode estimation. As such it is scalable to large volumes, and serves as a rapid initialization of a region segmentation algorithm.

6.1.3. Facade parsing

Participants: Olivier Teboul, Iasonas Kokkinos, Loic Simon, Panagiotis Katsourakis, Nikos Paragios.

In [17] we pursue a Reinforcement Learning-based approach to couple image observations with a grammar-based method to partitioning a building facade. For this we expressed 2D grammar- based image parsing as a Markov decision process where an agent has to take actions in an environment so as to maximize some notion of cumulative reward (reflecting the segmentation quality). This allowed us to accelerate previous stochastic hill-climbing approaches to image parsing by more than an order of magnitude.

6.1.4. Fast object detection

Participant: Iasonas Kokkinos.

In [27] we extended our previous work on fast object detection by developing a sparse-coding method for the efficient sharing of computation among multiple object models. In particular the first processing step of 'part score' computation was originally performed separate per object category; instead, we propose to do it 'in batch mode', so as to exploit the commonalities that exist among object parts. Building on recent developments in sparse coding we have managed to construct a compact basis for this task, which in the end gave us a two-fold acceleration over our previous fastest algorithms.

6.2. Machine Learning

6.2.1. Discriminative Parameter Estimation for Random Walks Segmentation

Participants: Pierre-Yves Baudin, Puneet Kumar, Noura Azzabou, Pierre Carlier, Nikos Paragios, M. Pawan Kumar Blaschko

In [19], we proposed a a novel discriminative learning framework that estimates the parameters of a random walks segmentation framework using a training dataset. The main challenge we face is that the training samples are not fully supervised. Specifically, they provide a hard segmentation of the medical images, instead of a probabilistic segmentation. We overcome this challenge by treating the optimal probabilistic segmentation that is compatible with the given hard segmentation as a latent variable. This allows us to employ the latent support vector machine (LSVM) formulation for parameter estimation.

6.2.2. Structured Sparsity & Applications

Participants: Katerina Gkirtzou, Wojciech Zaremba, Matthew Blaschko, M. Pawan Kumar, Nikos Paragios

We developed several machine learning applications to fMRI data, including graph representations [25] and structured sparsity regularization [26], [44]. A similar structured sparsity approach was applied in the development of a novel learning algorithm, the k-support regularized SVM, with applications to neuromuscular disease classification from diffusion tensor imaging [24]. Efficient training applications for taxonomic classification were developed in [21], while a fine grained taxonomic image classification task was introduced in [45]. The role of non-maximal suppression in accurate and efficient object detection cascades was elucidated in [20]. A fast, consistent two-sample test based on kernelized statistics was developed in [33].

6.2.3. Learning from M/EEG Data with Variable Brain Activation Delays

Participants: Wojciech Zaremba, Alexander Gramfort, M. Pawan Kumar, Matthew Blaschko

In [34], propose to address the misalignment of M/EEG samples by explicitly modeling time shifts of different brain responses in a classification setup. To this end, we use the LSVM formulation, where the latent shifts are inferred while learning the classifier parameters. The inferred shifts are further used to improve the signal-to-noise ratio of the M/EEG data, and to infer the chronometry and the sequence of activations across the brain regions that are involved in the experimental task.

6.3. Biomedical Image Analysis

6.3.1. Reconstruction

Participants: Helen Langet, Nikos Paragios

In [38] an overview of the methodological foundations of biomedical image analysis as well as their use to provide answers to a variety of clinical problems are presented. The problem of volumes of rotational angiography using non-linear sparsity constraints was studied in [28] where a novel method able to handle highly under-sampled acquisitions was introduced.

6.3.2. Graphical models and Image Segmentation

Participants: Bo Xiang, Nikos Paragios

[18] presents an overview of the use of graphical models in artificial vision where both inference, learning as well as applications are discussed. In [32] a max-margin dual decomposition method was used towards learning the compact, pose invariant shape representation using higher order graphs acting both on the connectivity of the graph as well its potentials. Graphical model was used as prior in [13] under a "curve" propagation principle for generic prior-constrained organ segmentation in 2D images. Similar inspiration driven from a higher order pose invariant graphical model learned according to [32] was considered in [31] where a novel segmentation method was proposed coupling model-based and pixel-based concepts while being pose invariant. The underlying idea was to consider a two-layer interconnected graphical model acting on pixel and on control points where segmentation consistency was imposed throgh penalties on label discrepancies of the different layers. Higher order graphical models were also employed in [14] for spine segmentation using an articulated graphical model where a non-linear approach/embedding towards reducing the complexity of the inference step was considered at training.

6.3.3. Deformable Registration and Fusion

Participants: Enzo Ferrante, Sarah Parisot, Nikos Paragios

In [16] a comprehensive survey of deformable registration was presented. It was organized in three sections: the first was studying the deformation model, the second the similarity criterion while the last section discussed the different optimization strategies. The problem of atlas-based segmentation/registration in the presence of brain tumors was studied in [29] an adaptive uncertainty-driven sampling strategy was proposed coupling segmentation and registration. Both sampling spaces (quantization of the search space, deformation grid) were determined according to the observed optimization min-marginals. The challenging problem of image to slice registration was proposed in [23] where an over-parameterized low rank graphical model acting both on the plan selection as well the in-plane deformations was introduced. The main strength of the method was its ability to simultaneously recover both the plane and the organ deformation.

MNEMOSYNE Team

6. New Results

6.1. Overview

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

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

Presently, our team is engaged on the following topics: Concerning perception loops, we are firstly considering the role of the hippocampus and of the posterior cortex in learning high level sensory cues that contribute to pavlovian conditioning in the amygdala. Secondly, we are investigating the role of the thalamus in attentional shifts in the cortex. This latter topic relies on recent advances we made in the Keops ANR project (*cf.* § 7.1) on a model of the retina that we also sum up here. Concerning Action loops, we are preparing a critical analysis of the current views of the interactions between the prefrontal cortex and the basal ganglia. Finally, we also report here more methodological achievements.

6.2. Pavlovian conditioning

The fundamental role of the amygdala in pavlovian conditioning is widely acknowledged, both on the motor, autonomic and hormonal expression of pavlovian responses and on the learning of the associations between conditional and unconditioned stimuli. This year, we have proposed models showing in a systemic view how errors of prediction in the amygdala might trigger, by cholinergic hormonal expression, episodic learning in the hippocampus [17] and semantic learning and attentional shift in the posterior cortex [8]. On this basis, we are currently studying how emotional values of sensory information are dynamically encoded in the basolateral nucleus of the amygdala [11], to fit with the multiple requirements for this kind of information in the brain.

6.3. The thalamus is more than a relay

Many recent results in neuroscience indicate that the role of the thalamus in the brain is certainly more important than it used to be considered, particularly concerning its relation with the cortex. Interestingly, we considered this question as a side effect of our work in the Keops project (*cf.* § 7.1) with our chilean neuroscientist colleagues studying non standard ganglion cells in the retina. Our modeling [6] and bibliographic studies led us to propose a biologically-founded algorithm [13] for the interplay between the modulatory and driving connections between the thalamus and the cortex, in the case of the projection of these ganglion cells on the thalamus [14]. This study has been carried out with the strong constraint of proposing a system working on a real visual flow. In the near future, we aim at developing this original view of the thalamus in the more general case of its pulvinar associative nucleus, learning to route to the cortex multimodal information flows.

6.4. Eyes are really smarter than believed

While it is known that the retina has not only standard ganglion cells devoted to contrast (parvo) or motion (magno) cues, but is able to perform sophisticated detection of spatial or temporal events in the visual scene (konio) [40], it was still to understand how such computation could be implemented as a robust effective processing of realistic natural image sequence and not only as a toy model of cartoonish stimuli. It has been shown in our group [19] that biologically plausible variational models of non-linear filtering coupled with optimized simple threshold mechanisms as derived from statistical learning mechanisms provides an efficient and realistic simulation of such non-standard retina output as observable at the level of konio-cells [14].

6.5. On the computational efficiency of Basal Ganglia models

Many valuable models have been proposed to capture the richness of the fundamental relations between the basal ganglia and prominent brain structures including the prefrontal cortex, the hippocampus and the superior colliculus. To choose among them the mechanisms on which to build the design of the motor pole of our brain-inspired system, a fundamental issue is to evaluate the efficiency of these models in more realistic cases than the ones which are generally considered by the authors. For this reason, we are presently preparing a comparative study of models, including a model developed in our lab with neuroscientist colleagues [2], in the more realistic case of large sensory and motor flows.

6.6. From distributed computing to distributed computing

One of the challenges in our systemic approach is to promote behaviors using distributed adaptive numerical computing that prevents (by definition) the specification of any behavior at a global level (no homunculus). This paradigm has been found to be very close to some class of problems in computer graphics where one tries to achieve a specific effect at the whole image level while the actual implementation is made at the pixel level (fragment). Transposing our expertise in distributed numerical computing, we thus experiment GPU programming (vertex/fragment programming) that offers hardware enforcement of distributed constraints: every fragment get the same program but their combination promotes a global image effect. This has been done for the case of text rendering [4] and dashed line rendering [5].

NEUROMATHCOMP Project-Team

5. New Results

5.1. Neural Networks as dynamical systems

5.1.1. Dynamics and spike trains statistics in conductance-based Integrate-and-Fire neural networks with chemical and electric synapses

Participants: Bruno Cessac, Rodrigo Cofré.

We investigate the effect of electric synapses (gap junctions) on collective neuronal dynamics and spike statistics in a conductance-based Integrate-and-Fire neural network, driven by a Brownian noise, where conductances depend upon spike history. We compute explicitly the time evolution operator and show that, given the spike-history of the network and the membrane potentials at a given time, the further dynamical evolution can be written in a closed form. We show that spike train statistics is described by a Gibbs distribution whose potential can be approximated with an explicit formula, when the noise is weak. This potential form encompasses existing models for spike trains statistics analysis such as maximum entropy models or Generalized Linear Models (GLM). We also discuss the different types of correlations: those induced by a shared stimulus and those induced by neurons interactions. This work has been presented in several conferences [40], [39], [39], [15] and published in Chaos, Solitons and Fractals [17].

5.2. Mean field approaches

5.2.1. Asymptotic description of neural networks with correlated synaptic weights

Participants: Olivier Faugeras, James Maclaurin.

We study the asymptotic law of a network of interacting neurons when the number of neurons becomes infinite. Given a completely connected network of neurons in which the synaptic weights are Gaussian correlated random variables, we describe the asymptotic law of the network when the number of neurons goes to infinity. We introduce the process-level empirical measure of the trajectories of the solutions to the equations of the finite network of neurons and the averaged law (with respect to the synaptic weights) of the trajectories of the solutions to the equations of the network of neurons. The main result of this work is that the image law through the empirical measure satisfies a large deviation principle with a good rate function which is shown to have a unique global minimum. Our analysis of the rate function allows us also to characterize the limit measure as the image of a stationary Gaussian measure defined on a transformed set of trajectories. This work is available on ArXiV and is under review for a Journal. A preliminary version has been presented at the CNS meeting [42].

5.2.2. Beyond dynamical mean-field theory of neural networks

Participants: Bruno Cessac, Massimiliano Muratori.

We consider a set of N firing rate neurons with discrete time dynamics and a leak term. The nonlinearity of the sigmoid is controlled by a parameter and each neuron has a firing threshold, Gaussian distributed (thresholds are uncorrelated). The network is fully connected with correlated Gaussian random synaptic weights, with mean zero and covariance matrix. When synaptic weights are uncorrelated the dynamic mean field theory allows us to draw the bifurcation diagram of the model in the thermodynamic limit (N tending to infinity): in particular there is sharp transition from fixed point to chaos characterized by the maximum Lyapunov exponent, which is known analytically in the thermodynamic limit. However, mean-field theory is exact only in the thermodynamic limit and when synaptic weights are uncorrelated. What are the deviations from mean-field theory observed when one departs from these hypotheses ? We have first studied the finite size dynamics.

For finite N the maximal Lyapunov exponent has a plateau at 0 corresponding to a transition to chaos by quasi-periodicity where dynamics is at the edge of chaos. This plateau disappears in the thermodynamic limit. Thus, mean-field theory neglects an important finite-sized effect since neuronal dynamics at the edge of chaos has strong implications on learning performances of the network. We also studied the effect of a weak correlation on dynamics. Even when correlation is small, one detects an important deviation on the maximal Lyapunov exponent. This work has been presented at the CNS conference in Paris, 2013 [43].

5.3. Neural fields theory

5.3.1. Existence of localized solutions

Participants: Pascal Chossat, Grégory Faye, James Rankin.

We have started to tackle the problem of rigorously proving the existence of localized solutions to the neural fields equations. Existence of such solutions had been assumed or guessed from numerical simulations by other researchers. In a series of articles starting with [55] we have used ideas from the theory of ordinary differential equations (existence of homoclinic orbits) [56], [19], and the theory of partial differential equations (Swift-Hohenberg equation) [16] to show the existence of localized solutions for an extended variety of neural fields equations. This is important both theoretically and for neuroscience since these solutions are considered to characterize working (short-term) memory.

5.3.2. A Center Manifold Result for Delayed Neural Fields Equations

Participants: Olivier Faugeras, Romain Veltz.

We have developed a framework for the study of delayed neural fields equations and proved a center manifold theorem for these equations. Specific properties of delayed neural fields equations make it difficult to apply existing methods from the literature concerning center manifold results for functional differential equations. Our approach for the proof of the center manifold theorem uses the original combination of results from Vanderbauwhede et al. [1992] together with a theory of linear functional differential equations in a history space larger than the commonly used set of time-continuous functions. This work has appeared in the SIAM Journal on Mathematical Analysis [24].

5.3.3. Interplay between synaptic delays and propagation delays in neural fields equations

Participant: Romain Veltz.

Neural field equations describe the activity of neural populations at a mesoscopic level. Although the early derivation of these equations introduced space dependent delays coming from the finite speed of signal propagation along axons, there has been few studies concerning their role in shaping the nonlinear dynamics of neural activity. This is mainly due to the lack of analytical tractable models. On the other hand, constant delays have been introduced to model the synaptic transmission and the spike initiation dynamics. By incorporating the two kind of delays in the neural fields equations, we are able to find the Hopf bifurcation curves analytically which produce many Hopf-Hopf interactions. We use normal theory to study two different types of connectivity that reveals a surprisingly rich dynamical portrait. In particular, the shape of the connectivity strongly influences the spatiotemporal dynamics. This work has appeared in SIAM Journal on Applied Dynamical Systems [25].

5.3.4. Stochastic neural field equations: A rigorous footing

Participants: James Inglis, Olivier Faugeras.

We extend the theory of neural fields which has been developed in a deterministic framework by considering the influence spatio-temporal noise. The outstanding problem that we address here is the development of a theory that gives rigorous meaning to stochastic neural field equations, and conditions ensuring that they are well-posed. Previous investigations in the field of computational and mathematical neuroscience have been numerical for the most part. Such questions have been considered for a long time in the theory of stochastic partial differential equations, where at least two different approaches have been developed, each having its advantages and disadvantages. It turns out that both approaches have also been used in computational and mathematical neuroscience, but with much less emphasis on the underlying theory. We present a review of two existing theories and show how they can be used to put the theory of stochastic neural fields on a rigorous footing. We also provide general conditions on the parameters of the stochastic neural field equations under which we guarantee that these equations are well-posed. In so doing, we relate each approach to previous work in computational and mathematical neuroscience. We hope this will provide a reference that will pave the way for future studies (both theoretical and applied) of these equations, where basic questions of existence and uniqueness will no longer be a cause for concern. This work is available on ArXiV and is under review for a Journal.

5.4. Spike trains statistics

5.4.1. Decoding the retina with the first wave of spikes

Participants: John Barrett [Institute of Neuroscience, Medical School, Newcastle University, Newcastle UK], Pierre Kornprobst, Geoffrey Portelli, Evelyne Sernagor [Institute of Neuroscience, Medical School, Newcastle University, Newcastle UK].

Understanding how the retina encodes visual information remains an open question. Using MEAs on salamander retinas [60] showed that the relative latencies between some neuron pairs carry sufficient information to identify the phase of square-wave gratings (Using gratings of varying phase, spatial frequency, and contrast on mouse retinas, we extended this idea by systematically considering the relative order of all spike latencies, i.e. the shape of the first wave of spikes after stimulus onset. The discrimination task was to identify the phase among gratings of identical spatial frequency. We compared the performance (fraction correct predictions) of our approach under classical Bayesian and LDA decoders to spike count and response latency of each recorded neuron. Best results were obtained for the lowest spatial frequency. There, results showed that the spike count discrimination performance was higher than for latency under both the Bayesian (0,95±0,02 and 0,75±0,11 respectively) and LDA (0,95±0,01 and 0,62±0,03 respectively) decoders. The first wave of spikes decoder is (0,46±0,06) less efficient than the spike count. Nevertheless, it accounts for 50% of the overall performance. Interestingly, these results tend to confirm the rank order coding hypothesis [59] which we are currently investigating further.

This work has been presented in [45].

5.4.2. Spike train statistics from empirical facts to theory: the case of the retina

Participants: Bruno Cessac, Adrian Palacios [CINV-Centro Interdisciplinario de Neurociencia de Valparaiso, Universidad de Valparaiso].

This work focuses on methods from statistical physics and probability theory allowing the analysis of spike trains in neural networks. Taking as an example the retina we present recent works attempting to understand how retina ganglion cells encode the information transmitted to the visual cortex via the optical nerve, by analyzing their spike train statistics. We compare the maximal entropy models used in the literature of retina spike train analysis to rigorous results establishing the exact form of spike train statistics in conductance-based Integrate-and-Fire neural networks. This work has been published in Mathematical Problems in Computational Biology and Biomedicine, F. Cazals and P. Kornprobst, Springer [29].

5.4.3. Hearing the Maximum Entropy Potential of neuronal networks

Participants: Bruno Cessac, Rodrigo Cofré.

We consider a spike-generating stationary Markov process whose transition probabilities are known. We show that there is a canonical potential whose Gibbs distribution, obtained from the Maximum Entropy Principle (MaxEnt), is the equilibrium distribution of this process. We provide a method to compute explicitly and exactly this potential as a linear combination of spatio-temporal interactions. The method is based on the Hammersley Clifford decomposition and on periodic orbits sampling. As an application, we establish an explicit correspondence between the parameters of the Ising model and the parameters of Markovian models like the Generalized-Linear Model. This work has been presented in several conferences [39], [27], and submitted to Phys. Rev. Letters [41], see also the research report [31].

5.4.4. Spatio-temporal spike trains analysis for large scale networks using maximum entropy principle and Monte-Carlo method

Participants: Bruno Cessac, Olivier Marre [Institut de la Vision, Paris, France], Hassan Nasser.

Understanding the dynamics of neural networks is a major challenge in experimental neuroscience. For that purpose, a modelling of the recorded activity that reproduces the main statistics of the data is required. We present a review on recent results dealing with spike train statistics analysis using maximum entropy models (MaxEnt). Most of these studies have been focusing on modelling synchronous spike patterns, leaving aside the temporal dynamics of the neural activity. However, the maximum entropy principle can be generalized to the temporal case, leading to Markovian models where memory effects and time correlations in the dynamics are properly taken into account. We also present a new method based on Monte-Carlo sampling which is suited for the fitting of large-scale spatio-temporal MaxEnt models. The formalism and the tools presented will be essential to fit MaxEnt spatio-temporal models to large neural ensembles. This work has been presented in several conferences [39], [15], [44] and published in Journal of Statistical Mechanics [20].

5.4.5. Spike train statistics and Gibbs distributions

Participants: Bruno Cessac, Rodrigo Cofré.

We introduce Gibbs distribution in a general setting, including non stationary dynamics, and present then three examples of such Gibbs distributions, in the context of neural networks spike train statistics: (i) Maximum entropy model with spatio-temporal constraints; (ii) Generalized Linear Models; (iii) Conductance based Integrate and Fire model with chemical synapses and gap junctions. This leads us to argue that Gibbs distributions might be canonical models for spike train statistics analysis. This work has published in J. Physiol. Paris [15].

5.4.6. A maximum likelihood estimator of neural network synaptic weights

Participants: Bruno Cessac, Wahiba Taouali.

Given a conductance-based Integrate-and-Fire model where the spike statistics dependence on synaptic weights is known, can one reconstruct this network of synaptic weights from the observation of a raster plot generated by the network ? We have solved this inverse problem using an explicit expression of a maximum likelihood estimator based on the Newton-Raphson method. This estimator uses analytically computed gradients and Hessian of the likelihood function given by the product of conditional probabilities. The explicit form of these conditional probabilities can be found in [49]. Our results show that this method allows to estimate the set of connections weights knowing the input, the noise distribution and the leak function. This work has been presented in the CNS conference in Paris, 2013 [47].

5.5. Synaptic plasticity

5.5.1. Effects of Cellular Homeostatic Intrinsic Plasticity on Dynamical and Computational Properties of Biological Recurrent Neural Networks

Participants: Hugues Berry, Bruno Cessac, Bruno Delord, Jérémie Naudé.

Homeostatic intrinsic plasticity (HIP) is a ubiquitous cellular mechanism regulating neuronal activity, cardinal for the proper functioning of nervous systems. In invertebrates, HIP is critical for orchestrating stereotyped activity patterns. The functional impact of HIP remains more obscure in vertebrate networks, where higher-order cognitive processes rely on complex neural dynamics. The hypothesis has emerged that HIP might control the complexity of activity dynamics in recurrent networks, with important computational consequences. However, conflicting results about the causal relationships between cellular HIP, network dynamics and computational performance have arisen from machine learning studies. in this work, we assess how cellular HIP effects translate into collective dynamics and computational properties in biological recurrent networks. We develop a realistic multi scale model including a generic HIP rule regulating the neuronal threshold with actual molecular signaling pathways kinetics, Dale's principle, sparse connectivity, synaptic balance and Hebbian synaptic plasticity (SP). Dynamic mean-field analysis and simulations unravel that HIP sets a working point at which inputs are transduced by large derivative ranges of the transfer function. This cellular mechanism insures increased network dynamics complexity, robust balance with SP at the edge of chaos, and improved input separability. Although critically dependent upon balanced excitatory and inhibitory drives, these effects display striking robustness to changes in network architecture, learning rates and input features. Thus, the mechanism we unveil might represent a ubiquitous cellular basis for complex dynamics in neural networks. Understanding this robustness is an important challenge to unravel principles underlying self-organization around criticality in biological recurrent neural networks. This work has been published in the Journal of Neuroscience [21].

5.5.2. Short-term synaptic plasticity in the deterministic Tsodyks-Markram model leads to unpredictable network dynamics

Participants: Jesus Cortes, Mathieu Desroches, Serafim Rodrigues, Romain Veltz, Miguel Munoz, Terrence Sejnowski.

Short-term synaptic plasticity strongly affects the neural dynamics of cortical networks. The Tsodyks and Markram (TM) model for short-term synaptic plasticity accurately accounts for a wide range of physiological responses at different types of cortical synapses. We report a route to chaotic behavior via a Shilnikov homoclinic bifurcation that dynamically organizes some of the responses in the TM model. In particular, the presence of such a homoclinic bifurcation strongly affects the shape of the trajectories in the phase space and induces highly irregular transient dynamics; indeed, in the vicinity of the Shilnikov homoclinic bifurcation, the number of population spikes and their precise timing are unpredictable and highly sensitive to the initial conditions. Such an irregular deterministic dynamics has its counterpart in stochastic/network versions of the TM model: The existence of the Shilnikov homoclinic bifurcation generates complex and irregular spiking patterns and acting as a sort of springboard facilitates transitions between the down-state and unstable periodic orbits. The interplay between the (deterministic) homoclinic bifurcation and stochastic effects may give rise to some of the complex dynamics observed in neural systems.

This work has been published in the Proceedings of the National Academy of Sciences [52].

5.6. Visual Neuroscience

5.6.1. Bifurcation Study of a Neural Fields Competition Model with an Application to Perceptual Switching in Motion Integration

Participants: James Rankin, Andrew Meso [Institut de Neurosciences de la Timone, UMR 6193, CNRS, Marseille, France], Guillaume S. Masson [Institut de Neurosciences de la Timone, UMR 6193, CNRS, Marseille, France], Olivier Faugeras, Pierre Kornprobst.

In this work we have investigated the underlying mechanisms that gate multistable perception, by focusing on the presentation of 1:1 barber pole during long presentations, which is perceived to move in a direction that changes every few seconds. This phenomenon has been studied from the perspective of dynamical systems modeling and human psychophysics: From a modeling point of view, numerical tools from bifurcations analysis were applied to the study of a competition model posed as a feature-only neural field equation (with a continuous feature space) where adaptation and noise are implemented as mechanisms that can drive activity switches. Human psychophysics experiments were jointly done by INT (Institut de Neurosciences de la Timone, Marseille): Human observers were presented a moving grating stimulus over 15s while eye movements and reports of perceptual switches were recorded. Investigating the stimulus contrast, we found that the peak in switching rate observed experimentally occurs close to a bifurcation in the model that separates two mechanistic regimes. By identifying signatures of the switching predicted by the model with the behavioural data at different parts of the transition-contrast curve, we found for the first time, evidence for a dominance of driving mechanisms which shifts from noise dominated at low contrasts to adaptation dominated at higher contrasts.

This work has been published in [22], [23].

5.6.2. A Retinotopic Neural Fields Model of Perceptual Switching in 2D Motion Integration

Participants: Pierre Kornprobst, Guillaume S. Masson [Institut de Neurosciences de la Timone, UMR 6193, CNRS, Marseille, France], Kartheek Medathati, James Rankin.

In perceptual multistability a fixed but ambiguous stimulus can invoke multiple interpretations although only one can be held at a time. Visual motion stimuli are inherently ambiguous, for instance due to the aperture problem, which makes motion perception a complex inference task. The underlying cortical dynamics that select one percept out of multiple competing possibilities are not fully understood. Recent studies by [22] and [68] have tried to address this problem using the neural fields formalism. In [22], a switching behaviour for a classical psychophysics stimulus, the multistable barberpole, was successfully captured in a feature-only, one-layer model of MT with adaptation and noise. However, without a representation of space, only some very specific stimulus could be considered. The work reported in [68] provides a much more general framework for motion integration in a two layer-model, however, it fails to capture the switching behaviour as the mechanisms of adaptation and noise were not considered. Building on the strengths of both studies, we propose a model that takes into account the spatial domain in a two-layer configuration whilst incorporating both adaptation and noise. Interactions between two layers processing local motion (V1 and MT) occurred through recurrent and lateral connections. The input stimuli are represented using direction of motion signals extracted using Reichardt detectors at corresponding 2D spatial locations. We use stimuli such as drifting bars and barberpoles to constrain the model to a suitable operating regime. In terms of computations, since the model is demanding, we implemented it using GPUs, extending the methods of [13]. Based on this implementation, we study dynamics of the model focusing on coherency in plaid motion (plaids and crossed barber pole).

This work has been presented in [28]

NEUROSYS Team

6. New Results

6.1. From the microscopic to the mesoscopic scale

Participants: Axel Hutt, Laure Buhry, Meysam Hashemi, Pedro Garcia Rodriguez, Peter beim Graben, James Wallace Sleigh.

Several previous studies focus on the derivation of neural population models. However most of these studies do not consider explicitly the microscopic properties of neurons, such as synaptic receptor dynamics or ion-channel distributions, although they may be implicit. The resulting models in some previous studies are poorly tractable analytically due to their complexity. Moreover, the complexity of previous models makes it difficult to discover those elements in the model that induce certain dynamical features as observed in experiments. Essentially most of previous studies do not consider the spatial interactions of neurons and, importantly, neglect delays present in biological networks. We aim to improve some previous models and a first step to a new statistical approach has been developed [11], [18], [17], [22] to bridge the scales between the network activity of coupled spiking single neurons and statistical quantities of populations, e.g., the mean membrane potential in the network and the networks population firing rate. Our work considers the specific effect of anaesthetics and takes into account the physiological effects of extra-synaptic GABA_A-receptors at single neurons, which are highly sensitive to anaesthetic drugs, such as propofol. We find numerically by simulation of a spiking neural network that propofol on single neuron level diminishes the network oscillation power in the $\alpha-$frequency band and affects strongly the spike coherence in the population. Such effects have been shown in previous experimental data obtained during propofol anaesthesia demonstrating the importance of extra-synaptic receptor dynamics in the understanding of experimental phenomena in anaesthesia.

The neural origin of generation and planning of motor action in humans is still unknown. In this context, psychophysical experiments and the neural modeling of the gained results may lead to further insight. We have participated in an experimental and theoretical study [8] to reveal the effect of temporal attention on non-conscious prime processing. Our stochastic accumulator model improves extensively the standard accumulator model for reaction time by involving additional stochastic neural accumulators, which permits an almost perfect fit to experimental data. The model indicates that motor action, which is generated on a population level, obeys a stochastic accumulation of activity of single neuron activity.

6.2. From the mesoscopic to the behavior scale

Participants: Axel Hutt, Laurent Bougrain, Eric Nichols, Maxime Rio, Carolina Saavedra, Louis Korczkovsky, Alexandre Martin, Pierre-Jean Morieux.

To link neural population activity to behavior, it is necessary to understand well the dynamic properties of population models which we have studied in general models [5], [14], [15], [16], [24]. To this end, we have analyzed a neural mass model [4] describing the neural population activity subject to synaptic anaesthetic action to explain characteristic signal features in measured EEG. The model explains the gain of power in the $\alpha-$ and $\delta-$frequency observed experimentally by a dynamic oscillatory instability (Hopf instability). The model considers a cortical population only and hence the result indicates that the experimental feature observed may originate in the cortex.

An extended population model considers not only the cortex but a feedback-loop to the thalamus. This model involves a delayed interaction. At first, we have studied the dynamics of delayed dynamical systems subject to additive stimuli [23], [7], [6] to learn more about the expected activity. Our first study of a linear thalamo-cortical feedback model [21], [20] reveals the descriptive power of neural mass models to describe EEG under anaesthesia.

In order to learn more about the effect of anaesthetics on neural populations, we have participated in the data analysis of an experimental study on anaesthesia in animals [9]. Moreover we have started developing new data analysis techniques to extract novel features from EEG. In his doctoral thesis, Maxime Rio has developed a new method to detect transient amplitude synchronization in multi-variate time series in a subset of time series [1]. Carolina Saavedra has conducted wavelet analysis in her thesis to improve the denoising in BCI-relevant measured signals [2]. Another study [3] proposes a new recurrence plot-technique based on symbolic dynamics. It extracts spatio-temporally recurrence patterns in a multivariate dataset which reflect underlying neural recurrent dynamics.

The event-related potentials (ERP) in EEG are important markers of cognitive processes in the brain and serve as features to control interfaces in BCI. We have performed advanced studies to improve the detection of ERP [13], [12].

PARIETAL Project-Team

6. New Results

6.1. Deformable Template estimation for joint anatomical and functional brain images

Participants: Bertrand Thirion [Correspondant], Hao Xu, Stéphanie Allassonnière.

Traditional analyses of Functional Magnetic Resonance Imaging (fMRI) use little anatomical information. The registration of the images to a template is based on the individual anatomy and ignores functional information; subsequently detected activations are not confined to gray matter (GM). In this work, we propose a statistical model to estimate a probabilistic atlas from functional and T1 MRIs that summarizes both anatomical and functional information and the geometric variability of the population. Registration and Segmentation are performed jointly along the atlas estimation and the functional activity is constrained to the GM, increasing the accuracy of the atlas.

More details can be found in [69].

6.2. Randomized parcellation-based inference

Participants: Gaël Varoquaux, Bertrand Thirion, Benoit Da Mota, Virgile Fritsch.

Neuroimaging group analyses are used to relate inter-subject signal differences observed in brain imaging with behavioral or genetic variables and to assess risks factors of brain diseases. The lack of stability and of sensitivity of current voxel-based analysis schemes may however lead to non-reproducible results. We introduce a new approach to overcome the limitations of standard methods, in which active voxels are detected according to a consensus on several random parcellations of the brain images, while a permutation test controls the false positive risk (see Fig. 3). Both on synthetic and real data, this approach shows higher sensitivity, better accuracy and higher reproducibility than state-of-the-art methods. In a neuroimaging-genetic application, we find that it succeeds in detecting a significant association between a genetic variant next to the COMT gene and the BOLD signal in the left thalamus for a functional Magnetic Resonance Imaging contrast associated with incorrect responses of the subjects from a Stop Signal Task protocol.

More details can be found in [55].

6.3. Group-level impacts of within- and between-subject hemodynamic variability in fMRI

Participants: Gaël Varoquaux, Solveig Badillo, Philippe Ciuciu [Correspondant].

Inter-subject fMRI analyses have specific issues regarding the reliability of the results concerning both the detection of brain activation patterns and the estimation of the underlying dynamics. Among these issues lies the variability of the hemodynamic response function (HRF), that is usually accounted for using functional basis sets in the general linear model context. Here, we use the joint detection-estimation approach (JDE) [76], [78], which combines regional nonparametric HRF inference with spatially adaptive regularization of activation clusters to avoid global smoothing of fMRI images (see Fig. 4). We show that the JDE-based inference brings a significant improvement in statistical sensitivity for detecting evoked activity in parietal regions. In contrast, the canonical HRF associated with spatially adaptive regularization is more sensitive in other regions, such as motor cortex. This different regional behavior is shown to reflect a larger discrepancy of HRF with the canonical model. By varying parallel imaging acceleration factor, SNR-specific region-based hemodynamic parameters (activation delay and duration) were extracted from the JDE inference. Complementary analyses highlighted their significant departure from the canonical parameters and the strongest between-subject variability that occurs in the parietal region, irrespective of the SNR value. Finally, statistical evidence that the fluctuation of the HRF shape is responsible for the significant change in activation detection performance is demonstrated using paired t-tests between hemodynamic parameters inferred by GLM and JDE.

`../../../../projets/parietal/IMG/method.png`

Figure 3. Overview of the randomized parcellation based inference framework on an example with few parcels. The variability of the parcels definition is used to obtain voxel-level statistics.

More details can be found in [49].

`../../../../projets/parietal/IMG/solveig.png`

Figure 4. General sketch summarizing the HRF computation at the subject and group-levels in activated regions r. Left: Position of the activation peak in r (here left motor cortex) given in mm in the Talairach space. Center: Individual weighted HRF time course extraction. Right: Computation of the group average normalized HRF time course with corresponding error bars ($\pm\sigma$).

6.4. Mapping cognitive ontologies to and from the brain

Participants: Gaël Varoquaux [Correspondant], Bertrand Thirion, Yannick Schwartz.

Imaging neuroscience links brain activation maps to behavior and cognition via correlational studies. Due to the nature of the individual experiments, based on eliciting neural response from a small number of stimuli, this link is incomplete, and unidirectional from the causal point of view. To come to conclusions on the function implied by the activation of brain regions, it is necessary to combine a wide exploration of the various brain functions and some inversion of the statistical inference. Here we introduce a methodology for accumulating knowledge towards a bidirectional link between observed brain activity and the corresponding function. We rely on a large corpus of imaging studies and a predictive engine. Technically, the challenges are to find commonality between the studies without denaturing the richness of the corpus. The key elements that we contribute are labeling the tasks performed with a cognitive ontology, and modeling the long tail of rare paradigms in the corpus. To our knowledge, our approach is the first demonstration of predicting the cognitive content of completely new brain images. To that end, we propose a method that predicts the experimental paradigms across different studies (see Fig. 5).

More details can be found in [63].

6.5. Implications of Inconsistencies between fMRI and dMRI on Multimodal Connectivity Estimation

Participants: Gaël Varoquaux [Correspondant], Bertrand Thirion, Bernard Ng.

There is a recent trend towards integrating resting state functional magnetic resonance imaging (RS-fMRI) and diffusion MRI (dMRI) for brain connectivity estimation, as motivated by how estimates from these modalities are presumably two views reflecting the same underlying brain circuitry. In this work, we show on a cohort of 60 subjects that conventional functional connectivity (FC) estimates based on Pearson's correlation and anatomical connectivity (AC) estimates based on fiber counts are actually not that highly correlated for typical RS-fMRI (7 min) and dMRI (32 gradient directions) data. The FC-AC correlation can be significantly increased by considering sparse partial correlation and modeling fiber endpoint uncertainty, but the resulting FC-AC correlation is still rather low in absolute terms. We further exemplify the inconsistencies between FC and AC estimates by integrating them as priors into activation detection and demonstrating significant differences in their detection sensitivity. Importantly, we illustrate that these inconsistencies can be useful in fMRI-dMRI integration for improving brain connectivity estimation.

More details can be found in [61]. See also [60].

6.6. Extracting brain regions from rest fMRI with Total-Variation constrained dictionary learning

Participants: Gaël Varoquaux [Correspondant], Alexandre Abraham.

Spontaneous brain activity reveals mechanisms of brain function and dysfunction. Its population-level statistical analysis based on functional images often relies on the definition of brain regions that must summarize efficiently the covariance structure between the multiple brain networks. In this paper, we extend a network-discovery approach, namely dictionary learning, to readily extract brain regions. To do so, we introduce a new tool drawing from clustering and linear decomposition methods by carefully crafting a penalty. Our approach automatically extracts regions from rest fMRI that better explain the data and are more stable across subjects than reference decomposition or clustering methods (see FIg. 6).

More details can be found in [47].

6.7. Cohort-level brain mapping: learning cognitive atoms to single out specialized regions

Participants: Gaël Varoquaux [Correspondant], Bertrand Thirion, Yannick Schwartz.

../../../../projets/parietal/IMG/forward_regions_stimulus_modality.png

../../../../projets/parietal/IMG/reverse_regions

../../../../projets/parietal/IMG/forward_regions_response_modality.png

../../../../projets/parietal/IMG/reverse_regions

../../../../projets/parietal/IMG/forward_regions_instructions.png

../../../../projets/parietal/IMG/reverse_regions

../../../../projets/parietal/IMG/kmeans_hard.png

../../../../projets/parietal/IMG/ward_hard.png

../../../../projets/parietal/IMG/msdl_hard.png

../../../../projets/parietal/IMG/msdl_unreg_hard

../../../../projets/parietal/IMG/smooth_msdl_hard.png

../../../../projets/parietal/IMG/icams_hard.png

Functional Magnetic Resonance Imaging (fMRI) studies map the human brain by testing the response of groups of individuals to carefully-crafted and contrasted tasks in order to delineate specialized brain regions and networks. The number of functional networks extracted is limited by the number of subject-level contrasts and does not grow with the cohort. Here, we introduce a new group-level brain mapping strategy to differentiate many regions reflecting the variety of brain network configurations observed in the population. Based on the principle of functional segregation, our approach singles out functionally-specialized brain regions by learning group-level functional profiles on which the response of brain regions can be represented sparsely. We use a dictionary-learning formulation that can be solved efficiently with on-line algorithms, scaling to arbitrary large datasets. Importantly, we model inter-subject correspondence as structure imposed in the estimated functional profiles, integrating a structure-inducing regularization with no additional computational cost. On a large multi-subject study, our approach extracts a large number of brain networks with meaningful functional profiles (see Fig. 7).

More details can be found in [66].

6.8. Identifying predictive regions from fMRI with TV-$\ell 1$ prior

Participants: Gaël Varoquaux [Correspondant], Bertrand Thirion, Alexandre Gramfort.

Decoding, i.e. predicting stimulus related quantities from functional brain images, is a powerful tool to demonstrate differences between brain activity across conditions. However, unlike standard brain mapping, it offers no guaranties on the localization of this information. Here, we consider decoding as a statistical estimation problem and show that injecting a spatial segmentation prior leads to unmatched performance in recovering predictive regions. Specifically, we use $\ell 1$ penalization to set voxels to zero and Total-Variation (TV) penalization to segment regions. Our contribution is two-fold. On the one hand, we show via extensive experiments that, amongst a large selection of decoding and brain-mapping strategies, TV+$\ell 1$ leads to best region recovery (see Fig. 8). On the other hand, we consider implementation issues related to this estimator. To tackle efficiently this joint prediction-segmentation problem we introduce a fast optimization algorithm based on a primal-dual approach. We also tackle automatic setting of hyper-parameters and fast computation of image operation on the irregular masks that arise in brain imaging.

More details can be found in [59].

6.9. Second order scattering descriptors predict fMRI activity due to visual textures

Participants: Michael Eickenberg, Bertrand Thirion [Correspondant], Alexandre Gramfort.

Second layer scattering descriptors are known to provide good classification performance on natural quasi-stationary processes such as visual textures due to their sensitivity to higher order moments and continuity with respect to small deformations. In a functional Magnetic Resonance Imaging (fMRI) experiment we present visual textures to subjects and evaluate the predictive power of these descriptors with respect to the predictive power of simple contour energy - the first scattering layer. We are able to conclude not only that invariant second layer scattering coefficients better encode voxel activity, but also that well predicted voxels need not necessarily lie in known retinotopic regions (see Fig. 9).

More details can be found in [56].

6.10. Bayesian Joint Detection-Estimation of cerebral vasoreactivity from ASL fMRI data

Participants: Thomas Vincent, Philippe Ciuciu [Correspondant].

../../../../projets/parietal/IMG/atlas.png

Figure 7. (Left) A brain functional atlas can be conceptualized as a parcellation of the brain volume into overlapping networks, where each functional network is characterized by a profile of activation for a set of functional contrasts. (Right) Such an atlas can be learned by applying an adapted dictionary learning to a set of images that display the activation observed in different subjects for a (very large) set of cognitive tasks.

../../../../projets/parietal/IMG/tvl1.png

Figure 8. Results on fMRI data from (from left to right F-test, ElasticNet and TV-ℓ_1). The TV-ℓ_1 regularized model segments neuroscientificly meaningful predictive regions in agreement with univariate statistics while the ElasticNet yields sparse although very scattered non-zero weights.

../../../../projets/parietal/IMG/scattering.png

Figure 9. Some brain regions are better explained by using two scattering layers rather than one (middle). These regions are symetric across hemispheres, and are observed mostly in the dorsal stream of the visual cortex. An atlas of the visual areas (left and right) shows that the mai foci are found in the V1, V2, V3AB and IPS0 regions.

Although the study of cerebral vasoreactivity using fMRI is mainly conducted through the BOLD fMRI modality, owing to its relatively high signal-to-noise ratio (SNR), ASL fMRI provides a more interpretable measure of cerebral vasoreactivity than BOLD fMRI. Still, ASL suffers from a low SNR and is hampered by a large amount of physiological noise. The current contribution aims at improving the recovery of the vasoreactive component from the ASL signal. To this end, a Bayesian hierarchical model is proposed, enabling the recovery of perfusion levels as well as fitting their dynamics. On a single-subject ASL real data set involving perfusion changes induced by hypercapnia, the approach is compared with a classical GLM-based analysis. A better goodness-of-fit is achieved, especially in the transitions between baseline and hypercapnia periods. Also, perfusion levels are recovered with higher sensitivity and show a better contrast between gray- and white matter.

More details can be found in [68].

Popix Team

6. New Results

6.1. Estimation in mixed-effects diffusion models

Participant: Marc Lavielle.

We have coupled the SAEM algorithm and the extended Kalman filter for maximum likelihood estimation in mixed-effects diffusion models: we have considered some general mixed-effects diffusion models, in which observations are made at discrete time points and include measurement errors. In these models, the observed likelihood is generally not explicit, making maximum likelihood estimation of the parameters particularly complex. We have proposed a specific inference methodology for these models. In particular, it combines the SAEM algorithm with the extended Kalman filter to estimate the population parameters. We have also provided some tools for estimating the individual parameters, for recovering the individual underlying diffusion trajectories and for evaluating the model. We evaluated the methods on simulations and applied them to a pharmacokinetics example.

6.2. Estimation in mixtures of models

Participant: Marc Lavielle.

We have proposed an improved SAEM algorithm for maximum likelihood estimation in mixtures of non linear mixed effects models. This involves a new methodology for maximum likelihood estimation in mixtures of non linear mixed effects models (NLMEM). Such mixtures of models include mixtures of distributions, mixtures of structural models and mixtures of residual error models. Since the individual parameters inside the NLMEM are not observed, we have proposed to combine the EM algorithm usually used for mixtures models when the mixture structure concerns an observed variable, with the Stochastic Approximation EM (SAEM) algorithm, which is known to be suitable for maximum likelihood estimation in NLMEM and also has nice theoretical properties. The main advantage of this hybrid procedure is to avoid a simulation step of unknown group labels required by a “full” version of SAEM. The resulting MSAEM (Mixture SAEM) algorithm is now implemented in the MONOLIX software. We have also proposed several criteria for classification of subjects and estimation of individual parameters. Our numerical experiments on simulated data have shown that MSAEM performs well in a general framework of mixtures of NLMEM. Indeed, MSAEM provides an estimator close to the maximum likelihood estimator in very few iterations and is robust with regards to initialization. Our application of the method to pharmacokinetic (PK) data demonstrated the potential of the method for practical applications.

6.3. Moving meshes with freefem++

Participants: Astrid Decoene, Bertrand Maury.

The Arbitrary Lagrangian-Eulerian framework allows to compute free surface flows with the Finite Element functions defined on a fittedmesh which follows the globalmotion of the fluid domain. We have described how freefem++ can be used to implement this method, and we have provided two and three dimensional illustrations in the context of water waves.

6.4. Modeling of the oxygen transfer in the respiratory process

Participant: Bertrand Maury.

We have proposed an integrated model for oxygen transfer into the blood, coupled with a lumped mechanical model for the ventilation process. We aim at investigating oxygen transfer into the blood at rest or exercise. The first task consists in describing nonlinear effects of the oxygen transfer under normal conditions. We also include the possible diffusion limitation in oxygen transfer observed in extreme regimes involving parameters such as alveolar and venous blood oxygen partial pressures, capillary volume, diffusing capacity of the membrane, oxygen binding by hemoglobin and transit time of the red blood cells in the capillaries. The second task consists in discussing the oxygen concentration heterogeneity along the path length in the acinus

6.5. Congestion-driven dendritic growth

Participant: Bertrand Maury.

In order to observe growth phenomena in biology where dendritic shapes appear, we have proposed a simple model where a given population evolves feeded by a diffusing nutriment, but is subject to a density constraint. The particles (e.g., cells) of the population spontaneously stay passive at rest, and only move in order to satisfy some constraint, by choosing the minimal correction velocity so as to prevent overcongestion. We treat this constraint by means of projections in the space of densities endowed with the Wasserstein distance, defined through optimal transport. This allows to provide an existence result and suggests some numerical computations, in the same spirit of what the authors did for crowd motion (but with extra difficulties, essentially due to the fact that the total mass may increase). The numerical simulations show, according to the values of the parameter and in particular of the diffusion coefficient of the nutriment, the formation of dendritic patterns in the space occupied by cells.

SHACRA Project-Team

6. New Results

6.1. Electrophysiology

Cardiac arrhythmia is a very frequent pathology that comes from an abnormal electrical activity in the myocardium. This work aims at developing a training simulator for interventional radiology and thermo-ablation of these arrhythmias. After tackling the issue of fast electrophysiology, a first version of our training simulator was proposed.

`../../../../projets/shacra/IMG/Electrophysiology.png`

Figure 3. Cardiac electrophysiology computed on a patient-specific geometry

The first main contribution of this work is the interactive catheter navigation inside a moving venous system and a beating heart. The virtual catheterization reproduces navigation issues that can be solved using a bending catheter. Second, our real-time GPU electrophysiology model allows interactions during the simulation such as extra-cellular potential measurement, RF ablation, and electrical stimulation. An innovative management of the computational units based on multithreading offers performances close to real-time. This framework is therefore a substantial step towards realistic and highly efficient virtual training systems in cardiology. As future work, we intend to use patient-specific data in our framework so that cardiologists could quantitatively assess the realism of our virtual training.

../../../../projets/shacra/IMG/TrainingSystem.jpg ../../../../projets/shacra/IMG/Navigation.png ../../../../projets/shacra/IMG/Sign

(a) Setup of the System (b) Simulation of Catheter Navigation (c) Electrophysiological Signals

Figure 4. The first simulation dedicated to electrocardiology training

6.2. Cryoablation

A new project started this year around cryotherapy. This technique consists in inserting needles that freezing the surrounding tissues, thus immediately leading to cellular death of the tissues. Cryoablation procedure is used in many medical fields for tumor ablation, and even starts being used in cardiology. In this scope, we build a simulator able to place the cryoprobes and run a simulation representing the evolution of iceballs in living tissues.

../../../../projets/shacra/IMG/CryoSimulator.jpg

Figure 5. Simulation framework for cryoablation planning

6.3. Stapedotomy

Stapedotomy is a challenging procedure of the middle ear microsurgery, since the surgeons is in direct contact with sensitive structures such as the ossicular chain. This procedures is taught and performed in the last phase of the surgical apprenticeship. To improve surgical teaching, we propose to use a virtual surgical simulator based on a finite element model of the middle ear. The static and dynamic behavior of the developed finite element model was successfully compared to published data on human temporal bones specimens. A semi-automatic algorithm was developed to perform a quick and accurate registration of our validated mechanical atlas to match the patient dataset. This method avoids a time-consuming work of manual segmentation, parameterization, and evaluation. A registration is obtained in less than 260 seconds with an accuracy close to a manual process and within the imagery resolution. The computation algorithms, allowing carving, deformation of soft and hard tissus, and collision response, are compatible with a real-time interactive simulation of a middle ear procedure. As a future work, we propose to investigate new robotized procedures of the middle ear surgery in order to develop new applications for the RobOtol device and to provide a training tool for the surgeons.

`../../../../projets/shacra/IMG/ENT.png`

Figure 6. Simulation of the stapedotomy procedure.

6.4. Radiotherapy planning

The main challenge of radiotherapy treatment is to irradiate the tumor while sparing the surrounding healthy tissues. In the case of throat cancer, the complexity of the therapy treatment is due to the proximity of organs at risk such as the two parotide glands. The parotide glands are the main salivary glands. An overdose of radiation in these glands may cause xerostomia, which is a medical term for the symptom of dryness in the mouth, or in other words, a lack of saliva. This disease affect significantly the life of the patient: difficulty talking, tasting, chewing, swallowing, excessive thirst, constant pain in the throat etcetera. A radiation therapy treatment of throat cancer takes from 5 to 7 weeks. The treatment is planned several days before the therapy. The planning consist in contouring each organ of the area on CT-scan images and defining the dose of radiation to deliver to each of these organs. This stage is lengthy and takes around two hours per patient. Yet, some anatomical variations occur in the course of the treatment, mainly due to the weight loss of the patients. These variations compromise the safety of the healthy tissues, because the planned treatments is no more up to date. For now, the physicians have no solution goog enough to handle these changes. Xerostomia affects around 20 per-cent of the patients suffering from throat cancer.

The main idea of this work is to create an interface that the physicians could use to redo the planning when it is needed, when the anatomical changes are significant. The purpose is to give to them the possibility to use what they see on images, to recreate the right shape of the contours without recontouring each images, and in a reasonable time. This interface will use their knowledge to determine the new shape of the organs. The work does not aim at providing a fully automatic method because it would reduce its acceptation by the physicians. As the method is based on the input of the physician, they can control the deformation based on images but also on their knowledge.

6.5. Image-based diagnoses

In the context of the female pelvic medicine, image-based diagnoses of pelvic floor disorders like prolapse or endometriosis rely on mechanical indicators, such as mobilities of organs and shear displacements between

../../../../projets/shacra/IMG/parotide22.png

Figure 7. Screenshot of our radotherapy planning tool.

organs. This information would be useful for both precise diagnoses and planning of surgical procedure. Involving numerical tools for diagnoses and surgery planing becomes increasingly interesting for physicians in clinical uses. The advantages of numerical models are not only in visualization, but also in quantitative measurements on a group of organs, such as their shapes and their relative movements. The processing pipeline includes patient data retrieval, image analysis, patient-specific modeling and biomechanical simulation. Our work consists in proposing new methods and algorithms for modeling the 3D anatomy of specific patients based on image data. This model should be compatible with the requirements of a biomechanical simulation. Moreover, we aim at developing new image processing tools for analyzing 2D dynamic MRI (to assess the mobilities of the pelvic system by extracting certain mechanical indicators from images) and for comparison with simulations.

Registration between geometric models and images remains a major challenge in these applications. We proposed a new model-to-image registration approach which was developed and tested for segmentation of organs in 2D images and for tracking the motion of pelvic organs from 2D dynamic MRI. Thanks to this technique, evaluation of the level of shear strain that is encountered by the fascias (connective tissues between organs) during the motion became possible. This tool could help in early diagnostic of prolapse. In the next step, our objective is to extend this method for adapting it to 3D reconstruction (with 3D geometric models and 3D MR images) and for the comparison of 3D simulations with deformable images.

6.6. Dynamic Deformations Simulated at Different Frequencies

The dynamic response of deformable bodies varies significantly in dependence on mechanical properties of the objects: while the dynamics of a stiff and light object (e. g. wire or needle) involves high-frequency phenomena such as vibrations, much lower frequencies are sufficient for capturing dynamic response of an object composed of a soft tissue. Yet, when simulating mechanical interactions between soft and stiff deformable models, a single time-step is usually employed to compute the time integration of dynamics of both objects. However, this can be a serious issue when haptic rendering of complex scenes composed of various bodies is considered. In this work, we present a novel method allowing for dynamic simulation of a scene composed of colliding objects modelled at different frequencies: typically, the dynamics of soft objects are calculated at frequency about 50 Hz, while the dynamics of stiff object is modeled at 1 kHz, being directly connected to the computation of haptic force feedback. The collision response is performed at both low and high frequencies employing data structures which describe the actual constraints and are shared between the high and low frequency loops. During the simulation, the realistic behaviour of the objects according to the mechanical principles (such as non-interpenetration and action-reaction principle) is guaranteed. We have shown several scenarios involving different bodies in interaction, demonstrating the benefits of the proposed method. This research has been published at IROS 2013.

6.7. Simulation of Lipofilling Reconstructive Surgery

We have developed a method to simulate the outcome of reconstructive facial surgery based on fat-filling. Facial anatomy is complex: the fat is constrained between layers of tissues which behave as walls along the face; in addition, connective tissues that are present between these different layers also influence the fat-filling procedure. To simulate the end result, we have proposed a method which couples a 2.5D Eulerian fluid model for the fat and a finite element model for the soft tissues. The two models are coupled using the computation of the mechanical compliance matrix. We had two contributions: a solver for fluids which couples properties of solid tissues and fluid pressure, and an application of this solver to fat-filling surgery procedure simulation. This research has been published at MICCAI 2013.

6.8. Real-time simulation of contact and cutting of heterogeneous soft-tissues

We have developed a new numerical method for interactive (real-time) simulations, which considerably improves the accuracy of the response of heterogeneous soft-tissue models undergoing contact, cutting and other topological changes. It provides an integrated methodology able to deal both with the ill-conditioning issues associated with material heterogeneities, contact boundary conditions which are one of the main sources of inaccuracies, and cutting which is one of the most challenging issues in interactive simulations. Our approach is based on an implicit time integration of a non-linear finite element model. To enable real-time computations, we propose a new preconditioning technique, based on an asynchronous update at low frequency. The preconditioner is not only used to improve the computation of the deformation of the tissues, but also to simulate the contact response of homogeneous and heterogeneous bodies with the same accuracy. We also address the problem of cutting the heterogeneous structures and propose a method to update the preconditioner according to the topological modifications. Finally, we have applied our approach to three challenging demonstrators: i) a simulation of cataract surgery ii) a simulation of laparoscopic hepatectomy iii) a brain tumor surgery. This research was done in collaboration with the University of Cardiff and has been published in the journal Media this year.

6.9. Control of Elastic Soft Robots

In this work, we present a new method for the control of soft robots with elastic behavior, piloted by several actuators. The central contribution of this work is the use of the Finite Element Method (FEM), computed in real-time, in the control algorithm. The FEM based simulation computes the nonlinear deformations of the robots at interactive rates. The model is completed by Lagrange multipliers at the actuation zones and at the end-effector position. A reduced compliance matrix is built in order to deal with the necessary inversion of the model. Then, an iterative algorithm uses this compliance matrix to find the contribution of the actuators (force and/or position) that will deform the structure so that the terminal end of the robot follows a given position. Additional constraints, like rigid or deformable obstacles, or the internal characteristics of the actuators are integrated in the control algorithm. We illustrate our method using simulated examples of both serial and parallel structures and we validate it on a real 3D soft robot made of silicone.

VISAGES Project-Team

6. New Results

6.1. Image Computing: Detection, Segmentation, Registration and Analysis

6.1.1. A Mathematical Framework for the Registration and Analysis of Multi-Fascicle Models for Population Studies of the Brain Microstructure

Participant: Olivier Commowick.

Diffusion tensor imaging (DTI) is unable to represent the diffusion signal arising from multiple crossing fascicles and freely diffusing water molecules. Generative models of the diffusion signal, such as multi-fascicle models, overcome this limitation by providing a parametric representation for the signal contribution of each population of water molecules. These models are of great interest in population studies to characterize and compare the brain microstructural properties. Central to population studies is the construction of an atlas and the registration of all subjects to it. However, the appropriate definition of registration and atlasing methods for multi-fascicle models have proven challenging. This paper proposes [32] a mathematical framework to register and analyze multi-fascicle models. Specifically, we define novel operators to achieve interpolation, smoothing and averaging of multi-fascicle models. We also define a novel similarity metric to spatially align multi-fascicle models. Our framework enables simultaneous comparisons of different microstructural properties that are confounded in conventional DTI. The framework is validated on multi-fascicle models from 24 healthy subjects and 38 patients with tuberous sclerosis complex, 10 of whom have autism. We demonstrate the use of the multi-fascicle models registration and analysis framework in a population study of autism spectrum disorder.

6.1.2. Multimodal rigid-body registration of 3D brain images using bilateral symmetry

Participants: Sylvain Prima, Olivier Commowick.

In this paper we show how to use the approximate bilateral symmetry of the brain with respect to its interhemispheric fissure for intra-subject (rigid-body) mono- and multimodal 3D image registration. We propose to define and compute an approximate symmetry plane in the two images to register and to use these two planes as constraints in the registration problem. This 6-parameter problem is thus turned into three successive 3-parameter problems. Our hope is that the lower dimension of the parameter space makes these three subproblems easier and faster to solve than the initial one. We implement two algorithms to solve these three subproblems in the exact same way, within a common intensity-based framework using mutual information as the similarity measure. We compare this symmetry-based strategy with the standard approach (i.e. direct estimation of a 6-parameter rigid-body transformation), also implemented within the same framework, using synthetic and real datasets. We show in [44] our symmetry-based method to achieve subvoxel accuracy with better robustness and larger capture range than the standard approach, while being slightly less accurate and slower. Our method also succeeds in registering clinical MR and PET images with a much better accuracy than the standard approach. Finally, we propose a third strategy to decrease the run time of the symmetry-based approach and we give some ideas, to be tested in future works, on how to improve its accuracy.

6.1.3. Distortion Correction in EPI Diffusion Weighted Images

Participants: Renaud Hedouin, Olivier Commowick.

We have compared and developed several methods which correct distortion of EPI images. The most popular method field map do not give optimal results. We have implemented and improved a method based on reversed phase encoding gradient which give good results. To correct diffusion weighted images this method only need one reversed phase encoding gradient $B0$ image which not need substantial additional acquisition time.

6.1.4. Using bilateral symmetry to improve non-local means denoising of MR brain images

Participants: Sylvain Prima, Olivier Commowick.

The popular NL-means denoising algorithm proposes to modify the intensity of each voxel of an image by a weighted sum of the intensities of similar voxels. The success of the NL-means rests on the fact that there are typically enough such similar voxels in natural, and even medical images; in other words, that there is some self-similarity/redundancy in such images. However, similarity between voxels (or rather, between patches around them) is usually only assessed in a spatial neighbourhood of the voxel under study. As the human brain exhibits approximate bilateral symmetry, one could wonder whether a voxel in a brain image could be more accurately denoised using information from both ipsi- and contralateral hemispheres. This is the idea we have investigated in this paper [45]. We define and compute a mid-sagittal plane which best superposes the brain with itself when mirrored about the plane. Then we use this plane to double the size of the neighbourhoods and hopefully find additional interesting voxels to be included in the weighted sum. We evaluate this strategy using an extensive set of experiments on both simulated and real datasets.

6.1.5. Detection of Multiple Sclerosis Lesions using Dictionary Learning

Participants: Hrishikesh Deshpande, Pierre Maurel, Christian Barillot.

Multiple sclerosis (MS) is a chronic, autoimmune, inflammatory disease of the central nervous system, in which certain areas of brain develop MS lesions, which are characterized by demyelination. Over the last years, various models combined with supervised and unsupervised classification methods have been proposed for detection of MS lesions using magnetic resonance images. Recently, signal modeling using sparse representations (SR) has gained tremendous attention and is an area of active research. SR allows coding data as sparse linear combinations of the elements of over-complete dictionary and has led to interesting image recognition results. The dictionary used for sparse coding plays a key role in the classification process. In this work, we have proposed to learn class specific dictionaries and develop new classification scheme, to automatically detect MS lesions in 3-D multi-channel magnetic resonance images.

6.1.6. Multiple Sclerosis Lesion Detection in Clinically Isolated Syndromes

Participants: Yogesh Karpate, Olivier Commowick, Christian Barillot.

Quantitative assessment of Multiple Sclerosis Lesions (MSL) in Clinically Isolated Syndromes (CIS) is important, as they are a precursor to subsequent stages of the disease. We address the problem of lesion patch detection with respect to Normally Appearing Brain Tissues (NABT). Our approach consists in learning rotationally invariant MSL and NABT multimodal intensity signatures based on 3D spherical gabor descriptors. This learning step, done once and for all, is followed by a testing step for the patient patches with an exemplar SVM. First, we develop a framework for selecting focused region of interest (fROI) using linear SVM for scoring. This allows an excellent trade-off between speed and accuracy. Second, building rotational invariant and scale independent features for accurate representation of image signatures. The extracted features are sensitive to the orientation of the analyzed image. This is a drawback in classification and retrieval applications. We handle this problem by using shperical Gabor descriptors. And last, we apply max pooling for down sampling of feature vectors. For the classification purpose we use a standard linear Support Vector Machine(SVM). The main contribution of the work is to build binary classifier to discriminate NABTs and MSLs based upon robust image representation. We have validated our approach on synthetic and real patient data. The synthetic lesion data is generated with noise, without noise and with bias field. Further, validation is carried out in three different scenarios. First, we evaluate our classifier using K-fold started with cross validation using NABT from healthy volunteers and MSL from CIS patients, then the detection of NABT and MSL from CIS patients on known patches is performed. The last evaluation concerned the full search algorithm.

6.1.7. Intensity Normalization in Longitudinal MS Patients

Participants: Yogesh Karpate, Olivier Commowick, Christian Barillot.

This work proposes a longitudinal intensity normalization algorithm for multi-channel MRI of brain of MS patient in the presence of lesions, aiming towards stable and consistent longitudinal segmentation. This approach is parametric and developed using two different forms of Robust Expectation Maximization (EM). The first is Spatio-Temporal Robust Expected Maximization (STREM) and other being EM with beta divergence. We validated our method on real longitudinal multiple sclerosis subjects.

6.2. Image processing on Diffusion Weighted Magnetic Resonance Imaging

6.2.1. Statistical Analysis of White Matter Integrity for the Clinical Study of Specific Language Impairment in Children

Participants: Olivier Commowick, Camille Maumet, Aymeric Stamm, Jean-Christophe Ferré, Christian Barillot.

Children affected by Specific Language Impairment (SLI) fail to develop a normal language capability. To date, the etiology of SLI remains largely unknown. It induces difficulties with oral language which cannot be directly attributed to intellectual deficit or other developmental delay. Whereas previous studies on SLI focused on the psychological and genetic aspects of the pathology, few imaging studies investigated defaults in neuroanatomy or brain function. We have proposed [53] to investigate the integrity of white matter in SLI thanks to diffusion Magnetic Res- onance Imaging. An exploratory analysis was performed without a priori on the impaired regions. A region of interest statistical analysis was performed based, first, on regions defined from Catani's atlas and, then, on tractography-based regions. Both the mean fractional anisotropy and mean apparent diffusion coefficient were compared across groups. To the best of our knowledge, this is the first study focusing on white matter integrity in specific language impairment. 22 children with SLI and 19 typically developing children were involved in this study. Overall, the tractography-based approach to group comparison was more sensitive than the classical ROI-based approach. Group differences between controls and SLI patients included decreases in FA in both the perisylvian and ventral pathways of language, comforting findings from previous functional studies.

6.2.2. Adaptive Multi-modal Particle Filtering for Probabilistic White Matter Tractography

Participants: Aymeric Stamm, Olivier Commowick, Christian Barillot.

Particle filtering has recently been introduced to perform probabilistic tractography in conjunction with DTI and Q-Ball models to estimate the diffusion information. Particle filters are particularly well adapted to the tractography problem as they offer a way to approximate a probability distribution over all paths originated from a specified voxel, given the diffusion information. In practice however, they often fail at consistently capturing the multi-modality of the target distribution. For brain white matter tractography, this means that multiple fiber pathways are unlikely to be tracked over extended volumes. We have proposed [51] to remedy this issue by formulating the filtering distribution as an adaptive M-component non-parametric mixture model. Such a formulation preserves all the properties of a classical particle filter while improving multi-modality capture. We apply this multi-modal particle filter to both DTI and Q-Ball models and propose to estimate dynamically the number of modes of the filtering distribution. We show on synthetic and real data how this algorithm outperforms the previous versions proposed in the literature.

6.2.3. Tracking the Cortico-Spinal Tract from Low Spatial and Angular Resolution Diffusion MRI

Participants: Aymeric Stamm, Olivier Commowick, Christian Barillot.

We have participated to the annual MICCAI workshop on DTI tractography [52]. We presented a pipeline to reconstruct the corticospinal tract (CST) that connects the spinal cord to the motor cortex. The proposed method combines a new geometry-based multi-compartment diffusion model coined Diffusion Directions Imaging and a new adaptive multi-modal particle filter for tractography. The DTI Tractography challenge proposes to test our methods in the context of neurosurgical planning of tumor removal, where very low spatial and angular resolution diffusion data is available due to severe acquisition time constraints. We took up the challenge and present our reconstructed CSTs derived from a single-shell acquisition scheme at b = 1000 s/mm2 with only 20 or 30 diffusion gradients (low angular resolution) and with images of 5 mm slice thickness (low spatial resolution).

6.3. Medical Image Computing in Brain Pathologies

6.3.1. Semi-Automatic Classification of Lesion Patterns in Patients with Clinically Isolated Syndrome

Participants: Olivier Commowick, Jean-Christophe Ferré, Gilles Edan, Christian Barillot.

Multiple sclerosis (MS) is neuro-degenerative disease of the Central Nervous System characterized by the loss of myelin. A Clinically Isolated Syndrome (CIS) is a first neurological episode caused by inflammation/demyelination in the central nervous system which may lead to MS. Better understanding of the disease at its onset will lead to a better discovery of pathogenic mechanisms, allowing suitable therapies at an early stage. We have proposed [37] an automatic segmentation algorithm for two different contrast agents, used within a framework for early characterization of CIS patients according to lesion patterns, and more specifically according to the nature of the inflammatory patterns of these lesions. We expect that the proposed framework can infer new prospective figures from the earliest imaging signs of MS since it can provide a classification of different types of lesions across patients. The lesion detection algorithm based on intensity normalization and subtraction of the used MRI data is a pivotal step, since it avoids the time-demanding task of manual delineation.

6.3.2. Multiple Sclerosis Lesions Evolution in Patients with Clinically Isolated Syndrome

Participants: Olivier Commowick, Jean-Christophe Ferré, Gilles Edan, Christian Barillot.

Multiple sclerosis (MS) is a disease with heterogeneous evolution among the patients. Some classifications have been carried out according to either the clinical course or the immunopathological profiles. Epidemiological data and imaging are showing that MS is a two-phase neurodegenerative inflammatory disease. At the early stage it is dominated by focal inflammation of the white matter (WM), and at a latter stage it is dominated by diffuse lesions of the grey matter and spinal cord. A Clinically Isolated Syndrome (CIS) is a first neurological episode caused by inflammation/demyelination in the central nervous system which may lead to MS. Few studies have been carried out so far about this initial stage. Better understanding of the disease at its onset will lead to a better discovery of pathogenic mechanisms, allowing suitable therapies at an early stage. We have proposed [36] a new data processing framework able to provide an early characterization of CIS patients according to lesion patterns, and more specifically according to the nature of the inflammatory patterns of these lesions. The method is based on a two layers classification. Initially, the spatio-temporal lesion patterns are classified using a tensor-like representation. The discovered lesion patterns are then used to identify group of patients and their correlation to 15 months follow-up total lesion loads (TLL), which is so far the only image-based figure that can potentially infer future evolution of the pathology. We expect that the proposed framework can infer new prospective figures from the earliest imaging sign of MS since it can provide a classification of different types of lesion across patients.

6.3.3. Arterial Spin Labeling at 3T in semantic dementia: perfusion abnormalities detection and comparison with FDG-PET

Participants: Isabelle Corouge, Jean-Christophe Ferré, Elise Bannier, Aymeric Stamm, Christian Barillot, Jean-Yves Gauvrit.

Arterial Spin Labeling (ASL) is a non invasive perfusion imaging technique which has shown great diagnosis potential in dementia. However, it has never been applied to semantic dementia (SD), a rare subtype of frontotemporal lobar degeneration characterized by the gradual loss of conceptual knowledge, which is actually explored by a now well established marker of SD: ^{18}fluorodeoxyglucose-positron emission tomography (FDG-PET) imaging. Although ASL and FDG-PET respectively measure perfusion and metabolism, they have been shown to be strongly correlated. In this work, we explore the ability of ASL to detect perfusion abnormalities in SD in comparison with FDG-PET. Using patients and healthy subjects data from an ongoing clinical study, we apply our analysis framework starting with visual comparison of ASL and FDG-PET, and focusing on ASL data preprocessing and statistical analysis at the individual and group level. Preliminary results yield concordant observations between ASL and FDG-PET as well as expected hypoperfusions in SD, namely in the left temporal lobe, thus suggesting the potential of ASL to assess perfusion impairments in SD.

6.4. Vascular Imaging and Arterial Spin Labelling

6.4.1. Patient-specific detection of perfusion abnormalities combining within-subject and between-subject variances in Arterial Spin Labeling.

Participants: Camille Maumet, Pierre Maurel, Jean-Christophe Ferré, Christian Barillot.

In this paper, patient-specific perfusion abnormalities in Arterial Spin Labeling (ASL) were identified by comparing a single patient to a group of healthy controls using a mixed-effect hierarchical General Linear Model (GLM). Two approaches are currently in use to solve hierarchical GLMs: (1) the homoscedastic approach assumes homogeneous variances across subjects and (2) the heteroscedastic approach is theoretically more efficient in the presence of heterogeneous variances but algorithmically more demanding. In practice, in functional magnetic resonance imaging studies, the superiority of the heteroscedastic approach is still under debate. Due to the low signal-to-noise ratio of ASL sequences, within-subject variances have a significant impact on the estimated perfusion maps and the heteroscedastic model might be better suited in this context. In this paper we studied how the homoscedastic and heteroscedastic approaches behave in terms of specificity and sensitivity in the detection of patient-specific ASL perfusion abnormalities. Validation was undertaken on a dataset of 25 patients diagnosed with brain tumors and 36 healthy volunteers. We showed evidence of heterogeneous within-subject variances in ASL and pointed out an increased false positive rate of the homoscedastic model. In the detection of patient-specific brain perfusion abnormalities with ASL, modeling heterogeneous variances increases the sensitivity at the same specificity level [24].

6.4.2. An a contrario approach for the detection of activated brain areas in fMRI

Participants: Camille Maumet, Pierre Maurel, Jean-Christophe Ferré, Christian Barillot.

BOLD functional MRI (fMRI) is now a widespread imaging technique to study task-related activity in the brain. However, getting the areas of activation at the individual subject level is still an open issue. The standard massively univariate statistical analysis is usually performed after smoothing the data and makes use of a single p-value for final thresholding of the results [1]. In group fMRI studies, the need for compensation of cross-subjects misregistrations clearly justifies the smoothing. However, at the individual level, where neat delineations of the activated areas are of interest, the use of Gaussian smoothing as a pre-processing step is more questionable. In this paper, we propose to study the ability of an a contrario approach, recently adapted for basal perfusion abnormalities detection [2], to correctly detect areas of functional activity [42].

6.4.3. Robust perfusion maps in Arterial Spin Labeling by means of M-estimators

Participants: Camille Maumet, Pierre Maurel, Jean-Christophe Ferré, Christian Barillot.

Non-invasive measurement of Cerebral Blood Flow (CBF) is now feasible thanks to the introduction of Arterial Spin Labeling (ASL) Magnetic Resonance Imaging (MRI) techniques. To date, due to the low signal-to-noise ratio of ASL, a single acquisition (pair of control/label scans) is not sufficient to estimate perfusion reliably. Instead, the acquisition is usually repeated several times and the perfusion information is calculated by averaging across the repetitions. However, due to its zero breakdown point, the sample mean is very sensitive to outliers. In this paper, we propose to compute ASL CBF maps using Huber's M-estimator, a robust statistical function that is not overly impacted by outliers. This method is compared to an empirical approach, introduced in [1], based on z-score thresholding [43].

6.4.4. Quantifying CBF from Aterial Spin Labeling via Diverse-TI: sampling diversity or repetitions ?

Participants: Lei Yu, Pierre Maurel, Christian Barillot.

Arterial Spin Labeling (ASL) is a noninvasive perfusion technique which allows the absolute quantification of Cerebral Blood Flow (CBF). The perfusion is obtained from the difference between images with and without magnetic spin labeling of the arterial blood and the captured signal is around 0.5-2% of the magnitude of the labeling images, so the noise is one of the main problems for further data analysis. Classical method, *Mono-TI*, for CBF quantification is averaging repetitions with only one Inversion Time (TI) - the time delay between labeling and acquisition to allow the labeled blood to arrive the imaging slice. It improves the robustness to noise, however, cannot compensate the variety of Arterial Arrival Time (AAT). In this paper, *Diverse-TI* is proposed to exploit different TI sampling instants (sampling diversity) to improve the robustness to variety of AAT and simultaneously average repetitions with each TI (sampling repetitions) to improve the robustness to noise. Generally, the sampling diversity is relatively small and can be considered as compressed measurements, thus the Compressive Matched Filter (CMF) enlightened from sparsity is exploited to directly reconstruct CBF and AAT directly from compressed measurements. Meanwhile, regarding the CBF quantification performance, the compromise between the sampling repetition and sampling diversity is discussed and the empirical protocol to determine the sampling diversity is proposed. Simulations are carried out to highlight our discussions. This is a joint work with Remi Gribonval (Panama Team) [56].

6.4.5. Peripheral angiography and neurovascular imaging

Participants: Hélène Raoult, Jean-Yves Gauvrit, Elise Bannier, Pierre Maurel, Clement Neyton, Christian Barillot, Jean-Christophe Ferré.

Vascular imaging contributions were performed on two different regions during the evaluation period: first on peripheral angiography, then on neurovascular imaging. Arteriography and MR angiography are routinely performed in patients presenting vascular pathologies. Yet, contrast agent injection is contraindicated in patients with renal insufficiency and the underlying risk of developing nephrogenic systemic fibrosis further encourages research on non-contrast enhanced MR angiography techniques (NCE MRA). In this context, we have been working on new MR sequences to reliably detect vascular abnormalities.

A first study [29] was published, where we assessed the feasibility and image quality of an improved non-gated carotid NATIVE TrueFISP NCE MRA sequence providing an extended field of view and a shorter acquisition time as compared to Time-of-Flight (TOF) imaging. A second study [48] was recently accepted for publication in Radiology on intracranial NCE MRA for arteriovenous malformation imaging with a high temporal resolution over 2 cardiac cycles. Combined with image post-processing, it allows improved depiction of venous drainage necessary to evaluate hemorrhagic risk and quantification. This ongoing work was just submitted.

ANGE Team

6. New Results

6.1. Geophysical flows

6.1.1. A numerical scheme for the Saint-Venant equations

Participants: Emmanuel Audusse, Christophe Chalons [Univ. Versailles], Philippe Ung.

In order to improve the numerical simulations of the shallow-water equations, one has to face three important issues related to the well-balanced, positivity and entropy-preserving properties, as well as the ability to handle vacuum states. In that purpose, we propose a Godunov-type method based on the design of a three-wave Approximate Riemann Solver (ARS) which satisfies all aformentioned properties.

6.1.2. Two-phase flows

Participants: Frédéric Coquel [CNRS], Jean-Marc Hérard [EDF], Khaled Saleh [IRSN], Nicolas Seguin.

After having developed numerical schemes for models of compressible two-phase flows [17], [19], we have proven some fundamental properties of these systems: symmetrizability and (non strict) convexity of the entropy [18]. This enables us now to address the well-posedness of these models when the relaxation terms are included.

6.1.3. Non-hydrostatic models

Participants: Marie-Odile Bristeau, Dena Kazerani, Anne Mangeney, Jacques Sainte-Marie, Nicolas Seguin.

The objective is to derive a model corresponding to a depth averaged version of the incompressible Euler equations with free surface. We have already contributed to this subject but the obtained results extend previous ones [29] in several directions:

- the derivation of the model is more rigorous and follows the entropy-based moment closures proposed in [28],
- the properties of the model and especially its connections with Green-Nagdhi model have been investigated,
- a family of analytical solutions for the proposed model have been obtained.

These analytical solutions emphasize the non-hydrostatic effects appearing for large slope variations.

6.1.4. Fluids with complex rheology

Participants: Anne Mangeney, Jacques Sainte-Marie.

We have been able

- to develop detection, characterization and localisation methods applicable to the seismic signals generated by rockfalls and thus to analyse ths spatio-temporal change of rockfall localisation and properties during several years, making it possible to show how rockfalls can be used as a precursor of volcanic activity,
- to propose an empirical “universal” law describing friction weakening in landslides over a broad range of volumes and geological contexts,
- to propose a new debris flow model with an energy balance,
- show the existence of a slow propagation phase in granular flows, playing a key role in their dynamics and in erosion processes.

6.1.5. Dynamics of sedimentary river beds with stochastic fluctuations

Participants: Emmanuel Audusse, Philippe Ung.

The Exner equation is a coarse model for the dynamics of sedimentary river beds, derived using both many heuristics and empirism. Though, it is also quite practical for hydraulic engineering applications, and efficient enough in numerous situations. Our goal in this work is to improve the model by including some effects that have been neglected so far in the heuristics. In particular, inline with other current research directions in the field, we study the possibility of introducing some stochasticity in the model. To this end, we suggest to numerically experiment some recently proposed variations of the Exner equation based on the introduction of stochastic fluctuations within the standard formulation.

This project has been the subject of a study during the 2013 session of the CEMRACS.

6.2. Ecology and sustainable energies

6.2.1. Hydrodynamic-biology coupling

Participants: Olivier Bernard [Inria BIOCORE], Anne-Céline Boulanger, Marie-Odile Bristeau, Raouf Hamouda, Jacques Sainte-Marie.

An important part of our research activity is built around a biological and industrial problem: the simulation of the coupling of hydrodynamics and biology in the context of industrial microalgae culture in outdoor raceways. The numerical modelling is adressed with the use of a multilayer vertical discretization of hydrostatic Navier-Stokes equations coupled with a light sensitive Droop model. Numerically, kinetic schemes allow for the development of efficient, positivity preserving, well balanced and entropy satisfying schemes. Simulations are carried out in 2D and 3D [1]. From a practical point of view, this model is capable of accounting for the utility of a paddlewheel and exhibits Lagrangian trajectories underwent by algae. Hence providing hints on the light history of algae in the pond, which is a key information to biologists, since it enables them to adapt their phytoplankton growth models to those particular, non natural conditions.

6.3. Coupling methods

6.3.1. Data assimilation for conservation laws associated with kinetic description

Participants: Anne-Céline Boulanger, Philippe Moireau [Inria M3DISIM], Jacques Sainte-Marie.

In order to take advantage of the kinetic description of conservation laws already used for the building of efficient schemes, an innovative data assimilation method for hyperbolic balance laws based in a Luenberger observer on the kinetic equation is developed. It provides a nice theoretical framework for scalar conservation laws, for which we study the cases of complete observations, partial observations in space, in time, and noisy observations. As far as systems are concerned, we focus on the Saint-Venant system, which is hyperbolic, nonlinear and has a topographic source term. We build an observer based only on water depths measurements. Numerical simulations are provided in the case of scalar laws and systems, in one and two dimensions, which validate the efficiency of the method [14].

6.3.2. Mach-parametrized flows

Participants: Stéphane Dellacherie [CEA], Bruno Després [UPMC Paris 6], Yohan Penel.

In order to enrich the modelling of fluid flows, we investigate in this paper a coupling between two models dedicated to distinct regimes. More precisely, we focus on the influence of the Mach number as the low Mach case is known to induce theoretical and numerical issues in a compressible framework. A moving interface is introduced to separate a compressible model (Euler with source term) and its low Mach counterpart through relevant transmission conditions. A global steady state for the coupled problem is exhibited. Numerical simulations are then performed to highlight the influence of the coupling by means of a robust numerical strategy [20].

6.3.3. Error analysis in a coupling strategy

Participants: Clément Cancès [UPMC Paris 6], Frédéric Coquel [CNRS], Edwige Godlewski, Hélène Mathis [Univ. Nantes], Nicolas Seguin.

We have proposed in a simplified framework an error analysis for an adaptive method which automatically selects the optimal model to use, the choice being between a reference model and an associated simplified one, see [15]. In particular, we are able to balance the thickness of the coupling buffer zone with the threshold on the modelling error which appears when introducing the coarse model.

6.4. Software development and assessments

6.4.1. Analytical solutions for the incompressible Euler system

Participants: Anne-Céline Boulanger, Marie-Odile Bristeau, Jacques Sainte-Marie.

We have proposed in [5] a large set of analytical solutions (FRESH-ASSESS) for the hydrostatic incompressible Euler system in 2d and 3d. These solutions mainly concern free surface flows but partially free surface flows are also considered. These analytical solutions can be especially useful for the validation of numerical schemes.

6.4.2. Software

Several tasks have been achieved in the FRESHKISS3D software (§ 5.1):

- First tests with a uniform pression before moving to the variable case;
- Rethinking of the C++ code with an object-oriented rewriting which provides a better memory management;
- Automatic boundary conditions handling in the case of a fluid/solid transition;
- New computations of the particule trajectories when they leave out the domain by means of directional interpolation procedures;
- Achievement of the 2nd-order space accuracy;
- Taking into account the wind.

BANG Project-Team

6. New Results

6.1. Proliferation dynamics and its control

6.1.1. Proliferation dynamics in cell populations

Participants: José Luis Avila Alonso [DISCO project-team, Inria Saclay IdF], Annabelle Ballesta, Gregory Batt [CONTRAINTES project-team], François Bertaux, Frédérique Billy, Frédéric Bonnans [Commands project-team, Inria Saclay IdF], Catherine Bonnet [DISCO project-team, Inria Saclay IdF], Jean Clairambault, Marie Doumic, Xavier Dupuis [Commands project-team], Ján Eliaš, Germain Gillet [IBCP, Université Cl. Bernard Lyon 1], Pierre Hirsch [INSERM Paris (Team18 of UMR 872) Cordeliers Research Centre and St. Antoine Hospital, Paris], Pierre Magal [University Bordeaux II], Anna Marciniak-Czochra [Institute of Applied Mathematics, Universität Heidelberg], Jean-Pierre Marie [INSERM Paris (Team18 of UMR 872) Cordeliers Research Centre and St. Antoine Hospital, Paris], Roberto Natalini [IAC-CNR, Università Sapienza, Rome], Silviu Niculescu [DISCO project-team, Inria Saclay IdF], Hitay Özbay [Bilkent University, Ankara, Turkey], Benoît Perthame, Szymon Stoma [CONTRAINTES project-team], Ruoping Tang [INSERM Paris (Team18 of UMR 872) Cordeliers Research Centre and St. Antoine Hospital, Paris], Vitaly Volpert [CNRS Lyon, UMR5208, Camille Jordan Institute, Lyon], Jorge Zubelli [IMPA, Rio de Janeiro].

1. **Transition kernels in a McKendrick model of the cell division cycle.** This theme, after a rich harvest of publications (most of them in 2013 and even 2014), is awaiting new developments, since of the main two young researchers on this theme, F. Billy has concluded her 2-year Inria postdoc at Bang, leaving for an industrial company in November 2012, while O. Fercoq (team MaxPlus, Saclay) has defended his PhD thesis at École Polytechnique in September 2012, only to leave for a postdoc position dedicated to optimisation theory in Edinburgh.
2. **Modelling haematopoiesis with applications to AML.** This theme has been active through a collaboration with Inria teams Commands (F. Bonnans, X. Dupuis) and Disco (J.L. Avila Alonso, C. Bonnet, Hitay Özbay, S. Niculescu), and J.-P. Marie's team at St Antoine Hospital leukaemic tumour bank, where A. Ballesta, Cancéropole IdF-Inria postdoc has been detached (ending in January 2013) to identify parameters of a model of acute myeloblastic leukaemia (AML) in patient fresh cell cultures with and without anticancer drugs. This work has led to several presentations, and publications are in preparation. In a book chapter summing up the PhD work of J.L. Avila Alonso [26], and in two submitted conference papers [28], [29], a new model of haematopoiesis for AML is presented, including phases of the cell division cycle and maturation stages, with targets for therapeutic control.
3. **Hybrid models.** Systems combining PDEs and discrete representations in hybrid models, with applications to cancer growth and therapy, in particular for AML, are the object of study of the ANR program *Bimod*, coordinated by V. Volpert (Lyon), associating CNRS (V. Volpert, Lyon), Bordeaux II University (P. Magal) and the Bang project-team.
4. **Molecular model of apoptosis.**

   With G. Gilllet (professor at IBCP/Lyon), A. Ballesta and M. Doumic have designed a mathematical ODE model for the mitochondrial pathway of apoptosis, focused on the early phase of apoptosis (before the cytochrome C release). This model has been validated by experimental data carried out in G. Gillet's lab and applied to propose new therapeutic strategies against cancer [6].
5. **Molecular model of the activity of the p53 protein.** This work, firstly the object of Luna Dimitrio's PhD thesis [37], who left in 2012 for the pharmaceutic industry (SANOFI), has been continued since a new PhD student, Ján Eliaš, has taken over this theme in September 2012 in a new PhD thesis at UPMC, under the supervision of J. Clairambault and B. Perthame. His work has given rise in 2013 to 2 publications [14], [32].

6. **TRAIL - induced apoptosis in HELA cells** Explaining cell-to-cell variability is a major step towards understanding how cancer cells escape action of chemotherapeutic drugs. We set up and studied an integrated model of stochastic gene expression, deterministic translation and protein degradation capable of explaining fractional killing and reversible resistance in Hela cells in response to treatment with TNF-Related Apoptosis Inducing Ligand, TRAIL (Bertaux, Stoma, Drasdo, and Batt, submitted). The results of the model suggests that stochastic fluctuations are a fundamental determinant in understanding cell-to-cell variability, and identified relations between the characteristic time scales of the processes at which stochasticity should play a particular important role.

6.1.2. Physiological and pharmacological control of cell proliferation

Participants: Annabelle Ballesta, Frédérique Billy, Jean Clairambault, Sandrine Dulong [INSERM Villejuif (U 776)], Olivier Fercoq [MaxPlus project-team], Stéphane Gaubert [MaxPlus project-team], Thomas Lepoutre [Dracula project-team], Francis Lévi [INSERM Villejuif (U 776)].

1. *Periodic (circadian) control of cell proliferation in a theoretical model of the McKendrick type.* This theme (cf. supra “transition kernels...”) has been continued [9], [27], [7], [8], [31]. Whereas transition kernels between cell cycle phases without control have been experimentally identified in cell cultures by FUCCI imaging [9], their circadian control remains elusive and has been modelled on the basis of gating by plain cosines representing the influence exerted on these transition kernels by circadian clocks. To go further, it would be necessary to have access by cell imaging to the activity of the best physiological candidates to such gating, namely the cyclin-Cdk complexes, together with the activities of the clock-controlled proteins Wee1 and p21, which thus far have remained unavailable to us through biological experimentation with imaging. A 12-year collaboration work with Francis Lévi on (circadian) chronotherapeutic optimisation in cancer is reported in [30].
2. *Intracellular pharmacokinetic-pharmacodynamic (PK-PD) models for anticancer drugs.* This theme has continued to be developed with new publications for the drugs irinotecan [5], 5-fluorouracil and oxaliplatin [31], and with a recent mini-review by A. Ballesta and J. Clairambault on mathematical models of treatment of metastatic colorectal cancer [4].

6.1.3. Optimisation of cancer chemotherapy and cancer radiotherapy

Participants: Juan Carlos Alfonso [University Complutense, Madrid, Spain], Annabelle Ballesta, Frédérique Billy, Frédéric Bonnans [Commands project-team], Rebecca Chisholm, Jean Clairambault, Sandrine Dulong [INSERM Villejuif (U 776)], Xavier Dupuis [Commands project-team], Alexandre Escargueil [INSERM and UPMC, St Antoine Hospital], Olivier Fercoq [MaxPlus project-team], Stéphane Gaubert [MaxPlus project-team], Miguel Angel Herrero [University Complutense, Madrid, Spain], Michael Hochberg [ISEM, CNRS, Montpellier], Dirk Drasdo, Nick Jagiella, Francis Lévi [INSERM U 776, Villejuif], Thomas Lepoutre [Dracula project-team], Tommaso Lorenzi, Alexander Lorz, Luis Núñez [University Complutense, Madrid, Spain], Benoît Perthame, Emmanuel Trélat [LJLL, UPMC].

1. **Limiting unwanted toxic side effects: age-structured models of the cell cycle.** Optimising cancer chemotherapy, in particular chronotherapy, is the final aim of these activities. A classical numerical method of optimization under the constraint of limiting toxicity to healthy tissues has been applied to the McKendrick model of the cell cycle divided in phases, endowed with physiologically based targets for both internal (circadian) and external (pharmacological) control. This model has been partly biologically identified on continuous FUCCI recordings of proliferating NIH3T3 cells in culture media; these data were made available to us within the C5Sys consortium, an ERASYSBIO+ European project. Then additional theoretical characteristics establishing hypothetical differences between healthy and cancer cell populations, relying on different responses to physiological circadian clock influences on gating by Cyclin-Cdk complexes between cell cycle phases, have been used to solve the optimization problem, proposing an optimal drug infusion regimen [7], [8], [27], [9]. Using an even more complex McKendrick-like model of the cell cycle, a connection with previously established PK-PD ODE models of the anticancer drugs 5-Fluorouracil and Oxaliplatin has

been established, proposing optimized combined drug delivery flows to solve the same optimization problem [31].

2. **Limiting drug resistance in cancer cell populations: cell Darwinism.** This theoretical activity has been continued also in more general settings taking into account another major issue of anticancer treatment, namely resistance to drugs in cancer cells. To this latter aim, we have developed another type of models based on integro-differential equations, which are inspired from those used in ecology for Darwinian evolution [22]. These are aimed at studying another major issue in cancer therapy: appearance of resistances to treatment in tumour cell populations. Indeed, these cell populations, because of their heterogeneity and genomic instability, present an ability to adapt and evolve (in the Darwinian sense) that is much higher than in healthy cell populations [7], [18], [35]. The time scales under investigation, much shorter than in ecology, are however much longer than in microbiology, and are those of clinical treatments. Theoretical optimization of external controls representing combined cytotoxic and cytostatic treatments on these models with the aim to limit the emergence of drug resistance are presently under assessment, in collaboration with Emmanuel Trélat (LJLL, UPMC), paper in preparation.
3. **Molecular aspects: ABC transporters.** From a molecular point of view, studying drug resistance leads to the study of ABC transporters, which is one of the tracks followed by A. Ballesta, following her PhD thesis, in collaboration with F. Lévi's INSERM team in Villejuif [4], [5].
4. **Optimisation of cell kill in AML.** Underway is also the use of methods of optimal control methods developed by the Commands project-team (Frédéric Bonnans, Xavier Dupuis) to optimise therapies in the treatment of Acute Myeloblastic Leukaemia (AML). X. Dupuis has lately produced a paper [40], accepted for publication in Math. Mod. Phys. Phenom, on optimisation of a combined treatment using a cytotoxic drug (representing Aracytin) and a cytostatic drug (representing AC220, an antagonist of Flt-3 receptors). This work is led in conjunction with the DISCO team, cf. supra "Modelling haematopoiesis with applications to AML").
5. **Estimating dose painting effects in radiotherapy: a mathematical model.** Tumor heterogeneity is widely considered to be a determinant factor in tumor progression and in particular in its recurrence after therapy. Unfortunately, current medical techniques are unable to deduce clinically relevant information about tumor heterogeneity by means of non-invasive methods. As a consequence, when radiotherapy is used as a treatment of choice, radiation dosimetries are prescribed under the assumption that the malignancy targeted is of a homogeneous nature. In this work we discuss the possible effects of different radiation dose distributions on heterogeneous tumors by means of an individual cell-based model. To that end, a case is considered where two tumor cell phenotypes are present, which strongly differ in their respective cell cycle duration and radiosensitivity properties. We show herein that, as a consequence of such differences, the spatial distribution of such phenotypes, as the resulting tumor heterogeneity, can be predicted as growth proceeds. As a consequence, heterogeneous dosimetries can be selected to enhance tumor control by boosting radiation in the region occupied by the more radioresistant tumor cell phenotype. It is also shown that, when compared with homogeneous dose distributions as those being currently delivered in clinical practice, such heterogeneous radiation dosimetries fare always better than their homogeneous counterparts (Alfonso et. al., PLoS One accepted [3]).

6.1.4. Protein polymerisation and application to amyloid diseases

Participants: Annabelle Ballesta, Vincent Calvez [ENS Lyon], Marie Doumic, Pierre Gabriel, Hadjer Wafaâ Haffaf, Benoît Perthame, Stéphanie Prigent [BPCP, INRA Jouy-en-Josas], Human Rezaei [BPCP, INRA Jouy-en-Josas], Léon Matar Tine [SIMPAF project-team, Inria Lille Nord-Europe].

Published in PLoS One in collaboration with the team of biologists led by H. Rezaei [44], a new and very complete PDE model for protein polymerisation has been designed. Following F. Charles's work, A. Ballesta has applied this model to Huntington's disease (PolyQ expansion) and compared it with its ODE counterpart, leading to a better understanding of the leading mechanisms responsible for PolyQ fibrillisation. New applications of this framework model are in progress with H.W. Haffaf and S. Prigent.

The eigenvalue problem playing a major role in the representation of Prion proliferation dynamics and, in a more general way, of many fragmentation-coalescence phenomena, the article [36] investigated the dependency of the principal eigenvector and eigenvalue upon its parameters. We exhibited possible nonmonotonic dependency on the parameters, opposite to what would have been conjectured on the basis of some simple cases.

6.1.5. Inverse problem in growth-fragmentation equations

Participants: Marie Doumic, Marc Hoffmann [ENSAE], Nathalie Krell [Univ. Rennes I], Patricia Reynaud [CNRS, Nice Univ.], Lydia Robert [UPMC], Vincent Rivoirard [Paris IX Univ.], Léon Matar Tine [SIMPAF project-team, Inria Lille Nord-Europe].

In collaboration with statisticians (M. Hoffman, Professor at Université de Marne-la-Vallée, V. Rivoirard, MC at Université d'Orsay, and P. Reynaud, CR CNRS at Université de Nice), in the article [38] published in SIAM Num. Anal., we explored a statistical viewpoint on the cell division problem. In contrast to a deterministic inverse problem approach, we take the perspective of statistical inference. By estimating statistically each term of the eigenvalue problem and by suitably inverting a certain linear operator, we are able to construct an estimator of the division rate that achieves the same optimal error bound as in related deterministic inverse problems. Our procedure relies on kernel methods with automatic bandwidth selection. It is inspired by model selection and recent results of Goldenschluger and Lepski.

An extension of this work, which consists of the statistical estimation of a branching process modelling the same growth and fragmentation dynamics, has been submitted in [12], in collaboration with N. Krell, M. Hoffmann and L. Robert. Such methods are indeed successfully applied to investigate bacterial growth, in collaboration with L. Robert (INRA and UPMC), see Figure 1 .

In [13], we generalised the inverse techniques proposed previously in [39], [43], in order to adapt them to general fragmentation kernels and growth speeds. The potential applications of this problem are numerous, ranging from polymerisation processes to the cell division cycle. An extension of this work, using refined estimates the Mellin transform of the equation, has just been accepted for publication in *Inverse Problems* [10].

6.2. Tissue growth, regeneration and cell movements

6.2.1. Chemotaxis, self-organisation of cell communities (KPP-Fisher and Keller-Segel)

Participants: Luís Lopes Neves de Almeida, Nikolaos Bournaveas [Univ. Edinburgh], Axel Buguin [UPMC, Institut Curie], Vincent Calvez [ENS Lyon], Casimir Emako-Kazianou, François James [univ. Orléans], Alexander Lorz, Grégoire Nadin [UPMC], Benoît Perthame, Jonathan Saragosti [Institut Curie], Pascal Silberzan [Institut Curie], Min Tang [Shanghai Jiaotong University], Nicolas Vauchelet.

Chemotaxis denotes the ability of some cells to undergo a directed movement in response to an extracellular chemical substance. A mathematical description of chemotaxis is a major issue in order to understand collective movements of bacterial colonies. Numerous mathematical models, at various scales, have been proposed, allowing for a good description of cell aggregation under chemotaxis at the macroscopic level, the first of all being that of Keller-Segel (1971), that is now at the centre of an abundant international scientific literature.

At the cell scale, one uses kinetic equations for which numerical simulations have been performed. Behaviour of solutions can be understood by performing a hydrodynamical limit of the kinetic equation. It leads to aggregation type equations for which finite time blow up is observed [42]. Then measure solutions for this system should be considered. A theoretical framework for the existence of weak solutions has then been developed [17], [34] where duality solutions for such system has been investigated which are equivalent to gradient flow solutions [33].

`../../../../projets/bang/IMG/FigureLydia.jpg`

Figure 1. Age and Size Distribution of a bacterial culture (E. coli): comparison between the experimental distribution (A) and the best-fit simulation (B). The methods developed in [38] and [12] allowed us to discriminate between a size-dependent and an age-dependent division rate.

Our understanding of traveling waves has progressed considerably in three directions: fitting continuous models and IBMs [21], fitting precisely models with experiments based on known biological values of parameters, and opening new paradigms: traveling waves can connect a dynamically unstable state to a Turing unstable state, certainly the stable wave connects the unstable state to a pulsating state.

6.2.2. Single-cell-based and continuum models of avascular tumours

Participants: Ibrahim Cheddadi, Dirk Drasdo, Benoît Perthame, Min Tang [Shanghai Jiaotong University], Nicolas Vauchelet, Irène Vignon-Clémentel [REO project-team].

The recent biomechanical theory of cancer growth considers solid tumours as liquid-like materials comprising elastic components. In this fluid mechanical view, the expansion ability of a solid tumour into a host tissue is mainly driven by either diffusion of cells (emerging on the mesoscopic scale by coarse graining from the cell micro-motility) or by cell division depending either on the local cell density (contact inhibition), on mechanical stress in the tumour, or both. For the two by two degenerate parabolic/elliptic reaction-diffusion system that results from this modelling, we prove there are always travelling waves above a minimal speed and we analyse their shapes. They appear to be complex with composite shapes and discontinuities. Several small parameters allow for analytical solutions; in particular the incompressible cells limit is very singular and related to the Hele-Shaw equation. These singular travelling waves are recovered numerically. See [21]. Besides this work, a direct comparison with agent-based and continuum models has been performed, showing very good agreement over a large parameter range.

6.2.3. Single cell-based models of tumour growth, tissue regeneration

Participants: Gregory Batt [CONTRAINTES project-team], François Bertaux, Noémie Boissier, Kai Breuhahn [German Cancer Centre, Heidelberg], Petru Bucur [Hopital Paul Brousse, Paris], Géraldine Cellière, Chadha Chettaoui, Ibrahim Cheddadi, Dirk Drasdo, Adrian Friebel, Rolf Gebhardt [Univ. of Leipzig, Germany], Adriano Henney [Director Virtual Liver Network and VLN consortium], Jan G. Hengstler [Leibniz Research Centre, Dortmund, Germany and CANCERSYS consortium], Stefan Höhme [Research Associate, University of Leipzig], Elmar Heinzle [University of Saarbrücken and NOTOX consortium], Nick Jagiella, Ursula Klingmüller [German Cancer Centre, Heidelberg and LungSys Consortium], Pierre Nassoy [Institut Curie, Paris and Univ. of Bordeaux], Johannes Neitsch, Benoît Perthame, Jens Timmer [University of Leipzig, Germany], Irène Vignon-Clémentel [REO project-team], Paul Van Liedekerke, Eric Vibert [Hôpital Paul Brousse, Villejuif], Ron Weiss [MIT, USA].

1. **Ammonia metabolism in healthy and damaged liver** The model on ammonia detoxification in liver, integrating a compartment model for the glutamine synthetase-active peri-central and the glutamine-inactive peri-portal liver lobule compartment (see Bang report 2012) with the spatial - temporal model of liver regeneration after drug-induced peri-central damage [41] has been extended to include the mass balance of other body compartments. The analysis shows that some body compartments that in the healthy liver produce ammonia, in the damaged liver detoxify blood from ammonia. The detoxification model of liver in combination with the body ammonia balance can be found in ref. (Schliess et. al., Hepatology, accepted [20]).
2. **Drug metabolism in hepatocytes** Since the begining of 2013 animal experiments for testing of cosmetics are forbidden within the EU. This has triggered initiatives towards how modeling may help to investigate drug toxicity, circumventing animal testing. The basic conceptual idea is to test drugs (cosmetics, perspectively also other drugs) in in-vitro systems such as monolayers, sandwich cultures, or multi-cellular spheroids, and use the emerging data to infer the expected toxicity in-vivo using novel experimental and computational approaches [16]. We have integrated an intracellular mathematical model of paracetamol drug metabolism in a mathematical agent-based cell model for monolayer and multi-cellular spheroids and compared simulation results with experimental findings in the same systems. We find that cell-to-cell variabilty can largely explain the experimentally observed cell population survival fractions. The mathematical model is now refined based on measurements of intermediate drug metabolites.

3. **Cell mechanics and its impact on cell proliferation** A novel numerical methodology has been developed to simulate the mechanics of cells and tissues using a continuum approach. Analogously to the Center Based Models, particles are used to represent (parts of) the cells but rather than discrete interactions they represent a continuum. This approach can be used for tissue mechanics simulations in where the individual cell-cell interactions are discarded but instead a constitutive law is proffered [23].
   Moreover, a new model in where cell adhesion dynamics is addressed. The cell model is constructed by a triangulated surface and a coarse-grained internal scaffolding structure. A model cell can adapt to realistic cell shapes, and is able to interact with a substrate or other cells. The parameters in this model can be determined by canonical experiments performed on cells informing about cell deformation, compression and cell-cell adhesion [19].
   A computational model for the confined growth of cells in a capsule has been developed. This model represents a realistic simulation tool for a novel experimental system (Institut Curie, Prof P. Nassoy) in where cells are grown in an elastic environment to mimic the effects of mechanical stress on cells and while monitoring their fate. Model parameter calibration is now ongoing to reproduce the correct quantitative behavior of the cells in order to unravel the relationship between cell mechanical stress and cell behavior.
4. **Playing the game of life with yeast cells** Within a collaboration with a synthetic biology lab at MIT, multicellular modelling of engineered yeast cell populations is performed. Those cells secrete a messenger molecule (IP) which diffuse in the medium, bind to other cells, and trigger a signalling cascade, which finally induces expression of lethal genes. A model has been established based on our single-cell-based model framework associated with PDE simulations, and it is currently used to explain and guide experiments conducted at the MIT. In 2013, the project has achieved significant progress on several aspects. First, we were able to quantitatively reproduce newly produced, rich data on the signaling cascade behavior with a kinetic model describing signaling reactions. Second, comparison between simulations and data allowed to identify key characteristics of the death module, which is positioned downstream of the signaling cascade: there is a rapid and stochastic commitment to death, followed by a deterministic and long delay (2-4 cell generations) needed before cells actually die. Finally, data production and analysis iterations with our collaborators alloweded to optimize the procedures for experimental measurements and the quantitative analysis of data in a synergistic manner.
5. **Other projects in short** Further progress have been achieved on the reconstruction of lung cancer micro-architecture from bright field micrographs. In partial hepatectomy (PHx), pig data on the changes of microarchitecture during regeneration after PHx have been generated and stained now being processed. The image processing chain for liver architecture reconstruction has been refined and extensive analysis has been performed on the architecture of the bile canaliculi network in healthy liver and in disease states of liver. Moreover, non-small-cell lung cancer cell invasion pattern have been analyzed leading to interesting observations now being studied by modelling.
   For multi-scale modeling of liver regeneration after drug-induced pericentral damage, integration of a molecular model of hepatocyte growth factor signalling with an agent-based model of liver regeneration has been extended to include blood flow in the lobule, as well as the contributions of the body compartment to the degradation and production of hepatocyte growth factor (HGF).

6.2.4. Modelling flows in tissues

Participants: Noémie Boissier, Lutz Brusch [TU Dresden], Dirk Drasdo, Adrian Friebel [IZBI, University of Leipzig], Stefan Hoehme [IZBI, University of Leipzig], Nick Jagiella [Inria and IZBI, University of Leipzig], Hans-Ulrich Kauczor [University of Heidelberg, Germany], Fabian Kiessling [University Clinics, Technical University of Aachen, Germany], Ursula Klingmueller [German Cancer Research Centre (DKFZ), Heidelberg, Germany], Hendrik Laue [Fraunhofer Mevis, Bremen, Germany], Ivo Sbazarini [MPI for Molecular Cell Biology and Genetics, Dresden, Germany], Irène Vignon-Clémentel [REO project-team], Marino Zerial [MPI for Molecular Cell Biology and Genetics, Dresden, Germany].

1. **Flow and perfusion scenarios in cancer.** We started reconstruction of the blood vessel system of lung cancers removed by surgery. For this purpose, patients underwent DCE-MRI prior to surgery. Part of the tumors after surgery was sliced and stained for nuclei, proliferation and endothelial cells. The slice data were recorded (Mevis, Luebeck) to allow identification of the position of the indiviudal structures in 3D space. The structures were then segmented. The work turned out to be particularly challenging because of staining artifacts for which image algorithms had to correct for. Nevertheless, last results look promising so that at least the network formed by larger vessels can be segmented and reconstructed in 3D. The so emerging data will be used for modeling of blood flow using the models developed in 2012.
2. **Flow in liver lobules.** We integrated blood flow in the new software CellSys (see above under software) and refined the algorithms. Moreover, we increased the resolution of the capillaries by triangulating them from high resolution confocal scanning micrographs.

6.2.5. Contraction of acto-myosin structures in morphogenesis and tissue repair

Participants: Luís Lopes Neves de Almeida, P. Bagnerini [Univ. Genova], A. Habbal [Univ. Nice], A. Jacinto [CEDOC, Lisbon], M. Novaga [Univ. Padova], A. Chambolle [École Polytechnique].

In 2013 we continued to investigate the dependence of physical and biological mechanisms of actomyosin cable formation and wound closure depending on the geometry of the wound, with particular emphasis on the effect of the wound edge curvature.

When the actomyosin cable starts to contract and the wound starts to close we have noticed that the behavior of the cable is related with the local curvature of the wound edge. This led us to study the curves evolving by positive part of their curvature in a Euclidean framework. A model where we consider viscous behavior and friction in the tissue plus boundary terms associated to cable and lamellipodial forces is under development. The numerical simulations obtained using this model are in good agreement with the previous experimental results and we are pursuing the model development by challenging it with new experiments.

6.3. Neurosciences

Participants: Jonathan Touboul, Gilles Wainrib, Tanguy Cabana, Mathieu Galtier, Luis Garcia Del Molino, Khashayar Pakdaman.

We pursued our studies of disordered networks of the brain and collective phenomena in neuroscience. We have been more interested this year in the role of disorder in the spontaneous emergence of synchronized activity. In order to study these phenomena, we have been establishing limit equations for randomly coupled networks [11], and the analysis of this equation reveal a number of transitions due to the level of disorder in the connectivity. A universal transition observed in such randomly coupled networks is a transition to chaotic activity for large levels of noise. These transitions were investigated [24] and were shown to be related to an explosion of complexity at the edge of chaos, i.e. the number of equilibria is exponentially large with the network size at the phase transition, and the exponential factor was related to the Lyapunov exponent. These large-scale limits give rise to nonlinear reduced equations that we have been introducing in [15]. Eventually, when considering that the network is structured into different populations and that the connectivity weights satisfy a balance condition, which is postulated as a natural scaling of the synaptic input, we have shown that the network shows random transitions to periodic activity depending on the spectrum of the random connectivity matrix [25], yielding up and down states or synchronized oscillations depending on the eigenvalue of larger real part of the connectivity matrix.

CASTOR Team

6. New Results

6.1. High order approximation of the two fluid Braginskii model

Participants: Sebastian Minjeaud, Richard Pasquetti.

We work on a two fluid physical model developed in close connection with Ph. Ghendrih (IRFM). It is based on the electrostatic assumption, i.e. the magnetic field is given (the magnetic field induced by the plasma itself is negligible), on the hypothesis of electroneutrality (the density of ions and electrons are proportional) and on the Braginskii closure. On the basis of the conservation equations of density, electron and ion velocities, electron and ion temperatures and electrical charges, a set of 10 non-linear and strongly anisotropic coupled partial differential equations (PDE) can be set up. A high order Fourier-SEM (Spectral Element Method) code is currently developed to address this problem. This Fourier-SEM code is close to be operational for the full set of PDEs in a 3D toroidal geometry. The torus section is discretized with quadrangular elements and Fourier expansions are used in the toroidal direction. In time one uses an RK3 (third order Runge-Kutta) IMEX (Implicit-Explicit), so that the Lorentz terms are handled implicitly. The capability of this code to handle a strongly anisotropic diffusion in a 3D toroidal geometry has already been tested. The Braginskii closure has been implemented. The Bohm boundary conditions at the plates are also considered. In 2013, we worked on a parallel version of the code and on the robustness of our algorithms, to be able to make long time computations, e.g. a few hundreds of thermal times.

6.1.1. Parallelization of the full Braginskii (FBGKI) code.

A first parallel version of the FBGKI code is now operational. Tests were made on the Computional center of the University of Nice-Sophia Antipolis. Tests on a large number of processors have however not yet been carried out, since presently our goal is to improve the robustness of our algorithms. Our parallelization strategy is based on a domain decomposition technique in the poloidal plane, where the spectral element approximation is local. On the contrary, in toroidal direction the approximation is global since based on Fourier expansions.

6.1.2. Numerical stabilization.

A spectral vanishing viscosity (SVV) technique was implemented in the sequential version of the code. Such a stabilization technique relies on the idea of introducing viscosity in the high frequency range of the spectral element / Fourier approximation. Such an approach was first proposed for hyperbolic problems, typically the Burgers equation (E. Tadmor, 1989). Later on, it was used for the large-eddy simulation of turbulent flows. Thus, we have a large experience of the SVV stabilization for the computation of turbulent wake flows.

6.1.3. Projection techniques.

A projection technique is used to enforce the divergence free constraint of the current. Projection techniques have been developed for a long time, in the frame of the Navier-Stokes equations to provide efficient algorithms when incompressible flows are concerned. For the Braginskii system, it appears natural to make use of such techniques for the current. Different projection techniques have been implemented in the FBGKI code, from the most classical one (Chorin-Temam, 1969) to the most recent. It turns out however that using projection techniques is less straightforward for Braginskii than for Navier-Stokes. We actively work on this point in order to cure some not yet understood failures of convergence with the time-step.

In the frame of the Eurofusion program, it is planned to check this version of the code on a simple configuration proposed by the EPFL (Paolo Ricci) where experimental as well as numerical results are available.

6.2. Equilibrium reconstruction and current density profile identification

Participants: Jacques Blum, Cédric Boulbe, Blaise Faugeras.

EQUINOX is a real-time equilibrium reconstruction code. It solves the equation satisfied by the poloidal flux in a computation domain, which can be the vacuum vessel for example, using a P1 finite element method and solves the inverse problem of the identification of the current density profile by minimizing a least square cost-function. It uses as minimal input the knowledge of the flux and its normal derivative on the boundary of the computation domain. It can also use supplementary constraints to solve the inverse problem: interferometric, polarimetric and MSE measurements. Part of the work reported here has been done in the frame of a RTM-JET contract [2]

6.2.1. Direct use of the magnetic measurements

The code EQUINOX was not originally designed to take as magnetic inputs directly the magnetic measurements, as it should be the case in the ITM (Integrated Tokamak Modeling European platform), but some outputs from the real-time codes APOLO at ToreSupra and XLOC at JET. These codes provide EQUINOX with the values of the flux and its normal derivative on a closed contour defining the boundary of the computation domain (this contour can be the limiter for example). As a consequence the main difficulty arising in the objective of integrating the code EQUINOX in the ITM structure is to interpolate between the magnetic measurements (flux loops and poloidal B-probes) with a machine independent method. A solution to do this is to use toroidal harmonics functions as a basis for the decomposition of the poloidal flux in the vacuum region in complement to the contribution of the PF coils. The first version of the algorithm implementing this method for EQUINOX-ITM developed in 2012 has been updated and tested during 2013:

- WEST and JET: This method can provide an alternative tool, comparable to APOLO (for Tore Supra) and FELIX (for JET), to compute the plasma boundary in real time from the magnetic measurements. Some twin experiments for WEST have been successfully conducted. In a first step the equivalents of magnetic measurements were generated using the free boundary equilibrium code CEDRES++. In a second step these measurements were used by the toroidal harmonics algorithm to reconstruct the plasma boundary. Additional calculations aiming at validating the design of the WEST magnetic diagnostics have been performed. They consisted in checking the equilibrium reconstruction accuracy with respect to: (i) a reduced number of magnetic sensors; (ii) noise on magnetic sensor and/or current measurements. Then, experiments on the possibility to reconstruct not only the plasma boundary but also the current density have been conducted. A paper on this subject is accepted for publication [13]. The same algorithm has been tested using real JET measurements in order to provide an equilibrium reconstruction code that directly uses the magnetic measurements instead of using FELIX as an intermediate preprocessing of the measurements.
- EFDA-ITM (Task WP13-ITM-IMP12-ACT3): EQUINOX-ITM has been upgraded and tested on the new gateway machine of the ITM. The Kepler actor was tested and used on 3 different tokamaks (JET, Tore Supra and WEST) (with F. Imbeaux, T. Aniel, P. Moreau, E. Nardon (CEA)). A benchmark work is on going between the codes Equal, Efit and Equinox on JET shot 74221 (with Dimitriy Yadykin and Wolfgang Zwingmann).

6.3. Evolutive equilibrium and transport coupling and optimization of scenarii

Participants: Jacques Blum, Cédric Boulbe, Blaise Faugeras.

6.3.1. New developments in the direct evolutive version of CEDRES++

6.3.1.1. External circuits and saddle currents in the blankets

In the previous version of the free boundary equilibrium code CEDRES++, each PF-coil (Poloidal Field Coil) was considered separately. In the evolutive version, a voltage was applied to each coil. In the machine, PF-coils can be connected in series or in parallel with one or several power supplies. In order to consider more realistic configurations of the PF system, the model used in CEDRES++ has been generalized to circuits involving several coils and supplies connected in series or in parallel. This model has been implemented in CEDRES++ and has been tested on simple configurations with circuits composed of one supply and several coils connected in series. More complicated configurations like circuits with several supplies and coils in series and in parallel in the same circuit have to be tested. This will be done on a WEST test case. A model for saddle currents in the blankets has also been implemented. This model is actually under validation on DEMO geometry.

6.3.1.2. Coupling CEDRES++ with a feedback controller (task ITM-WP13-ITM-IMP12-ACT1-T3)

Cedres++ has been successfully coupled with a controller generated from the true TCV hybrid Simulink controller in an ITM (Integrated Tokamak Modeling) workflow. The "yoyo" discharge (shot 40475) has been reproduced. In Figure 1, the vertical position of the magnetic axis simulated matches the experimental one.

`../../../../projets/castor/IMG/controlcedres.png`

Figure 1. Simulation of Yoyo discharge on TCV - Comparison between Z axis simulated and Z axis obtained from experiments

6.3.1.3. Cedres++ - transport coupling

Last year, different algorithms coupling free boundary equilibrium solvers and transport solvers (CEDRES++-diff, CEDRES++-CRONOS, CEDRES++-ETS, FREEBIE-CRONOS) have been developed. The ETS-C coupling between CEDRES++ and the transport solver under the ITM environment has been finalized this year. A simulation of a VDE test case has been performed (task WP13-ITM-IMP12-ACT1-T2). A benchmark between the different strategies has been performed in order to validate the numerical methods required to ensure stability of the coupling system and to compare the physical simulations of each model. A benchmark between CEDRES++-diff solving the resistive diffusion equation coupled to CEDRES and the CEDRES-CRONOS coupling has been performed on an ITER test case. Some divergences between the two codes appear and are not fully understood despite long investigations. This difficulty has led us to delay the introduction of non inductive terms in the resistive diffusion equation implemented in CEDRES++-diff. These developments will be realized when the different coupling strategies will be fully validated.

6.3.2. Reasearch of optimal trajectories for the preparation of Tokamak discharges

A new approach has been developed for the optimization of dynamic plasma scenarios in tokamaks. This task has been formulated as an optimal control problem, using numerical solution methods for optimization problems with PDE constraints. Due to free boundary setting, a new linearization of the non-linear equations has been introduced, which is consistent with the numerical discretization. It is this consistency that guarantees that the method converges to the optimum.

6.4. Parallel Kelvin-Helhmohltz-like instability in edge plasma

Participants: Hervé Guillard, Boniface Nkonga, Marco Bilanceri, Giorgio Giorgiani.

A large part of this year activities have been devoted to the investigation of the Kelvin-Helmholtz-like instabilities that can be triggered at the core/SOL transition, the sheath acceleration at the limiter or divertor plates leading to a radial shear for the velocity. Linear stability analysis developed in Schwander et al [2] actually reveals that unstable modes at the edge-core transition can develop in the presence of core rotation. This study was also an opportunity for a benchmarking comparison with the TOKAM3X code developed in IRFM and at Marseille. Both codes have confirmed the linear stability analysis and have shown that large fluctuations grow in the shear layer downstream of the limiter on the LFS (Low Field Side), the growth of these fluctuations being accompanied by a radial drift away from the limiter. Fig. 2 displays a 3D representation of these unstable

---

[2] Schwander et al., J. Nucl. Materials, doi:10.1016/j.jnucmat.2010.10.073 2011

density fluctuations. Apart from the physical results, this study has also shown the large sensitivity of the solutions to the discretization and to the implementation of the Bohm's boundary conditions. It has also shown that the study of these parallel KH like instabilities is a very demanding benchmark : these simulations require large meshes and since the growth rate of these instabilities is very weak, this results in an extremely long simulation time involving an extremely large number of time steps. In the future, it is planned to investigate the saturation of the instability as well as its possible presence in diverted plasmas.

../../../../projets/castor/IMG/parallelKHinstabilities.png

Figure 2. 3D representation of unstable density fluctuations in edge plasma tokamak with limiter. The core plasma rotates in the anti-clockwise direction M//central=0.75, safety factor q=6

6.5. Development of a two temperature model

Participants: Hervé Guillard, Afeintou Sangam, Elise Estibals.

A two temperature (ions - electrons) version of the code is in development. At present an approximate Riemann solver using the total energy equation and the electron entropy as main variables has been designed. This Riemann solver has been validated against standard shock tube problems and incorporated in the PlaTo platform. The implementation of this solver in toroidal geometry is in progress.

6.6. Implementation of a Taylor-Galerkin stabilized Finite Element

Participants: José Costa, Marie Martin, Boniface Nkonga.

The theoretical part of Taylor-Galerkin (TG) stabilized strategy applied to MHD and reduced MHD modeling has been realized. The final method amounts to add in the formulation, a self-adjoint operator associated to the most critical hyperbolic component of the system to be solved. The design of the critical contours and the identification of associated waves to be stabilized is problem dependent and related to the Jacobian matrix. A simplified version has been developed for reduced MHD and takes into account the high anisotropy of strongly magnetized plasma under consideration here. This first implementation of the TG stabilization in Jorek, has made possible efficient and robust simulations of Edge Localized Modes (Elms) and their mitigation by Resonant Magnetic Perturbations (RMP) and pellets injections. Work under progress will use more elaborated TG formulations that will be applied to reduced and full MHD models.

6.7. Development of a full MHD Modeling

Participants: José Costa, Jeaniffer Vides, Boniface Nkonga.

The single fluid full MHD numerical model has been developed. The divergence free constraint on the magnetic field is achieved by introduction of a potential vector. The use of the potential vector has the additional advantage that the toroidal component is the magnetic flux of the Grad-Shafranov equilibrium. Therefore, using the same finite element for the computation of initial equilibrium and the evolution of perturbed system, the numerical scheme is well balanced when the projection of the momentum equation use a component parallel to the magnetic field. Indeed, at the discrete level the projection is exactly orthogonal to equilibrium sub-space. Using the potential vector as variable introduces third order derivatives in the system and classical C0 finite elements cannot be directly applied. This is why our finite element strategy uses shape/test functions whose derivatives have global continuity in space. Finite element method is designed for poloidal plane discretization using quadrangles or triangles. Validations have been performed for internal kink and tearing modes instabilities in tokamak with a circular plasma. For this configuration, all magnetic surfaces are closed and simple boundary conditions are used. Future work will address X-point configurations with Bohm boundary conditions.

6.8. Environmental and Astrophysical flows

Participants: Hervé Guillard, Boniface Nkonga, Marco Bilanceri, Maria-Vittoria Salvetti [University of Pisa, Italy], Karim Elhakim [University Ains Shams, Egypt].

The numerical approximation of a model coupling the shallow-water equations with a sediment transport equation for the morphodynamics has been studied. In shallow-water problems, time advancing can be carried out by explicit schemes. However, if the interaction with the mobile bed is weak, the characteristic time scales of the flow and of the sediment transport can be very different introducing time stiffness in the global problem. For this case, it is of great interest to use implicit schemes. The time integration stategy that we have devised is based on a defect-correction approach and on a time linearization, in which the flux Jacobians are computed through automatic differentiation. This work has been published in this reference [3]. The aim of the present work is to investigate the behavior of this time scheme for different Riemann solvers, sediment transport models and situations related to environmental flows [12]. This activity takes place in the framework of the Euromediterranée 3+3 MedLagoon program and a PHC Imhotep program. A preliminary work has begun to apply this strategy to the study of the Burulus lake in Egypt.

6.9. Ionospheric plasma

Participants: Didier Auroux, Sebastian Minjeaud.

In order to guarantee the integrity of the european positionning system Galileo, it is fundamental to identify all the potential sources of system unavailability. One of the main sources that has been identified is the phenomenon of ionospheric scintillations which causes radio frequency signal amplitude and phase variations when satellite signals pass through the ionosphere. Scintillations appear as the turbulent aspect of a larger disturbance of the ionospheric plasma density, which has the shape of a plasma bubble. In this context and in the framework of the ANR IODISSEE, a model hierarchy aimed at representing the evolution of the ionospheric plasma was proposed (Besse and al., 2004). It is based on an asymptotic analysis of the Euler-Maxwell system thanks to typical scales of the physical parameters involved in this framework. Among these models the simplest, referred to as the Striation model, describes the evolution of the quasineutral plasma in a plane perpendicular to the earth magnetic field. The magnetic field is assumed constant, and both electron and ions inertia are neglected. In this model the mobility of charge particles is assumed infinite along the magnetic field lines so that they constitute equipotential for the electric field. This property allows the computation of the electric field by means of a two dimensional elliptic equation with coefficients integrated along the magnetic field lines. This equation is coupled to either a two or a three dimensional transport equation for the evolution of the plasma density.

---

[3] Linearized implicit time advancing and defect correction applied to sediment transport simulations Marco Bilanceri; François Beux; Imad Elmahi; Hervé Guillard; Maria Vittoria Salvetti Computers and Fluids, elsevier, 2012, 63, pp. 82-104

6.9.1. Data assimilation

We worked on data-models coupling method to identify the parameters of the Striation model (especially, the initial data for the electronic density and the ion/neutral collision frequency). Some measurements acquired during the mission of satellite DEMETER will constitute the set of observed data. We consider this problem from an optimal control point of view. We define a cost function measuring the misfit between the observed data and the corresponding model states. This function can be seen as a function of only the input model parameters. The previous inverse problem is then equivalent to minimizing this cost function. Of course, this problem can be ill posed (over or under-determined, non convex...) and we need to add some regularization terms (mainly Tikhonov terms), using some a priori estimation of the model parameters. The effective minimization of the cost function (in order to estimate the best possible set of model parameters) relies on the computation of the adjoint state. Thanks to the help of L. Hascoet, the adjoint of the Striation code was derived and validated with the automatic differentiation tool TAPENADE. We obtained the first results of coefficients identification in very simple situations.

6.9.2. Wave Propagation

The electronic density fluctuations of the ionospheric plasma have been identified as the main causes of the scintillation phenomena since they induce some variations of the amplitudes and phases of the signals passing through the ionosphere. These fluctuations have indeed a direct influence on the refractive index of the medium. A code was developed to simulate the propagation of wave signals in the perturbed ionosphere (whose representation is obtained thanks to the Striation code). We chose to use the method of the Rytov approximations which allow to obtain sufficiently accurate results (since, e.g., it takes into account the diffraction due to the small structures) whithin a reasonable computational time (contrary to the resolution of the whole Maxwell system). This work was carried out during the Master 2 internship of Gonzalo José Carracedo Carballal (advised with P. Lafitte, Ecole Centrale Paris).

6.10. Mesh adaptative MG Methods

Participants: Gautier Brethes [Projet Ecuador], Alain Dervieux, Olivier Allain [Lemma].

Anisotropic tetrahedrization, Continuous metric

This activity concerns the use of mesh adaptation and multigrid for simplified plasma models in the context of ANEMOS ANR project.

6.11. Turbulence models

Participants: Alain Dervieux, Bruno Koobus [University of Montpellier 2], Carine Moussaed [University of Montpellier 2], Maria-Vittoria Salvetti [University of Pisa], Stephen Wornom [Lemma], Marianna Braza [IMF-Toulouse].

Large Eddy Simulation, Variational Multi-scale, hybrid models, unstructured meshes, vortex shedding

The purpose of our works in hybrid RANS/LES is to develop new approaches for industrial applications of LES-based analyses. This year, a lot of experiments have validated the dynamic version of our VMS-LES. The quality of simulations is either comparable to non-dynamic, or better. In the foreseen applications (aeronautics, hydraulics), the Reynolds number can be as high as several tenth millions, a far too large number for pure LES models. However, certain regions in the flow can be much better predicted with LES than with usual statistical RANS (Reynolds averaged Navier-Stokes) models. These are mainly vortical separated regions as assumed in one of the most popular hybrid model, the hybrid Detached Eddy Simulation model. Here, “hybrid” means that a blending is applied between LES and RANS. The french-italian team is working on a novel type of hybrid model between the VMS-LES model and a $k-\varepsilon$ one. The team has this year concentrated on the shift between the RANS boundary region and the VLES boundary one. This is also the problematic on the IDDES studies. We are working on propositions relying not only on the value of the RANS viscosity but also on its gradient. A paper is in preparation on this subject.

CLIME Project-Team

6. New Results

6.1. New methods for data assimilation

One major objective of Clime is the conception of new techniques for data assimilation in geophysical sciences. Clime is active on several of the most challenging theoretical aspects of data assimilation: data assimilation methods based on non-Gaussian assumptions, methods for estimating errors, ensemble filtering techniques, 4D variational assimilation approaches, ensemble-variational methods, etc.

This year, focus was on ensemble-variational methods. We introduced a new method known as the iterative ensemble Kalman smoother. It is an ensemble method with an underlying cost function; it does not require the use of the adjoint; and it is flow-dependent. Because of these propreties, the IEnKS outperforms other data assimilation methods when tested with perfect meteorological toy-models. Its potential for parameter estimation has also been demonstrated.

6.1.1. An iterative ensemble Kalman smoother

Participants: Marc Bocquet, Pavel Sakov [BOM, Australia].

The iterative ensemble Kalman filter (IEnKF) was recently proposed to improve the performance of ensemble Kalman filtering with strongly nonlinear geophysical models. IEnKF can be used as a lag-one smoother and extended to a fixed-lag smoother: the iterative ensemble Kalman smoother (IEnKS [12]). IEnKS is an ensemble variational method. It does not require the use of the tangent linear of the evolution and observation models, nor the adjoint of these models: the required sensitivities (gradient and Hessian) are obtained from the ensemble. Looking for the optimal performance, we consider a quasi-static algorithm, out of the many possible extensions. IEnKS is explored on the Lorenz'95 model and on a 2D turbulence model. As a logical extension of IEnKF, IEnKS significantly outperforms standard Kalman filters and smoothers in strongly nonlinear regimes. In mildly nonlinear regimes (typically synoptic scale meteorology), its filtering performance is marginally but clearly better than the standard ensemble Kalman filter, and it keeps improving as the length of the temporal data assimilation window is increased. For long windows, its smoothing performance very significantly outranks the standard smoothers, which is believed to stem from the variational but flow-dependent nature of the algorithm. For very long windows, the use of a multiple data assimilation variant of the scheme, where observations are assimilated several times, is advocated. This paves the way for finer re-analysis freed from the static prior assumption of 4D-Var, but also partially freed from the Gaussian assumptions that usually impede standard ensemble Kalman filtering and smoothing.

6.1.2. Joint state and parameter estimation with an iterative ensemble Kalman smoother

Participants: Marc Bocquet, Pavel Sakov [BOM, Australia].

Both ensemble filtering and variational data assimilation methods have proven being useful in the joint estimation of state variables and parameters of geophysical models. Yet, their respective benefits and drawbacks in this task are distinct. An ensemble variational method, known as the iterative ensemble Kalman smoother (IEnKS), has recently been introduced. It is based on an adjoint-free variational but flow-dependent scheme. As such, IEnKS is a candidate tool for joint state and parameter estimation that may inherit the benefits from both the ensemble filtering and variational approaches.

In this study [13], an augmented state IEnKS is tested on the estimation of the forcing parameter of the Lorenz'95 model. Since joint state and parameter estimation is especially useful in applications where the forcings are uncertain but nevertheless determining, typically in atmospheric chemistry, the augmented state IEnKS is tested on a new low-order model that combines the Lorenz'95 model, representing its meteorological part, and the advection diffusion of a tracer for its chemical part. In these experiments, IEnKS is compared to the ensemble Kalman filter, the ensemble Kalman smoother and a 4D-Var method, that are considered choices to solve these joint estimation problems. In this low-order model context, IEnKS is shown to significantly outperform those methods, for any length of the data assimilation window, and for present time analysis as well as retrospective analysis. Besides, the performance of IEnKS is even more striking on parameter estimation, whereas getting close to the same performance with 4D-Var is likely to require both a long data assimilation window and a complex modeling of the background statistics.

6.1.3. Data assimilation applied to air quality at urban scale

Participants: Vivien Mallet, Raphaël Périllat, Anne Tilloy, Fabien Brocheton [Numtech], David Poulet [Numtech], Frédéric Mahé [Airparif], Pierre Pernot [Airparif], Fabrice Joly [Airparif].

Based on Verdandi [14], Polyphemus and the "Urban Air Quality Analysis" software, data assimilation was further developed at urban scale. The Best Linear Unbiased Estimator (BLUE) is computed to merge the outputs of the ADMS Urban model and the observations of a sparse monitoring network [19]. We improved the modeling of the covariance of the model state error. The assimilation was applied for part of Paris (see Fig. 2) and for Paris region, in the context of the PRIMEQUAL project PREQUALIF ("Multidisciplinary Program on Air Quality Research in Île-de-France").

../../../../projets/clime/IMG/adms_2012/09/01_18.png

../../../../projets/clime/IMG/analyse_201

Figure 2. Left: Map of [NO_2] ($\mu g\ m^{-3}$), before assimilation, at a given date in September 2012. Right: Map of [NO_2] ($\mu g\ m^{-3}$), after assimilation of the observations (disks).

It was applied to nitrogen dioxide, particulate matter and black carbon. Specific investigations were carried out to estimate the variance of the a posteriori error and to determine the impact of each monitoring station on the final results.

6.2. Inverse modeling

We continued research on inverse modelling techniques, with a focus on hyperparameter estimation when the statistics are non-Gaussian. We applied these methods to the estimation of the caesium-137 Fukushima source term using heterogenous datasets. We applied similar methods to the estimation of Volatile Organic Compounds (VOC) at the European scale by assimilation of the EMEP VOC observations over one year. We also studied the estimation of several hyperparameters in the context of CO_2 flux inversions.

6.2.1. Estimation of the caesium-137 source term from the Fukushima Daiichi nuclear power plant using a consistent joint assimilation of air concentration and deposition observations

Participants: Victor Winiarek, Marc Bocquet, Nora Duhanyan [CEREA], Yelva Roustan [CEREA], Olivier Saunier [IRSN], Anne Mathieu [IRSN].

To estimate the amount of radionuclides and the temporal profile of the source term released in the atmosphere during the accident of the Fukushima Daiichi nuclear power plant in March 2011, inverse modeling techniques have been used and have proven their ability in this context. In a previous study, the lower bounds of the caesium-137 and iodine-131 source terms were estimated with such techniques, using activity concentration observations. The importance of an objective assessment of prior errors (the observation errors and the background errors) was emphasised for a reliable inversion. In such critical context where the meteorological conditions can make the source term partly unobservable and where only a few observations are available, such prior estimation techniques are mandatory, the retrieved source term being very sensitive to this estimation.

We propose to extend the use of these techniques to the estimation of prior errors when assimilating observations from several data sets [21]. The aim is to compute an estimate of the caesium-137 source term jointly using all available data about this radionuclide, such as activity concentrations in the air, but also daily fallout measurements and total cumulated fallout measurements. It is crucial to properly and simultaneously estimate the background errors and the prior errors relative to each data set. A proper estimation of prior errors is also a necessary condition to reliably estimate the a posteriori uncertainty of the estimated source term. Using such techniques, we retrieve a total released quantity of caesium-137 in the interval $11.6 - 19.3$ PBq with an estimated standard deviation range of $15 - 20\%$ depending on the method and the data sets. The “blind” time intervals of the source term have also been strongly mitigated compared to the first estimations with only activity concentration data.

6.2.2. An inverse modeling method to assess the source term of the Fukushima Nuclear Power Plant accident using gamma dose rate observations

Participants: Olivier Saunier [IRSN], Anne Mathieu [IRSN], Damien Didier [IRSN], Maryline Tombette [IRSN], Denis Quélo [IRSN], Victor Winiarek, Marc Bocquet.

The Chernobyl nuclear accident, and more recently the Fukushima accident, highlighted that the largest source of error on consequences assessment is the source term, including the time evolution of the release rate and its distribution between radioisotopes. Inverse modeling methods, which combine environmental measurements and atmospheric dispersion models, have proven being efficient in assessing source term due to an accidental situation. Most existing approaches are designed to use air sampling measurements and some of them also use deposition measurements [21]. Some studies have been conceived to use dose rate measurements, but none of the developed methods were carried out to assess the complex source term of a real accident situation like the Fukushima accident. However, dose rate measurements are generated by the most widespread measurement system and, in the event of a nuclear accident, these data constitute the main source of measurements of the plume and radioactive fallout during releases. This study [18], [23] proposes a method to use dose rate measurements as part of an inverse modeling approach to assess source terms. The method is proven efficient and reliable when applied to the accident at the Fukushima Daiichi Nuclear Power Plant (FD-NPP). The emissions for the eight main isotopes have been assessed. Accordingly, 105.9 PBq of ^{131}I, 35.8 PBq of ^{132}I, 15.5 PBq of ^{137}Cs and 12,134 PBq of noble gases were released. The events at FD-NPP (such as venting, explosions, etc.) known to have caused atmospheric releases are well identified in the retrieved source term. The estimated source term is validated by comparing simulations of atmospheric dispersion and deposition with environmental observations. In total, it was found that for 80 % of the measurements, simulated and observed dose rates agreed within a factor of 2. Changes in dose rates over time have been overall properly reconstructed, especially in the most contaminated areas to the northwest and south of the FD-NPP. A comparison with observed atmospheric activity concentration and surface deposition shows that the emissions of caesiums and ^{131}I are realistic but that ^{132}I and ^{132}Te are probably underestimated and noble gases are likely overestimated. Finally, an important outcome of this study is that the method proved to be perfectly suited to emergency management and could contribute to improve emergency response in the event of a nuclear accident.

6.2.3. Estimation of volatile organic compound emissions for Europe using data assimilation

Participants: Mohammad Reza Koohkan, Marc Bocquet, Yelva Roustan [CEREA], Yougseob Kim [CEREA], Christian Seigneur [CEREA].

The emissions of non-methane volatile organic compounds (VOCs) over western Europe for the year 2005 are estimated via inverse modeling by assimilation of in situ observations of concentration and they are subsequently compared to a standard emission inventory. The study [16] focuses on fifteen VOC species: five aromatics, six alkanes, two alkenes, one alkyne and one biogenic diene. The inversion relies on a validated fast adjoint of the chemical transport model used to simulate the fate and transport of these VOCs. The assimilated ground-based measurements over Europe are provided by the European Monitoring and Evaluation Programme (EMEP) network. The background emissions errors and the prior observational errors are estimated by maximum likelihood approaches. The positivity assumption on the VOC emission fluxes is pivotal for a successful inversion and this maximum likelihood approach consistently accounts for the positivity of the fluxes. For most species, the retrieved emissions lead to a significant reduction of the bias, which underlines the misfit between the standard inventories and the observed concentrations. The results are validated through a forecast test and a cross-validation test. An estimation of the posterior uncertainty is also provided. It is shown that the statistically consistent non-Gaussian approach, based on a reliable estimation of the errors, offers the best performance. The efficiency in correcting the inventory depends on the lifetime of the VOCs and the accuracy of the boundary conditions. In particular, it is shown that the use of in situ observations using a sparse monitoring network to estimate emissions of isoprene is inadequate because its short chemical lifetime significantly limits the spatial radius of influence of the monitoring data. For species with longer lifetime (a few days), successful, albeit partial, emission corrections can reach regions hundreds of kilometres away from the stations. Domainwide corrections of the emissions inventories of some VOCs are significant, with underestimations on the order of a factor of two for propane, ethane, ethylene and acetylene.

6.2.4. Hyperparameter estimation for uncertainty quantification in mesoscale carbon dioxide inversions

Participants: Lin Wu [LSCE, France], Marc Bocquet, Frédéric Chevallier [LSCE, France], Thomas Lauvaux [Department of Meteorology, Pennsylvania State University, USA], Kenneth Davies [Department of Meteorology, Pennsylvania State University, USA].

Uncertainty quantification is critical in the inversion of CO_2 surface fluxes from atmospheric concentration measurements. We estimate the main hyperparameters of the error covariance matrices for a priori fluxes and CO_2 concentrations, that is, the variances and the correlation lengths, using real, continuous hourly CO_2 concentration data in the context of the Ring 2 experiment of the North American Carbon Program Mid Continent Intensive. Several criteria, namely maximum likelihood (ML), general cross-validation (GCV) and χ^2 test are compared for the first time under a realistic setting in a mesoscale CO_2 inversion. It is shown [22] that the optimal hyperparameters under the ML criterion assure perfect χ^2 consistency of the inverted fluxes. Inversions using the ML error variances estimates rather than the prescribed default values are less weighted by the observations, because the default values underestimate the model-data mismatch error, which is assumed to be dominated by the atmospheric transport error. As for the spatial correlation length in prior flux errors, the Ring 2 network is sparse for GCV and this method fails to reach an optimum. In contrast, the ML estimate (e.g. an optimum of 20 km for the first week of June 2007) does not support long spatial correlations that are usually assumed in the default values.

6.3. Monitoring network design

In this section, we report studies that are related to the evaluation of monitoring networks and to new monitoring strategies. This year, we studied the impact of using lidar observation for particulate matter forecasting.

6.3.1. Assimilation of ground versus lidar observations for PM_{10} forecasting

Participants: Yiguo Wang [CEREA], Karine Sartelet [CEREA], Marc Bocquet, Patrick Chazette [LSCE, France].

This study [20] investigates the potential impact of future ground-based lidar networks on analysis and short-term forecasts of PM_{10}. To do so, an Observing System Simulation Experiment (OSSE) is built for PM_{10} data assimilation using optimal interpolation over Europe for one month in 2001. First, we estimate the efficiency of the assimilation of lidar network measurements in improving PM_{10} concentration analysis and forecast. It is compared to the efficiency of assimilating concentration measurements from the AirBase ground network, which includes about 500 stations in western Europe. It is found that the assimilation of lidar observations is more efficient at improving PM_{10} concentrations in terms of root mean square error and correlation after 12 hours of assimilation than the assimilation of AirBase measurements. Moreover, the spatial and temporal influence of the assimilation of lidar observations is larger and longer. In our experiments, the assimilation of lidar products improves PM_{10} forecast for 108 hours against 60 hours for AirBase assimilation. The results show a potentially powerful impact of the future lidar networks. Secondly, since a lidar is a very costly instrument, a sensitivity study on the number of required lidars is performed to help defining an optimal lidar network for PM_{10} forecast. The results suggest 12 lidar stations over western Europe, because a network with 26 lidar stations is more expensive and offers a limited improvement (less than $1\ \mu g\ m^{-3}$ of root mean square error on average) over the lidar network. A comparison of two networks with 12 lidar stations at different locations does not lead to substantial differences.

6.4. Reduction and emulation

The use of environmental models raise a number of problems due to:

- the dimension of the inputs, which can easily be $10^5 - -10^8$ at every time step;
- the dimension of the state vector, which is usually $10^5 - -10^7$;
- the high computational cost.

In particular, the application of data assimilation methods and uncertainty quantification techniques may require dimension reduction and cost reduction. The dimension reduction consists in projecting the inputs and the state vector to low-dimensional subspaces. The cost reduction can be carried out by emulation, i.e., the replacement of costly components with fast surrogates.

6.4.1. Reduction and emulation of a chemistry-transport model

Participants: Vivien Mallet, Serge Guillas [University College London].

Both reduction and emulation were applied to the dynamic air quality model Polair3D from Polyphemus. The reduction relied on proper orthogonal decomposition (POD) on the input data and on the state vector. The dimension of the reduced subspace for the input data is about 80, while the dimension of the reduced state vector is less than 10. The projection of the state vector on its reduced subspace can be carried out before every integration time step, so that one can reproduce a full state trajectory (in time) using the reduced model.

Significant advances were made to emulate the reduced model, which requires about 90 inputs (reduced input data and reduced state vector) and computes about 10 outputs (reduced state vector). 90 inputs is however a large number to build an emulator using a classical approaches. Promising results were however obtained with radial basis functions and an adapted kriging-based method.

6.4.2. Reduction and emulation of a static air quality model

Participants: Vivien Mallet, Anne Tilloy, Fabien Brocheton [Numtech], David Poulet [Numtech].

The dimension reduction was applied to the outputs of the urban air quality model ADMS Urban, which is a static model with low-dimensional inputs and high-dimensional outputs. A proper orthogonal decomposition (POD) on the outputs allowed us to drastically reduce their dimension, from 10^4 to just a few scalars. The emulation of the reduced model itself was successfully carried out with radial basis functions or an adapted kriging-based method. The resulting reduced/emulated model exhibited meaningful response to all variables. Its performance compared to observations was the same as the original model. The computational cost of the full model is about 8 minutes on 16 cores (for a single time step), while the reduced/emulated model requires only 50 ms on one core [29].

6.4.3. Motion estimation from images with a waveforms reduced model

Participants: Etienne Huot, Isabelle Herlin, Giuseppe Papari [Lithicon, Norway], Karim Drifi.

Dimension reduction is applied to an image model, composed of Lagrangian constancy of velocity and transport of image brightness. Waveforms basis are obtained on the image domain for subspaces of images, motion fields and divergence-free motion fields, as eigenvectors of quadratic functions. Image assimilation with th reduced model allows to estimate velocity fields satisfying space-time properties defined by user and traduced as a quadratic function. This approach also solves the issue of complex geographical domains and the difficulty of applying boundary conditions on these domains. Results are obtained with a reduced dimension of motion to a few scalars, to be compared with the original problem that has the size of image domain [31], [26], [25].

6.5. Ensemble forecasting with sequential aggregation

The aggregation of an ensemble of forecasts is an approach where the members of an ensemble are given a weight before every forecast time, and where the corresponding weighted linear combination of the forecasts provides an improved forecast. A robust aggregation can be carried out so as to guarantee that the aggregated forecast performs better, in the long run, than any linear combination of the ensemble members with time-independent weights. The approaches are then based on machine learning. The aggregation algorithms can be applied to forecast analyses (generated from a data assimilation system), so that the aggregated forecasts are naturally multivariate fields.

6.5.1. Application of sequential aggregation to meteorology

Participants: Paul Baudin, Vivien Mallet, Gilles Stoltz [CNRS], Laurent Descamps [Météo France].

Nowadays, it is standard procedure to generate an ensemble of simulations for a meteorological forecast. Usually, meteorological centers produce a single forecast, out of the ensemble forecasts, computing the ensemble mean (where every model receives an equal weight). It is however possible to apply aggregation methods. When new observations are available, the meteorological centers also compute analyses. Therefore, we can apply the ensemble forecast of analyses. Ensembles of forecasts for wind velocity and mean sea level pressure, from Météo France, were aggregated. Preliminary results show significant improvements for mean sea level pressure.

6.5.2. Sequential aggregation with uncertainty estimation

Participants: Vivien Mallet, Sergiy Zhuk [IBM research, Ireland], Paul Baudin, Gilles Stoltz [CNRS].

An important issue is the estimation of the uncertainties associated with the aggregated forecasts. One investigated direction relies on the framework of machine learning, with the aggregation of an ensemble of probability density functions instead of the point forecasts of the ensemble.

Another direction is to reformulate the aggregation problem in a filtering problem for the weights. The weights are supposed to satisfy some dynamics with unknown model error, which defines the state equation of a filter. An observation equation compares the aggregated forecast with the observations (or analyses) with known observational error variance. The filter finally computes estimates for the weights and quantifies their uncertainties. We applied a Kalman filter and a minimax filter for air quality forecasting. We also introduced a criterion that the filter results should satisfy if they are representative of the uncertainties [17].

6.6. Uncertainty quantification

Many uncertainties limit the forecast skills of geophysical simulations: limited understanding of physical phenomena, simplified representation of a system state and of the physical processes, inaccurate data and approximate numerical solutions. In many applications, a deterministic forecast or analysis is not enough a result since its uncertainties may be very large. It is of high interest to evaluate the quality of a forecast, before observations are available, and the quality of an analysis at any location, observed or not. An even more desirable result is the full probability density of system state, which can only be derived from a fully stochastic approach.

6.6.1. Sensitivity analysis in the dispersion of radionuclides

Participants: Sylvain Girard [IRSN], Vivien Mallet, Irène Korsakissok [IRSN].

We carried out a sensitivity analysis of the dispersion of radionuclides during Fukushima disaster. We considered the dispersion at regional scale, with the Eulerian transport model Polair3D from Polyphemus. The sensitivities to most input parameters were computed using the Morris method (with 8 levels and 100 trajectories). The influences of 19 scalar parameters were quantified. The scalar parameters were additive terms or multiplicative factors applied to 1D, 2D or 3D fields such as emission rates, precipitations, cloud height, wind velocity. It was shown that, depending on the output quantities of interest (various aggregated atmospheric and ground dose rates), the sensitivity to the inputs may greatly vary. Very few parameters show low sensitivity in any case. The vertical diffusion coefficient, the scavenging factors, the winds and precipitation intensity were found to be the most influential inputs. Most input variables related to the source term (emission rates, emission dates) also had a strong influence.

6.7. Image assimilation

Sequences of images, such as satellite acquisitions, display structures evolving in time. This information is recognized of major interest by forecasters (meteorologists, oceanographers, etc.) in order to improve the information provided by numerical models. However, these satellite images are mostly assimilated in geophysical models on a point-wise basis, discarding the space-time coherence visualized by the evolution of structures such as clouds. Assimilating in an optimal way image data is of major interest and this issue should be considered in two ways:

- from the model's viewpoint, the location of structures on the observations is used to control the state vector.
- from the image's viewpoint, a model of the dynamics and structures is built from the observations.

6.7.1. Divergence-free motion estimation

Participants: Dominique Béréziat [UPMC], Isabelle Herlin, Sergiy Zhuk [IBM Research, Ireland].

This research addresses the issue of divergence-free motion estimation on an image sequence, acquired over a given temporal window. Unlike most state-of-the-art technics, which constrain the divergence to be small thanks to Tikhonov regularization terms, a method that imposes a null value of divergence of the estimated motion is defined.

Motion is either characterized by its vorticity value or by its coefficients on a divergence-free basis and assumed to satisfy the Lagragian constancy hypothesis. An image assimilation method, based on the 4D-Var technic, is defined that estimates motion as a compromise between the evolution equations of vorticity or projection coefficients and the observed sequence of images.

The method is applied on Sea Surface Temperature (SST) images acquired over Black Sea by NOAA-AVHRR sensors. The divergence-free assumption is roughly valid on these acquisitions, due to the small values of vertical velocity at the surface.

6.7.2. Model error and motion estimation

Participants: Dominique Béréziat [UPMC], Isabelle Herlin.

Data assimilation technics are used to retrieve motion from image sequences. These methods require a model of the underlying dynamics, displayed by the evolution of image data. In order to quantify the approximation linked to the chosen dynamic model, an error term is included in the evolution equation of motion and a weak formulation of 4D-Var data assimilation is designed. The cost function to be minimized simultaneously depends on the initial motion field, at the begining of the studied temporal window, and on the error value at each time step. The result allows to assess the model error and analyze its impact on motion estimation.

The approach has been used to estimate the impact of geophysical forces (gravity, Coriolis, diffusion) and better assess the surface dynamics [24].

6.7.3. *Tracking of structures from an image sequence*

Participants: Yann Lepoittevin, Isabelle Herlin, Dominique Béréziat [UPMC].

The research concerns an approach to estimate velocity on an image sequence and simultaneously segment and track a given structure. It relies on the underlying dynamics' equations of the studied physical system. A data assimilation method is designed to solve evolution equations of image brightness, those of motion's dynamics, and those of the distance map modeling the tracked structures. The method is applied on meteorological satellite data, in order to track tropical clouds on image sequences and estimate their motion, as seen on Fig. 3 .

../../../../projets/clime/IMG/tracking-3.png ../../../../projets/clime/IMG/tracking-9.png ../../../../projets/clime/IMG/

Figure 3. Tracking a tropical cloud. Frames 3, 9, 18 of the sequence.

Quantification is obtained on synthetic experiments by comparing trajectories of characteristic points. The respective position of these points on the last image of the sequence for different methods may be compared to that obtained with ground truth as seen on Fig. 4 .

../../../../projets/clime/IMG/ComPts-17-ell.png

Figure 4. Red point: ground truth. Blue point: our method. Green point: Sun's optical flow. Blue ellipse: our method is the best. Green ellipse: Sun's result is the best. Grey ellipse : results are equivalent.

Data assimilation is performed either with a 4D-Var variational approach [27], [30], [28] or with an ensemble approach. In the last case, computation of the ensemble from optical flow methods of the literature is currently studied.

6.8. Minimax filtering

In minimax filtering for state estimation, the initial state error, the model error and the observation errors are supposed to belong to one joint ellipsoid. It is only assumed that the errors, stochastic or deterministic, are bounded. During the assimilation process, the filter computes an ellipsoid where one will find at least all states compatible with observations and errors description. The state estimate is taken as the center of the ellipsoid. No assumption on the actual distribution of the errors in needed and the state estimate minimizes the worst-case error, which makes the filter robust.

6.8.1. Retrieval of a continuous image function and a posteriori minimax motion estimation

Participants: Sergiy Zhuk [IBM Research, Ireland], Isabelle Herlin, Olexander Nakonechnyi [Taras Shevchenko National University of Kyiv], Jason Frank [CWI, the Netherlands].

An iterative minimax method is developed for the problem of motion estimation from an image sequence. The main idea of the algorithm is to use the "bi-linear" structure of the Navier-Stokes equations and optical flow constraint in order to iteratively estimate the velocity. The algorithm consists of the following parts:

1) we construct a continuous image function $\widehat{I}$, solving the optical flow constraint, such that $\widehat{I}$ fits (in the sense of least-squares) the observed sequence of images. To do so, we set the velocity field in the optical flow constraint to be the current minimax estimate of the velocity field $\mathbf{w}$, obtained at the previous iteration of the algorithm, and construct the minimax estimate $\widehat{I}$ of the resulting linear advection equation using the observed image sequence as discrete measurements of the brightness function;

2) we plug the estimate of the image gradient, obtained out of pseudo-observations $\widehat{I}$ in 1), into the optical flow constraint and the current minimax estimate $\mathbf{w}$ of the velocity field into the non linear part of Navier-Stokes equations so that we end up with a system of linear PDEs, which represents an extended state equation: it contains a linear parabolic equation for the velocity field and linear advection equation for the image brightness function. We construct the minimax estimate of the velocity field from the extended state equation using again the observed image sequence as discrete measurements of the brightness function;

3) we use the minimax estimate of the velocity field obtained in 2) in order to start 1) again.

Alternatively, point 1) may be used to retrieve a continuous image function from sparse and noisy image snapshots, based on previous motion estimation with a 4D-Var technic as seen on Fig. 5 , that displays ground truth, noisy image observation, image estimation at the end of the studied intervall.

Figure 5. From left to right: Ground truth, image observation, result.

6.9. Fire application

6.9.1. Model evaluation for fire propagation

Participants: Vivien Mallet, Jean-Baptiste Fillipi [CNRS], Bahaa Nader [University of Corsica].

In the field of forest fires risk management, important challenges exist in terms of people and goods preservation. Answering to strong needs from different actors (firefighters, foresters), researchers focus their efforts to develop operational decision support system tools that may forecast wildfire behavior. This requires the evaluation of model performance, but currently, simulation errors are not sufficiently qualified and quantified.

We consider that the proper evaluation of a model requires to apply it to a large number of fires – instead of carrying out a fine tuning on just one fire. We implemented a software to simulate a large number of fires (from the Prométhée database, http://www.promethee.com/) with the simulation model ForeFire (CNRS/University of Corsica) and evaluate the results with error measures [15]. One simulation requires mainly the following data: the ignition point, the ground elevation, the vegetation cover and the wind field. See illustration in Fig. 6 . We simulated 80 fires with four physical models, which proved that the most advanced models performed better overall, even though the input data is often inaccurate. We also carried out Monte Carlo simulations to evaluate the impact of the uncertainty in input data. We showed that the Monte Carlo approach led to a reliable forecasting system, which suggests that the probability densities derived from the simulations (see Fig. 6) may be useful information for preventive actions in an operational context.

../../../../projets/clime/IMG/San-Giovanni-di-Moriani_04-08-2003.jpg

../../../../projets/clime/IMG/probability

Figure 6. Left: Fire simulation (using ForeFire) in red elevated contour, and observation (from Prométhée) of the burned area in filled red contour, for a 2003 fire near San-Giovanni-di-Moriani (Corsica). Right: Burn probability as computed by a Monte Carlo simulation for a wildfire that was observed (red contour) in Corsica in 2003.

COFFEE Project-Team (section vide)

FLUMINANCE Project-Team

6. New Results

6.1. Fluid motion estimation

6.1.1. Stochastic uncertainty models for motion estimation

Participants: Sébastien Beyou, Etienne Mémin, Emmanuel Saunier.

In this study we have proposed a stochastic formulation of the brightness consistency used principally in motion estimation problems. In this formalization the image luminance is modeled as a continuous function transported by a flow known only up to some uncertainties. Stochastic calculus then enables to built conservation principles which take into account the motion uncertainties. These uncertainties defined either from isotropic or anisotropic models can be estimated jointly to the motion estimates. Such a formulation besides providing estimates of the velocity field and of its associated uncertainties allows us to naturally define a linear multiresolution scale-space framework. The corresponding estimator, implemented within a local least squares approach, has shown to improve significantly the results of the corresponding deterministic estimator (Lucas and Kanade estimator). This fast local motion estimator provides results that are of the same order of accuracy than state-of-the-art dense fluid flow motion estimator for particle images. The uncertainties estimated supply a useful piece of information in the context of data assimilation. This ability has been exploited to define multiscale incremental data assimilation filtering schemes. This work has been recently published in Numerical Mathematics: Theory, Methods and Applications [14]. It is also described in Sébastien Beyou's PhD dissertation [11]. The development of an efficient GPU based version of this estimator recently started through the Inria ADT project FLUMILAB

6.1.2. 3D flows reconstruction from image data

Participants: Ioana Barbu, Kai Berger, Cédric Herzet, Etienne Mémin.

Our work focuses on the design of new tools for the estimation of 3D turbulent flow motion in the experimental setup of Tomo-PIV. This task includes both the study of physically-sound models on the observations and the fluid motion, and the design of low-complexity and accurate estimation algorithms. On the one hand, we investigate state-of-the-art methodologies such as "sparse representations" for the characterization of the observation and fluid motion models. On the other hand, we place the estimation problem into a probabilistic Bayesian framework and use state-of- the-art inference tools to effectively exploit the strong time-dependence on the fluid motion. In our previous work, we have focussed on the problem of reconstructing the particle positions from several two-dimensional images. Our approach was based on the exploitation of a particular family of sparse representation algorithms, leading to a good trade-off between performance and complexity. Moreover, we also tackled the problem of estimating the 3D velocity field of the fluid flow from two instances of reconstructed volumes of particles. Our approach was based on a generalization of the well-known Lucas-Kanade's motion estimator to 3D problems. A potential strength of the proposed approach is the possibility to consider a fully parallelized (and therefore very fast) hardware implementation. This year, we have focused on the design of new methodologies to jointly estimate the volume of particles and the velocity field from the received image data. Our approach is based on the minimization (with respect to both the position of the particles and the velocity field) of a cost function penalizing both the discrepancies with respect to a conservation equation and some prior estimates of particle positions. This work has led to one publication in an international conference (PIV13) [27] and one publication in a national conference (Fluvisu13) [31].

Since October 2013, with our new postdoctoral fellow Kai Berger, we have started a new direction of research targeting the volume reconstruction problem. In particular, we address the question of devising effective reconstruction procedures taking into account the limited computational budget available in practice. Our approach is based on the design of simple thresholding operators, allowing to reduce the dimension of the initial problem and amenable to fast parallel implementations.

6.1.3. Motion estimation techniques for turbulent fluid flows

Participants: Patrick Héas, Dominique Heitz, Cédric Herzet, Etienne Mémin.

In this study we have devised smoothing functional adapted to the multiscale structure of homogeneous turbulent flows. These regularization constraints ensue from a classical phenomenological description of turbulence. The smoothing is in practice achieved by imposing some scale-invariance principles between histograms of motion increments computed at different scales. Relying on a Bayesian formulation, an inference technique, based on likelihood maximization and marginalization of the motion variable, has been proposed to jointly estimate the fluid motion, the regularization parameters and a proper physical models. The performance of the proposed Bayesian estimator has been assessed on several image sequences depicting synthetic and real turbulent fluid flows. The results obtained in the context of fully developed turbulence show that an improvement in terms of small-scale motion estimation can be achieved as compared to classical motion estimator. This work, performed within a collaboration with Pablo Mininni from the university of Buenos Aires, have been published in the IEEE Transactions on Pattern Analysis And Machine Learning [22].

6.1.4. Wavelet basis for multiscale motion estimation

Participants: Patrick Héas, Cédric Herzet, Etienne Mémin.

In this study we focused on the implementation of a simple wavelet-based optical-flow motion estimator dedicated to the recovery of fluid motions. The wavelet representation of the unknown velocity field is considered. This scale-space representation, associated to a simple gradient-based optimization algorithm, sets up a natural multiscale/multigrid optimization framework for the optical flow estimation that can be combined to more traditional incremental multiresolution approaches. Moreover, a very simple closure mechanism, approximating locally the solution by high-order polynomials, is provided by truncating the wavelet basis at intermediate scales. This offers a very interesting alternative to traditional Particle Image Velocimetry techniques. As another alternative to this medium-scale estimator, we explored strategies to define estimation at finer scales. These strategies rely on the encoding of high-order smoothing functional on divergence free wavelet basis. This study has been published in the journal of Numerical Mathematics: Theory, Methods and Applications [19] and in the international Journal of Computer Vision [23]. This work has strongly benefited from a collaboration with Souleyman Kadri-Harouna (University of La Rochelle and who was formerly on a post-doctoral position in our team). The divergence free wavelets basis proposed in [24] constitutes the building blocks on which we have elaborated our wavelet based motion estimation solutions. We have otherwise pursue our collaboration with Chico university through the post-doc of Pierre Dérian on the GPU implementation of such motion estimator for Lidar data.

6.1.5. Sparse-representation algorithms

Participant: Cédric Herzet.

The paradigm of sparse representations is a rather new concept which turns out to be central in many domains of signal processing. In particular, in the field of fluid motion estimation, sparse representation appears to be potentially useful at several levels: i) it provides a relevant model for the characterization of the velocity field in some scenarios; ii) it plays a crucial role in the recovery of volumes of particles in the 3D Tomo-PIV problem.

Unfortunately, the standard sparse representation problem is known to be NP hard. Therefore, heuristic procedures have to be devised to try and access to the solution of this problem. Among the popular methods available in the literature, one can mention orthogonal matching pursuit, orthogonal least squares and the family of procedures based on the minimization of ℓ_p norms. In order to assess and improve the performance of these algorithms, theoretical works have been undertaken in order to understand under which conditions these procedures can succeed in recovering the "true" sparse vector.

This year, we have contributed to this research axis by deriving conditions of success for the algorithms mentioned above when some partial information is available about the position of the nonzero coefficients in the sparse vector. This paradigm is of interest in the Tomographic-PIV volume reconstruction problem: one can indeed expect volumes of particles at two successive instants to be pretty similar; any estimate of the position of the particles at one given instant can therefore serve as a prior estimate about their position at the next instant. The conditions of success of such procedure have been rigorously formalized in two publications in the IEEE Transactions on Information Theory [21], [26] and one publication in an international conference (SPARS13) [28].

6.2. Tracking, Data assimilation and model-data coupling

6.2.1. Stochastic filtering for fluid motion tracking

Participants: Sébastien Beyou, Anne Cuzol, Etienne Mémin.

Within the PhD thesis of Sébastien Beyou [11], we investigated the study of a recursive Bayesian filter for tracking velocity fields of fluid flows. We resort in this study to Monte-Carlo approximations based on the particle filtering paradigm. In particular, we investigated the use of the so-called ensemble Kalman filtering for fluid tracking problems. This kind of filters introduced for the analysis of geophysical fluids is based on the Kalman filter update equations. Nevertheless, unlike traditional Kalman filtering setting, the covariances of the estimation errors, required to compute the so-called Kalman gain, relies on an ensemble of forecasts. Such a process gives rise to a Monte-Carlo approximation for a family of non-linear stochastic filters enabling to handle state spaces of large dimension. The method we proposed can be seen as an extension of this technique that combines sequential importance sampling and the propagation law of an ensemble Kalman filter. This technique leads to an ensemble Kalman filter with an improved efficiency. Within this type of scheme, we have in particular investigated the introduction of a nonlinear direct image measurement operator. This modification of the filter provides very good results on 2D numerical and experimental flows even in the presence of strong noises. We assessed successfully its application to oceanic satellite images for the recovering of ocean streams. We have also studied the impact on the stochastic dynamics of auto-similar Gaussian noise mimicking statistical properties of turbulence and the introduction within an incremental ensemble analysis scheme of multiscale motion measurements. This work has been published in the Tellus A journal [15].

6.2.2. Stochastic filtering technique for the tracking of closed curves

Participant: Etienne Mémin.

We have studied a stochastic filtering technique for the tracking of closed curves along an image sequence. In that aim, we designed a continuous-time stochastic dynamics that allows us to infer inter-frame deformations. The curve is defined by an implicit level-set representation and the stochastic dynamics is expressed properly on the level-set function. It takes the form of a stochastic partial differential equation with a Brownian motion of low dimension. The evolution model we proposed combines local photometric information, deformations induced by the curve displacement and an uncertainty modeling of the dynamics. Specific choices of noise models and drift terms lead to an evolution law based on mean curvature as in classic level set methods, while other choices yield new evolution laws. The approach we propose is implemented through a particle filter, which includes color measurements characterizing the target and the background photometric probability densities respectively. The merit of this parameter free filter is demonstrated on various satellite image sequences depicting the evolution of complex geophysical flows. This work has been recently published in the Journal of Mathematical Imaging and Vision [13]. Let us note the method provides an empirical dynamical model learned recursively from a data flow. Its short time forecasting skills have been used in the context of weather-watch radar images within a fruitful collaboration with MeteoFrance.

6.2.3. Sequential smoothing for fluid motion

Participants: Anne Cuzol, Etienne Mémin.

In parallel to the construction of stochastic filtering techniques for fluid motions, we have proposed a new sequential smoothing method within a Monte-Carlo framework. This smoothing aims at reducing the temporal discontinuities induced by the sequential assimilation of discrete time data into continuous time dynamical models. The time step between observations can indeed be long in environmental applications for instance, and much longer than the time step used to discretize the model equations. While the filtering aims at estimating the state of the system at observations times in an optimal way, the objective of the smoothing is to improve the estimation of the hidden state between observation times. The method is based on a Monte-Carlo approximation of the filtering and smoothing distributions, and relies on a simulation technique of conditional diffusions. The proposed smoother can be applied to general non linear and multidimensional models. It has been applied to a turbulent flow in a high-dimensional context, in order to smooth the filtering results obtained from a particle filter with a proposal density built from an Ensemble Kalman procedure. This conditional simulation framework can also be used for filtering problem with low measurement noise. This has been explored through a collaboration with Jean-Louis Marchand (ENS Bretagne) in the context of vorticity tracking from image data.

6.2.4. Stochastic fluid flow dynamics under uncertainty

Participants: Etienne Mémin, Valentin Resseguier.

In this research axis we aim at devising Eulerian expressions for the description of fluid flow evolution laws under uncertainties. Such an uncertainty is modeled through the introduction of a random term that allows taking into account large-scale approximations or truncation effects performed within the dynamics analytical constitution steps. This includes for instance the modeling of unresolved scales interaction in large eddies simulation (LES) or in Reynolds average numerical simulation (RANS), but also uncertainties attached to non-uniform grid discretization. This model is mainly based on a stochastic version of the Reynolds transport theorem. Within this framework various simple expressions of the drift component can be exhibited for different models of the random field carrying the uncertainties we have on the flow. We aim at using such a formalization within image-based data assimilation framework and to derive appropriate stochastic versions of geophysical flow dynamical modeling. This formalization has been published in the journal Geophysical and Astrophysical Fluid Dynamics [25]. Numerical simulation on divergence free wavelets basis of 3D viscous Taylor-Green vortex and Crow instability have been performed within a collaboration with Souleymane Kadri-Harouna. First promising results have been published in the TSFP8 conference [30]. Besides, we explore in the context of Valentin Resseguier's PhD the extension of such framework to oceanic models and to satellite image data assimilation. This PhD thesis takes place within a fruitful collaboration with Bertrand Chapron (CERSAT/IFREMER).

6.2.5. Free surface flows reconstruction and tracking

Participants: Dominique Heitz, Etienne Mémin, Cordelia Robinson, Yin Yang.

Characterizing a free-surface flow (space and time-dependent velocity and geometry) given observations/measures at successive times is an ubiquitous problem in fluid mechanic and in hydrology. Observations can consist of e.g. measurements of velocity, or like in this work of measurements of the geometry of the free-surface. Indeed, recently developed depth/range sensors allow to capture directly a rough 3D geometry of surfaces with high space and time resolution. We have investigated the performance of the Kinect and have shown that it is likely to capture temporal sequences of depth observations of wave-like surfaces with wavelengths and amplitudes sufficiently small to characterize medium/large scale flows. Several data assimilation methods have been experimented and compared to estimate both time dependent geometry and displacement field associated to a free-surface flow from a temporal sequence of Kinect data. This study has been conducted on synthetic and real-world data. Finally, we explored the application of such techniques to hydrological applications. These results have been recently submitted to Journal of Computational Physics.

6.2.6. Variationnal ensemble methods for data assimilation

Participants: Dominique Heitz, Etienne Mémin, Cordelia Robinson, Yin Yang.

In this work, we aim at studying an ensemble based optimal control strategy for data assimilation. Such a formulation nicely combines the ingredients of ensemble Kalman filters and variational assimilation. In the same way as standard variational assimilation, it is formulated as the minimization of an objective function. However, similarly to ensemble filters, it introduces in its objective function an empirical ensemble-based background-error covariance and works in an off-line smoothing mode rather than sequentially like filtering approaches in a sequential filter. These techniques have the great advantage to avoid the introduction of tangent linear and adjoint models, which are necessary for standard incremental variational techniques. As the background error covariance matrix plays a key role in the variational process, our study particularly focuses on the generation of the analysis ensemble state with localization techniques. We compared the performances of both methods in different cases in which the system's component are fully observed or only partially. The comparisons have been leaded on the basis of a Shallow Water model.

6.2.7. Optimal control techniques for the coupling of large scale dynamical systems and image data

Participants: Dominique Heitz, Etienne Mémin, Cordelia Robinson.

This work aims at investigating the use of optimal control techniques for the coupling of Large Eddies Simulation (LES) techniques and 2D image data. The objective is to reconstruct a 3D flow from a set of simultaneous time resolved 2D image sequences visualizing the flow on a set of 2D plans enlightened with laser sheets. This approach will be experimented on shear layer flows and on wake flows generated on the wind tunnel of Irstea Rennes. Within this study we wish also to explore techniques to enrich large-scale dynamical models by the introduction of uncertainty terms or through the definition of subgrid models from the image data. This research theme is related to the issue of turbulence characterization from image sequences. Instead of predefined turbulence models, we aim here at tuning from the data the value of coefficients involved in traditional LES subgrid models or in longer-term goal to learn empirical subgrid models directly from image data. An accurate modeling of this term is essential for Large Eddies Simulation as it models all the non resolved motion scales and their interactions with the large scales.

We have pursued the first investigations on a 4DVar assimilation technique, integrating PIV data and Direct Numerical Simulation (DNS), to reconstruct two-dimensional turbulent flows. The problem we are dealing with consists in recovering a flow obeying Navier-Stokes equations, given some noisy and possibly incomplete PIV measurements of the flow. By modifying the initial and inflow conditions of the system, the proposed method reconstructs the flow on the basis of a DNS model and noisy measurements. The technique has been evaluated in the wake of a circular cylinder. It denoises the measurements and increases the spatiotemporal resolution of PIV time series. These results have been recently published in the Journal of Computational Physics [20]. A paper has been also recently published on the denoising aspect in the (PIV13) international conference [29]. Along the same line of studies the 3D case is ongoing. The goal consists here to reconstruct a 3D flow from a set of simultaneous time resolved 2D images of planar sections of the 3D volume. This work is mainly conducted within the PhD of Cordelia Robinson. The development of the variational assimilation code has been initiated within a collaboration with A. Gronskis, S. Laizé (lecturer, Imperial College, UK) and Eric Lamballais (institut P' Poitiers). A High Reynolds number simulation of the wake behind a cylinder has been recently performed within this collaboration.

6.2.8. Variational assimilation of images for large scale fluid flow dynamics with uncertainty

Participant: Etienne Mémin.

In this work we explore the assimilation of a large scale representation of the flow dynamics with image data provided at a finer resolution. The velocity field at large scales is described as a regular smooth components whereas the complement component is a highly oscillating random velocity field defined on the image grid but living at all the scales. Following this route we have started to assess the performance of a variational assimilation technique with direct image data observation. Preliminary encouraging results obtained for a wavelet-based 2D Navier Stokes implementation and images of a passive scalar transported by the flow have been obtained. Large-scale simulation under uncertainty for the 3D viscous Taylor-Green vortex flow have been carried out and show promising results of the approach.

6.2.9. *Reduced-order models for flows representation from image data*

Participants: Cédric Herzet, Etienne Mémin, Véronique Souchaud.

One of the possibilities to neglect the influence of some degrees of freedom over the main characteristics of a flow consists in representing it as a sum of K orthonormal spatial basis functions weighted with temporal coefficients. To determine the basis function of this expansion, one of the usual approaches relies on the Karhunen-Loeve decomposition (refered to as proper orthogonal decomposition – POD – in the fluid mechanics domain). In practice, the spatial basis functions, also called modes, are the eigenvectors of an empirical auto-correlation matrix which is built from “snapshots" of the considered physical process.

In this axis of work we focus on the case where one does not have a direct access to snapshots of the considered physical process. Instead, the POD has to be built from the partial and noisy observation of the physical phenomenon of interest. Instances of such scenarios include situations where real instantaneous vector-field snapshots are estimated from a sequence of images. We have been working on several approaches dealing with such a new paradigm. A first approach consists in extending standard penalized motion-estimation algorithms to the case where the sought velocity field is constrained to span a low-dimensional subspace. In particular, we have considered scenarios where the standard optical flow constraint (OFC) is no longer satisfied and one has therefore to resort to a Discrete Finite Difference (DFD) model. The non-linearity of the latter leads to several practical issues that we have addressed this year.

Within a collaboration with the University of Buenos Aires, we have also explored, a method that combines Proper Orthogonal Decomposition with a spectral technique to analyze and extract reduced order models of flows from time resolved data of velocity fields. This methodology, relying on the eigenfunctions of the Koopman operator, is specifically adapted to flows with quasi periodic orbits in the phase space. The technique is particularly suited to cases requiring a discretization with a high spatial and temporal resolution. The proposed analysis enables to decompose the flow dynamics into modes that oscillate at a single frequency. For each modes an energy content and a spatial structure can be put in correspondence. This approach has been assessed for a wake flow behind a cylinder at Reynolds number 3900 and has been recently published to the journal of Theoretical and Computational Fluid Dynamics [16]. The assessment of this method on oceanic model simulation data is on going.

6.3. Analysis and modeling of turbulent flows

6.3.1. *Hot-wire anemometry at low velocities*

Participant: Dominique Heitz.

A new dynamical calibration technique has been developed for hot-wire probes. The technique permits, in a short time range, the combined calibration of velocity, temperature and direction calibration of single and multiple hot-wire probes. The calibration and measurements uncertainties were modeled, simulated and controlled, in order to reduce their estimated values. Based on a market study the french patent application has been extended this year to a Patent Cooperation Treaty (PCT) application.

6.3.2. *Numerical and experimental image and flow database*

Participant: Dominique Heitz.

The goal was to design a database for the evaluation of the different techniques developed in the Fluminance group. The main challenge was to enlarge a database mainly based on two-dimensional flows, with three-dimensional turbulent flows. New synthetic image sequences based on homogeneous isotropic turbulence and on circular cylinder wake have been provided. These images have been completed with real image sequences based on wake and mixing layers flows. This new database provides different realistic conditions to analyse the performance of the methods: time steps between images, level of noise, Reynolds number, large-scale images. A Wake flow at high Reynolds number has been also simulated on one of the IDRISS super computer. This simulation, whose results analysis is on going, has been performed within a collaboration with Sylvain Laizet (Imperial College).

6.4. Visual servoing approach for fluid flow control

6.4.1. Minimization of the kinetic energy density in the 2D plane Poiseuille flow

Participants: Christophe Collewet, Xuan Quy Dao.

This works concerns the PhD thesis of Xuan-Quy Dao. This year we have focused on a way to ensure a strict decreasing of the kinetic energy density. In that purpose, we have first proposed an approach to increase the controlled degrees of freedom. Indeed, the classical way to model this flow leads to only two degrees of freedom. With so few degrees of freedom it is obviously impossible to reach high desired performances as the strict minimization of the kinetic energy density. This way to proceed leads to a better minimization of the kinetic energy density. We have also proposed on approach based on a local decoupling of the controlled degree of freedom of the system so that an exponential decoupled decrease of each components of the state vector is locally obtained. This work has been presented at the CFM conference (Congrés Français de Mécanique) [32].

6.4.2. Control behind a backward-facing step

Participant: Christophe Collewet.

This work is performed in the context of the PhD thesis of Nicolas Gautier from ESCPI in collaboration with J.L. Aider. The separated flow downstream a backward-facing step is studied using visual information for feedback. More precisely, flow velocity fields are computed from a real time optical flow algorithm. The control law we used is a simple PID controller. Even a better control law could be used, this study validates that visual servo control is an effective approach to control a flow.

6.4.3. Control of systems described by partial differential equations

Participants: Tudor-Bogdan Airimitoaie, Christophe Collewet.

This work concerns principally the post-doctoral research of Tudor-Bogdan Airimiţoaie. It aims at controlling continuously evolving systems described by partial differential equations (PDEs). This is relevant in the context of the Fluminance team because fluid flows are infinite dimensional systems and can be rigorously described only through PDEs. In spite of this, practical approaches of flow control are based on low order numerical implementation relying on space and time discretization of the continuous system. This implies to setup strategies for model reduction that must be then in return properly understood with respect to the convergence of the control law. For finite dimensional implementations, one of the research directions pursued concerns the study on the benefit of increasing the controlled degrees of freedom (see the work of Xuan-Quy Dao). Another research direction, started recently, consists in improving control by using real-time estimation of a finite number of parameters related to the original infinite dimensional system. Indeed, this opens the possibility of improving performances by using more advanced robust linear parametric varying (LPV) control techniques existing in the literature. Two conference papers on these works have been submitted at the "7th AIAA Flow Control Conference".

MAGIQUE-3D Project-Team

6. New Results

6.1. Inverse Problems

6.1.1. Reconstruction of an elastic scatterer immersed in a homogeneous fluid

Participants: Hélène Barucq, Rabia Djellouli, Élodie Estécahandy.

The determination of the shape of an obstacle from its effects on known acoustic or electromagnetic waves is an important problem in many technologies such as sonar, radar, geophysical exploration, medical imaging and nondestructive testing. This inverse obstacle problem (IOP) is difficult to solve, especially from a numerical viewpoint, because it is ill-posed and nonlinear. Its investigation requires as a prerequisite the fundamental understanding of the theory for the associated direct scattering problem, and the mastery of the corresponding numerical solution methods.

In this work, we are interested in retrieving the shape of an elastic obstacle from the knowledge of some scattered far-field patterns, and assuming certain characteristics of the surface of the obstacle. The corresponding direct elasto-acoustic scattering problem consists in the scattering of time-harmonic acoustic waves by an elastic obstacle Ω^s embedded in a homogeneous medium Ω^f, that can be formulated as follows:

$$\begin{array}{rl} \Delta p + (\omega^2/c_f^2)\, p = 0 & \text{in } \Omega^f \\ \nabla\cdot\sigma(u) + \omega^2\rho_s\, u = 0 & \text{in } \Omega^s \\ \omega^2\rho_f u\cdot n = \partial p/\partial n + \partial e^{i\,(\omega/c_f)\,x\cdot d}/\partial n & \text{on } \Gamma \\ \sigma(u)n = -pn - e^{i\,(\omega/c_f)\,x\cdot d} n & \text{on } \Gamma \\ \lim\limits_{r\to+\infty} r\,(\partial p/\partial r - i\,(\omega/c_f)\,p) = 0 & \end{array} \tag{11}$$

where p is the fluid pressure in Ω^f whereas u is the displacement field in Ω^s, and $\sigma(u)$ represents the stress tensor of the elastic material.

This boundary value problem has been investigated mathematically and results pertaining to the existence, uniqueness and regularity can be found in [92] and the references therein, among others. We have obtained a new result proving the well-posedness of the problem when the fluid-solid interface is only lipschitzian. This has been published in the Journal of Mathematical Analysis and Applications [20]. We then propose a solution methodology based on a regularized Newton-type method for solving the IOP. The proposed method is an extension of the regularized Newton algorithm developed for solving the case where only the Helmholtz equation is involved, that is the acoustic case by impenetrable scatterers [86]. The direct elasto-acoustic scattering problem defines an operator $F:\Gamma\to p_\infty$ which maps the boundary Γ of the scatterer Ω^s onto the far-field pattern p_∞. Hence, given one or several measured far-field patterns $\widetilde{p}_\infty(\widehat{x})$, corresponding to one or several given directions d and wavenumbers k, one can formulate IOP as follows:

$$\text{Find a shape } \Gamma \text{ such that } \quad F(\Gamma)(\widehat{x}) = \widetilde{p}_\infty(\widehat{x}); \quad \widehat{x}\in S^1.$$

At each Newton iteration, we solve the forward problem using a finite element solver based on discontinuous Galerkin approximations, and equipped with high-order absorbing boundary conditions. We have first characterized the Fréchet derivatives of the scattered field and the characterization has been published in the Journal of Inverse and Ill-posed problems [18]. It is worth noting that they are solutions to the same boundary value problem as the direct problem with other transmission conditions. This work has been the object of several talks [63], [50], [36]. Elodie Estécahandy has defended her PhD thesis [14] in September 2013 and two papers will be submitted soon.

6.1.2. hp-adaptive inversion of magnetotelluric measurements

Participants: Hélène Barucq, Julen Alvarez Aramberri, David Pardo.

The magnetotelluric (MT) method is a passive electromagnetic exploration technique. It makes use of natural electric fields which propagate permanently into the Earth. Electric fields induce magnetic waves which can be detected at the surface to produce a map of the subsurface from the determination of the resistivity distribution. Magnetotelluric method is based on the mathematical relation between the magnetic and telluric variations which involve the electric resistivity of the subsurface. It is particularly relevant for the detection of metallic ores and for the study of geothermal sites. It is also used for oil and gas exploration because it provides information on sedimentary basins. It performs well on depth scales varying between a few tens of meters to hundred of kilometers, following the pioneering works of Tikhonov and Cagniard . Magnetotelluric measurements are governed by polarized Maxwell's equations in such a way that Helmholtz equations have to be solved. The geological mapping is constructed from the solution of an Inverse problem which requires computing the Impedance and/or the Resistivity distributions. In this work, we assimilate Earth with a horizontally layered model with possible 2D heterogeneities. Both the size of the direct problem and the required computational times may be excessively large. Indeed, on the one hand, the model of the source requires defining a horizontally sufficiently large thick plate to avoid undesirable effects that could take place around the edges. On the other hand, the inversion of MT measurements typically requires the computation of an accurate solution at the receivers located at different positions. Since traditional hp-goal oriented techniques [98], [97] provide an accurate solution in one single point, we use a multi-goal-oriented algorithm [99] to obtain accurate solutions at all receivers. To get accurate quantities at several positions, it is necessary to increase the size of the mesh. This induces high computational costs in particular because the solution of the inverse problem is based on reiterated solutions of the direct problem. To decrease the computational costs required to perform the inversion, we propose an adaptive multi-dimensional inversion algorithm, which consists in increasing step by step the dimension in which the direct problem and the inversion are solved. At first step, we compute the 1D primary field with a semi-analytical solution and we invert the 1D problem. After that, we introduce the 2D heterogeneities. Regarding the direct problem, we compute the secondary field, thereby, drastically reducing the size of the computational domain for this problem. Then, we perform the inversion using the solution to the 1D Inverse Problem as a regularization term, increasing the robustness of the inversion algorithm.

6.1.2.1. *Reverse Time Migration with Elastic Wave Equations*

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

Even if RTM has enjoyed the tremendous progresses of scientific computing, its performances can still be improved, in particular when applied to strong heterogeneous media. In this case, images have been mainly obtained by using direct arrivals of acoustic waves and the transition to elastic waves including multiples is not obvious, essentially because elastic waves equations are still more computationally consuming. We have thus chosen to consider high-order Discontinuous Galerkin Methods which are known to be well-adapted to provide accurate solutions based upon parallel computing. Now, one of the main drawback of RTM is the need of storing a huge quantity of information which is prohibitive when using elastic waves. For that purpose, we apply the Griewank algorithm [88] following Symes' ideas [101] for the acoustic RTM. The idea is to find a compromise between the number of wave equations to solve and the number of numerical waves that we have to store. This is the so-called Optimal Checkpointing. By reducing the occupancy of the memory, RTM should be efficient even when using elastic waves. The next step is the derivation of accurate imaging conditions, which could take advantage of all the information contained in the elastic wavefield. For acoustic media, Claerbout [83] proposed an imaging condition which is widely used and turns out to be sufficient to accurately reproduce interfaces. But Claerbout conditions do not take wave conversions into account and, since P-wave and S-wave interact with each other, it might be relevant to use an imaging condition including these interactions. This has been done successfully by J. Tromp and C. Morency [102] for seismology applications based upon the inversion of the global Earth. Their approach is based upon the adjoint state and it involves sensitivity kernels which are defined from the propagated and the back-propagated fields. Now, it has been shown in [93] that full wave form inversions using these sensitivity kernels may be polluted by numerical artifacts. One solution is to use a linear combination of the sensitivity kernels to delete artifacts. In this work, we propose then a new imaging condition which construction is inspired from [93] with some approximations required to keep admissible computational costs. We illustrate the properties of the new imaging condition on industrial benchmarks like the Marmousi model. In particular, we compare the new imaging condition with other imaging conditions by using as criteria the quality of the image and the computational costs required by the RTM. The results will be presented at the 2014 ECCOMAS conference in Barcelona.

6.2. Modeling

6.2.1. Implementation of a non-reflecting boundary condition on ellipsoidal boundary

Participants: Hélène Barucq, Anne-Gaëlle Saint-Guirons, Sébastien Tordeux.

The modeling of wave propagation problems using finite element methods usually requires the truncation of the computational domain around the scatterer of interest. Absorbing boundary condition are classically considered in order to avoid spurious reflections. This year we have proposed and tested a formulation which allows to take into account with no extra-cost a quasi-exact radiation condition based on a non local Dirichlet to Neumann operator.

6.2.2. Modeling of small heterogeneities in the context of the time domain wave equation

Participants: Vanessa Mattesi, Sébastien Tordeux.

We have proposed an approximate model to take into account small heterogeneities for the three dimensional time dependent wave equation. One of the most important result of this work is the generalization of the multipole theory (classically written for the Helmholtz equation) to the wave equation. This work has been presented at Waves 2013 and at JSA 2013 [58], [39]

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

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

In order to obtain high-order time-schemes, we are considering an alternative approach to the ADER schemes and to the modified equation technique described in section 3.2 . The two first steps of the construction of the schemes are similar to the previous schemes : we apply a Taylor expansion in time to the solution of the wave equation and we replace the high-order derivatives with respect to the time by high order space operators, using the wave equation. The difference is that we do not use auxiliary variables and we choose to discretize directly the high-order operators in space. These operators can not be discretized by classical finite elements. For the discretization of the biharmonic operator in an homogeneous acoustic medium, both C1 finite elements, such as the Hermite ones, and Discontinuous Galerkin Finite Elements (DGFE) can be used, while in a discontinuous medium, or for higher-order operators, DGFE should be preferred [80]. We have applied this method to the second order wave equation [15] and the numerical results showed that this technique induced less computational burden than the modified equation scheme or the ADER scheme.

In the framework of the PhD thesis of Florent Ventimiglia, we have extended the new method involving p-harmonic operator to the first order formulation of the acoustic wave equation, which is the formulation discretized in the DIVA platform of TOTAL. In this case, the high order operators in space are not are not powers of the Laplace operator but powers of the gradient. Hence, we also had to adapt the space discretization, and we have extended the DG formulation with centered fluxes proposed in [87] to higher order operators. A numerical analysis of performance in 2D indicates that, for a given accuracy, this method requires less computational costs and less storage than the High-Order ADER Scheme. These results have been presented to the SMAI conference [49], Waves 2013 [61], Numerico IV [61] and HF2013 [34]. A paper has been accepted in ESAIM Proceedings [49].

6.2.4. Constructing and using Absorbing Boundary Conditions

6.2.4.1. Higher Order On-Surface Radiation Conditions for elastic scatterers

Participants: Hélène Barucq, Chokri Bekkey, Juliette Chabassier, Julien Diaz.

The numerical simulation of wave propagation is generally performed by truncating the propagation medium and the team works on new Absorbing Boundary Conditions (ABCs), trying to improve the performance of existing conditions. As we explained at Section 3.2 , item **Boundary conditions**, we are developing ABCs for curved boundaries, based on the full factorization of the wave equation. These ABCs should take propagating, grazing and evanescent waves into account. In [17], we have considered the issue of constructing high-order ABCs taking into account both propagating and evanescent waves for the Helmholtz equation. In case of the simulation of acoustic waves diffracted by a solid immersed in a fluid, we investigate the performance of the new ABCs when used as On-Surface Radiation Conditions. The ABCs are set directly on the boundary of the solid. The unbounded problem is then replaced by a problem involving an acoustic pressure computed on the surface of the solid only. Preliminary results have been obtained by considering the toy problem where the scatterer is a disk. Analytic solutions are then available and we show that that taking into account evanescent waves in the ABC could improve the accuracy of classical ABCs by two orders of magnitude at mid-frequency range, for ka between 1 and 100, k being the frequency and a the typical size of the diffracting obstacle. These results have been presented to the Waves 2013 conference [47].

6.2.4.2. Radiation boundary condition at high frequency

Participants: Hélène Barucq, Elodie Estécahandy, Juliette Chabassier, Julien Diaz.

Regarding the solution of the Helmholtz equation at high frequency with finite element methods, it is current to refine the mesh in order to limit the effect of numerical pollution. It is then interesting to dispose of radiation boundary conditions which do not require to set the artificial boundary far from the scatterer. In this work, we have investigated the possibility of bringing closer the artificial boundary when standard conditions are enhanced by the modeling of grazing waves. The preliminary results we obtained show that this new ABC outperforms classical ABCs at high-frequency, for $ka > 50$, and that it is highly accurate, even when the boundary is very close to the scatterer. These results have been presented to the Waves 2013 conference [45]. Now, the next step is to consider an ABC taking the three types of waves into account.

6.2.4.3. Absorbing Boundary Conditions for Tilted Transverse Isotropic Elastic Media

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

The simulation of wave propagation in geophysical media is often performed in domains which are huge compared to the wavelengths of the problem. It is then necessary to reduce the computational domain to a box. When considering acoustic or elastic isotropic media, this can be done by applying an Absorbing Boundary Condition (ABC) or by adding a Perfectly Matched Layer (PML). However, a realistic representation of the Earth subsurface must include anisotropy and, in particular, the so-called Tilted Transverse Isotropy. Perfectly Matched Layers are known to be unstable for this kind of media and, to the best of our knowledge, no ABC have been proposed yet. We have thus proposed a low-order ABC for TTI media. The construction is based on comparing and then connecting the slowness curves for isotropic and elliptic TTI waves. Numerical experiments illustrate the performance of the new ABC. They are performed by integrating the ABC in a DG formulation of Elastodynamics. When applied in a TTI medium, the new ABC performs well with the same level of accuracy than the standard isotropic ABC set in an isotropic medium. The condition demonstrates also a good robustness when applied for large times of simulation. These results have been presented to the Smai and to the Waves conferences [67], [68] and a paper has been submitted.

6.2.5. Modeling of the damping factor of Multiperforated plates

Participants: Estelle Piot, Vincent Popie, Sébastien Tordeux.

Multiperforated plates are classically used as a damping material. Melling has proposed in [94] a model to estimate the energy dissipated by these devises. However its well-known result should be corrected by a factor two to fit with experimental data. We have proposed a correct way to compute the energy dissipated by the multiperforated plates. This work has been presented at the Fifth International Scientific Conference and Young Scientists School "Theory and Computational Methods for Inverse and Ill-posed Problems.

6.2.6. Performance Assessment of IPDG for the solution of an elasto-acoustic scattering problem

Participants: Hélène Barucq, Rabia Djellouli, Élodie Estécahandy.

We present a solution methodology for the direct elasto-acoustic scattering problem that falls in the category of Discontinuous Galerkin methods. The method distinguishes itself from the existing methods by combining high-order Discontinuous Galerkin approximations, local stabilizations for the coupled problem and the use of curved element edges on the boundaries. We present some numerical results that illustrate the salient features and highlight the performance of the proposed solution methodology on the resonance phenomenon existing in the elastic scatterer for simple geometries such as circles. Moreover, the designed method ensures a convergence order with a gain of two order of magnitude compared to polygonal boundaries, and a potential to address both mid- and high-frequency regimes. These results have been accepted for publication in International Journal for Numerical Methods in Engineering [19].

6.2.7. *On the influence of curvature on transmission conditions*

Participants: Hélène Barucq, Martin Gander, Yingxiang Xu.

Domain decomposition methods are both highly successful parallel solvers and also important modeling tools, since problems in subdomains can be treated by adapted methods to the physics in each subdomain. Subdomain boundaries are therefore rarely straight lines. The focus of this paper is to study the influence of curvature on transmission conditions used in optimized Schwarz methods. For straight interfaces and simple geometries, optimized interface conditions are typically determined using Fourier analysis.Asymptotically, these optimized conditions are still valid for curved interfaces.Since however the curvature is the most important information for a smooth curve, we want to study in this paper if and how the interface curvature influences the constants in the optimized parameters.

We consider the model problem

$$(\Delta - \eta)u = f, \quad \text{on } \Omega = \mathbb{R}^2,\ \eta > 0, \tag{12}$$

and we require the solution to decay at infinity.We decompose Ω into two overlapping subdomains $\Omega_1 = (-\infty, a(y)) \times \mathbb{R}$ and $\Omega_2 = (b(y), \infty) \times \mathbb{R}$, where Γ_1 given by $a(y)$ and Γ_2 given by $b(y)$ are smooth curves satisfying $a(y) \geq b(y)$. A general parallel Schwarz algorithm is then given by

$$\begin{array}{rcll} (\Delta - \eta)u_i^n & = & f & \text{in } \Omega_i, \\ \mathcal{B}_i(u_i^n) & = & \mathcal{B}_i(u_j^{n-1}) & \text{on } \Gamma_i,\ 1 \leq i \neq j \leq 2, \end{array} \tag{13}$$

where $\mathcal{B}_i, i = 1, 2$, are transmission conditions to be chosen. If $\mathcal{B}_i$, $i = 1, 2$ are chosen as $\partial_{n_i} + DtN_i$, with DtN_i the Dirichlet to Neumann operators, the iterates will converge in two steps. These operators are however non-local, and thus difficult to use in practice. Therefore, local approximations are used in optimized Schwarz methods. We presented two different approaches to take the curvature of interfaces into account in the transmission conditions of optimized Schwarz methods: micro-local analysis, and analysis using a circular model problem. In both cases, we obtained curvature dependent transmission conditions. A preliminary comparison shows that the transmission conditions based on optimization perform better on the model problem, and that it could be important to take the curvature into account in transmission conditions. In our opinion it is however essential to do a more thorough theoretical and numerical study on more general geometry, where micro-local analysis is still applicable, before we can definitely draw conclusions. This work has been published in the peer-reviewed proceedings of the conference Decomposition Methods in Science and Engineering XXI [51].

6.2.8. *Operator Based Upscaling for Discontinuous Galerkin Methods*

Participants: Hélène Barucq, Théophile Chaumont, Christian Gout.

Realistic numerical simulations of seismic wave propagation are difficult to handle because they must be performed in strongly heterogeneous media. Two different scales must then be taken into account. Indeed, the medium heterogeneities are currently very small compared to the characteristic dimensions of the propagation medium. To get accurate numerical solutions, engineers are then forced to use meshes that match the finest scale representing the heterogeneities. Meshing the whole domain with a fine grid leads then to huge linear systems and the computational cost of the numerical method is then too high to consider 3D realistic simulations. To dispose of a numerical method allowing to represent the heterogeneity of the medium accurately while computing on a coarse grid is thus relevant. This is the challenge of multiscale approaches like homogenization or upscaling. The ultimate objective of this work is to develop a software package for the Helmholtz equation set in heterogeneous domains. We focus on the operator-based upscaling method. Operator-based upscaling methods were first developed for elliptic flow problems (see [81]) and then extended to hyperbolic problems (see [90], [104], [103]). Operator-based upscaling method consists in splitting the solution into a coarse and a fine part. The coarse part is defined on a coarse mesh while the fine part is computed on a fine mesh. In order to speed up calculations, artificial boundary conditions (ABC) are imposed. By enforcing suitable ABCs on the boundary of every cells of the coarse mesh, calculations on the fine grid can be carried out locally. The coarse part is next computed globally on the coarse mesh. Operator-based upscaling methods were so far developed in joint with standard finite element discretization strategy. In this work, we investigate the idea of combining an operator based upscaling method with discontinuous Galerkin finite element methods(DGFEM). During this year, we have performed the mathematical analysis of the Helmholtz equation set in a domain represented by a discontinuous velocity. The analysis has been achieved both for the continuous and the discretized problem which is based on a quadrature scheme which allows to take the discontinuities of the velocity into account. We get new stability results for stratified-like domains and numerical experiments show that in case of industrial benchmarks, the quadrature scheme leads to lower computational costs with a very good level of accuracy. A paper is in preparation and the results will be presented at the 2014 ECCOMAS conference in Barcelona.

6.2.9. Efficient solution methodology based on a local wave tracking strategy for high-frequency Helmholtz problems.

Participants: Mohamed Amara, Sharang Chaudhry, Julien Diaz, Rabia Djellouli, Steven Fiedler.

We have proposed a procedure for selecting basis function orientation to improve the efficiency of solution methodologies that employ local plane-wave approximations. The proposed adaptive approach consists of a local wave tracking strategy. Each plane-wave basis set, within considered elements of the mesh partition, is individually or collectively rotated to best align one function of the set with the local propagation direction of the field. Systematic determination of the direction of the field inside the computational domain is formulated as a minimization problem. As the resultant system is nonlinear with respect to the directions of propagation, the Newton method is employed with exact characterization of the Jacobian and Hessian. To illustrate the salient features and evaluate the performance of the proposed wave tracking approach, we present error estimates as well as numerical results obtained by incorporating the procedure into a prototypical plane-wave based approach, the least-squares method (LSM) developed by Monk and Wang [95]. The numerical results obtained for the case of a two-dimensional rigid scattering problem indicate that (a) convergence was achievable to a prescribed level of accuracy, even upon initial application of the tracking wave strategy outside the pre-asymptotic convergence region, and (b) the proposed approach reduced the size of the resulting system by up to two orders of magnitude, depending on the frequency range, with respect to the size of the standard LSM system. These results has been presented to the Waves 2013 [43]and in a Research Report [71]. A paper has been submitted.

6.2.10. Mesh Free Frontier-based Formulation (MF3) for High Frequency Helmholtz Problems.

Participants: Mohamed Amara, Julien Diaz, Rabia Djellouli.

We have proposed a novel approach for solving efficiently Helmholtz problems. The proposed solution method employs a boundary-type formulation without however involving Green functions and/or incurring singular integrals. In addition, this approach does not necessitate the use of a mesh. For these reasons, the method is named Mesh Free Frontier-based Formulation (MF3). Furthermore, the sought-after field is locally approximated using a set of basis of functions that consists of a Bessel kind function computed at a prescribed finite set of points. The number of the functions determines mainly the size of the resulting system, the complexity, as well as the computational cost of the proposed method. Preliminary numerical results obtained in the case of 2D-Helmholtz problems in the high-frequency regime are presented to illustrate the computational efficiency of MF3 (the method delivers results with high accuracy level, about 10^-8 on theL^2 relative error, while requiring the solution of small linear systems). In addition, these results tend to suggest that MF3 is pollution free. These results has been presented to two conferences [32], [44]. A paper is in preparation.

6.2.11. Energy based simulation of a Timoshenko beam in non-forced rotation. Application to the flexible piano hammer shank.

Participants: Juliette Chabassier, Marc Duruflé.

A nonlinear model for a vibrating Timoshenko beam in non-forced unknown rotation is derived from the virtual work principle applied to a system of beam with mass at the end. The system represents a piano hammer shank coupled to a hammer head. An energy-based numerical scheme is then provided, obtained by non classical approaches. A major difficulty for time discretization comes from the nonlinear behavior of the kinetic energy of the system. Numerical illustrations are obtained by coupling this new numerical scheme to a global energy-preserving numerical solution for the whole piano. These numerical results show that the pianistic touch clearly influences the spectrum of the piano sound of equally loud isolated notes. These differences do not come from a possible shock excitation on the structure, nor from a changing impact point, nor a “longitudinal rubbing motion” on the string, since neither of these features are modeled in our study.

This work has been submitted for publication in Journal of Sound and Vibration [79]. discretization

6.2.12. Simulating the propagation of ultra short laser pulses in a dispersive non linear medium.

Participants: Juliette Chabassier, Marc Duruflé, Nayla Herran.

This collaboration with CEA-CESTA aimed at evaluating the limits of validity of the MIRÓ software provided by CEA. The MIRÓ software implements the propagation of laser pulses with a non-linear Schrodinger-like equation obtained from Maxwell’s equations in non-linear dispersive medium, assuming that the pulse spectrum is narrow and the non-linearity small enough. When considering intense ultra-short pulses, the spectrum bandwidth and the amplitude of the pulse may violate the constitutive assumptions of the model used by MIRÓ. This collaboration began with the internship of Nayla Herran (march 2013- august 2013). During this internship, different alternative models have been explored, and some of them were able to provide a solution much more accurate than MIRÓ’s models. The comparisons between alternative models and MIRÓ’s model have been performed in 1-D on small cases, for which the reference solution could be obtained from a direct numerical simulation of non-linear Maxwell’s equations. An efficient and accurate solver for non-linear Maxwell’s equations has been implemented in Montjoie software. This solver uses high order finite element in space, and high order Runge-Kutta-Nystrom scheme in time. The space grid moves with respect to the group velocity such that the computational domain stays relatively small and centered around the pulse. Thanks to this solver, we have been able to validate the different models and compare them. These 1-D promising results encourage us to continue this collaboration in order to obtain efficient and accurate numerical methods in 3-D. Our desire is also to be able to perform realistic cases involving physically relevant phenomena.

6.2.13. Asymptotic Modeling for Elasto-Acoustics

Participants: Julien Diaz, Victor Péron.

In the papers [31], [74], we derive equivalent conditions and asymptotic models for the diffraction problem of elastic and acoustic waves in a solid medium surrounded by a thin layer of fluid medium. Due to the thinness of the layer with respect to the wavelength, this problem is well suited for the notion of equivalent conditions and the effect of the fluid medium on the solid is as a first approximation local. We derive and validate equivalent conditions up to the fourth order for the elastic displacement. These conditions approximate the acoustic waves which propagate in the fluid region. This approach leads to solve only elastic equations. The construction of equivalent conditions is based on a multiscale expansion in power series of the thickness of the layer for the solution of the transmission problem.

Questions regarding the implementation of the conditions have been addressed carefully and the boundary conditions have been integrated without changing the structure of the code Hou10ni.

This work has been presented in two international conferences and Workshops : Workshop HPC-GA and WAVES'2013 [55].

A paper with numerical results for the elasto-acoustic problem with several configurations (a thin layer of variable thickness; coupling with an exterior "acoustic" medium) is in preparation.

6.2.14. Thin layer models for electromagnetics

Participants: Marc Duruflé, Victor Péron, Clair Poignard.

We have considered the transmission of electromagnetic waves through a thin layer. This thin layer can be replaced by transmission conditions. Media with thin inclusions appear in many domains: geophysical applications, microwave imaging, biomedical applications, cell phone radiations, radar applications, non-destructive testing. Our application concerns also media for which the conductivity drops inside the layer, such that the low-frequency regime has a different behavior from the mid-frequency regime. Different models are compared for these two regimes, drawbacks and disadvantages of each model are detailed. This work has been accepted for publication [27].

6.2.15. Corner Asymptotics of the Magnetic Potential in the Eddy-Current Model

Participants: Monique Dauge, Patrick Dular, Laurent Krähenbühl, Victor Péron, Ronan Perrussel, Clair Poignard.

The following results rely on a problematic developed in section 3.2 , item **Asymptotic modeling**.

In the paper [25], we describe the magnetic potential in the vicinity of a corner of a conducting body embedded in a dielectric medium in a bidimensional setting. We make explicit the corner asymptotic expansion for this potential as the distance to the corner goes to zero. This expansion involves singular functions and singular coefficients. We introduce a method for the calculation of the singular functions near the corner and we provide two methods to compute the singular coefficients: the method of moments and the method of quasi-dual singular functions. Estimates for the convergence of both approximate methods are proven. We eventually illustrate the theoretical results with finite element computations. The specific non-standard feature of this problem lies in the structure of its singular functions: They have the form of series whose first terms are harmonic polynomials and further terms are genuine non-smooth functions generated by the piecewise constant zeroth order term of the operator. This work has been presented in the international conference JSA 2013 [38].

6.2.16. Finite Element Subproblem Method

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

In the paper [26], we develop a finite element subproblem method to correct the inaccuracies proper to perfect conductor and impedance boundary condition models, in particular near edges and corners of conductors, for a large range of conductivities and frequencies. Successive local corrections, supported by fine local meshes, can be obtained from each model to a more accurate one, allowing efficient extensions of their domains of validity. This work has been presented in the international conference Compumag 2013 [57].

We develop also a finite element subproblem method for progressive eddy current modeling. The modeling of eddy currents in conductors is split into a sequence of progressive finite element subproblems. The source fields generated by the inductors alone are calculated at first via either the Biot-Savart law or finite elements. The associated reaction fields for each added conductive region, and in return for the source regions themselves when massive, are then calculated with finite element models, possibly with initial perfect conductor and/or impedance boundary conditions to be further corrected. The resulting subproblem method allows efficient solving of parameterized analyses thanks to a proper mesh for each subproblem and the reuse of previous solutions to be locally corrected. This work has been presented in the international conference ISEF'2013 [56].

6.3. High Performance methods for solving wave equations

6.3.1. Coupling the DG code with task programming libraries

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

Last year we optimized the DG code implemented in the DIVA platform of Total by reducing the number of communications between each processors. This optimization, coupled with the use of Hybrid MPI and OpenMP parallel programming has allowed to prove the scalability of the code up to 512 cores. We are now planning to extend these tests up to 4000 cores. However, preliminary results emphasized the limitations due to low level issues such as threads placement, data communications and cache utilization. Therefore, we are now considering the implementation in DIVA of task programming libraries such as StarPU (http://runtime.bordeaux.inria.fr/StarPU/ or DAGuE (http://icl.cs.utk.edu/dague/index.html). These libraries handle the low level issues directly at the runtime level and allow the programmer to focus on the algorithm itself. They are also provide a valuable help to improve the portability of the code from one architecture to another, which will allow us to port DIVA on heterogeneous architectures such as CPU/GPU and Intel Xeon Phi. We have already coupled DIVA and DAGuE on the Symmetric Multiprocessor System (SMP) of Plafrim (https://plafrim.bordeaux.inria.fr) and compared the performance of the code with MPI and an OpenMP implementations [66], [64].

MOISE Project-Team

6. New Results

6.1. Mathematical Modelling of the Ocean Dynamics

6.1.1. Coupling Methods for Oceanic and Atmospheric Models

Participants: Eric Blayo, Mehdi-Pierre Daou, Laurent Debreu, Florian Lemarié, Antoine Rousseau, Manel Tayachi.

6.1.1.1. Interface conditions for coupling ocean models

Many physical situations require coupling two models with not only different resolutions, but also different physics. Such a coupling can be studied within the framework of global-in-time Schwarz methods. However, the efficiency of these iterative algorithms is strongly dependent on interface conditions. As a first step towards coupling a regional scale primitive equations ocean model with a local Navier-Stokes model, a study on the derivation of interface conditions for 2-D $x - z$ Navier-Stokes equations has been performed in D. Cherel PhD thesis. It has been shown theoretically that several usual conditions lead to divergent algorithms, and that a convergent algorithm is obtained when using transmission conditions given by a variational calculation.

D. Cherel has implemented a Schwarz-based domain decomposition method, for which he developed optimized absorbing boundary conditions that mix the velocity and pressure variables on an Arakawa-C grid. The numerical results confirm the rate of convergence that has been obtained theoretically, thanks to a Fourier analysis of the semi-discretized problem. New optimized conditions offer much better convergence than classical Dirichlet-Dirichlet conditions, even when domains overlap. A paper is now ready for submission.

6.1.1.2. Coupling dimensionally heterogeneous models

The coupling of different types of models is gaining more and more attention recently. This is due, in particular, to the needs of more global models encompassing different disciplines (*e.g.* multi-physics) and different approaches (*e.g.* multi-scale, nesting). Also, the possibility to assemble different modeling units inside a friendly modelling software platform is an attractive solution compared to developing more and more global complex models. More specifically one may want to couple 1D to 2D or 3D models, such as Shallow Water and Navier Stokes models: this was the framework of our partnership with EDF in the project MECSICO, now extended with ARTELIA Group. In her PhD (defended Oct. 28th, 2013) [4], M. Tayachi built a theoretical and numerical framework to couple 1D, 2D and 3D models for river flows.

In [103] (now accepted for publication), we propose and analyze an efficient iterative coupling method for a dimensionally heterogeneous problem. We consider the case of a 2-D Laplace equation with non symmetric boundary conditions with a corresponding 1-D Laplace equation. We first show how to obtain the 1-D model from the 2-D one by integration along one direction, by analogy with the link between shallow water equations and the Navier-Stokes system. Then we focus on the design of a Schwarz-like iterative coupling method. We discuss the choice of boundary conditions at coupling interfaces. We prove the convergence of such algorithms and give some theoretical results related to the choice of the location of the coupling interface, and to the control of the difference between a global 2-D reference solution and the 2-D coupled one. These theoretical results are illustrated numerically. The extension of this work to shallow water equations has been started in 2013 with the PhD thesis of Medhi Pierre Daou (funded by ARTELIA). An extension to primitive equations is envisaged: a post-doc position has been proposed in 2013 (not funded) and will renewed in 2014.

6.1.1.3. Ocean-Atmosphere coupling

Coupling methods routinely used in regional and global climate models do not provide the exact solution to the ocean-atmosphere problem, but an approached one [75]. This finding has motivated a deep numerical analysis of multi-physics coupling problems, first on simplified academic cases based on diffusion equations. In this context, Schwarz-like iterative domain decomposition methods have been analyzed and efficient interface conditions have been determined to optimize the convergence rate of the method [19], [20], [53], [79]. This method has then been applied to the coupling of realistic oceanic and atmospheric models to simulate the propagation of a tropical cyclone (cyclone erica, Fig. 1). Sensitivity tests to the coupling method have been been carried out in an ensemblist approach. We showed that with a mathematically consistent coupling, compared to coupling methods en vogue in existing coupled models, the spread of the ensemble is reduced, thus indicating a much reduced uncertainty in the physical solution [24], [75].

The next step is now to complete the theoretical work done on a diffusion problem by including the formulation of physical parameterizations to tackle a problem more representative of the realistic models: a PhD thesis should start on this subject in fall 2014. In parallel, an important perspective is to assess the impact of our work on IPCC-like climate models, this task will be initiated in 2014 through a collaboration between the MOISE project-team and the LSCE (Laboratoire des Sciences du Climat et de l'Environnement).

In collaboration with geophysicists, number of studies to investigate important small-scales air-sea feedbacks are in progress [45]. Through those studies, the aim will be to mathematically derive a metamodel able to represent important processes in the marine atmospheric boundary layers. The medium term objective will be to use this metamodel to force high-resolution oceanic operational models for which the use of a full atmospheric model is not possible due to a prohibitive computational cost.

6.1.2. Numerical schemes for ocean modelling

Participants: Laurent Debreu, Jérémie Demange, Florian Lemarié.

Reducing the traditional errors in terrain-following vertical coordinate ocean models (or sigma models) has been a focus of interest for the last two decades. The objective is to use this class of model in regional domains which include not only the continental shelf, but the slope and deep ocean as well. Two general types of error have been identified: 1) the pressure-gradient error and 2) spurious diapycnal diffusion associated with steepness of the vertical coordinate. In ([92],[46]) we have studied the problem of diapycnal mixing. The solution to this problem requires a specifically designed advection scheme. We propose and validate a new scheme, where diffusion is split from advection and is represented by a rotated biharmonic diffusion scheme with flow-dependent hyperdiffusivity. The main numerical development was to render the biharmonic diffusion operator scheme unconditionally stable. This is particularly needed when the slopes between coordinates lines and isopycnal surfaces are important so that the rotation of the biharmonic leads to strong stability condition along the vertical coordinate where the grid size is relatively small. This work also extends more classical results on the stability of Laplacian diffusion with mixed derivatives.

In his PhD, Jérémie Demange works on advection-diffusion schemes for ocean models (Supervisors : L. Debreu, P. Marchesiello (IRD)). His work focuses on the link between tracers (temperature and salinity) and momentum advection and diffusion in the non hyperbolic system of equations typically used in ocean models (the so called primitive equations with hydrostatic and Boussinesq assumptions). We also investigated the use of a depth dependent barotropic mode in free surface ocean models. When most ocean models assume that this mode is vertically constant, we have shown that the use of the true barotropic mode, derived from a normal mode decomposition, allows more stability and accuracy in the representation of external gravity waves.

Salinity at 1000 m in the Southwest Pacific ocean is shown in figure 2 . The use of traditional upwind biased schemes (middle) exhibits a strong drift in the salinity field in comparison with climatology (left). The introduction of high order diffusion rotated along geopotential surfaces prevents this drift while maintaining high resolution features (right).

../../../../projets/moise/IMG/wrf_roms_domain_reduced2.png

Figure 1. Snapshots (March 12, 2003 at 8 p.m. GMT) of (a) oceanic sea surface temperature (b) atmospheric 10 meter winds during a coupled simulation.

../../../../projets/moise/IMG/diffusion.png

Figure 2. Salinity at 1000m in the Southwest Pacific ocean.

6.2. Data Assimilation for Geophysical Models

6.2.1. Development of a Variational Data Assimilation System for OPA9/NEMO

Participants: Arthur Vidard, Bénédicte Lemieux-Dudon, Pierre-Antoine Bouttier, Eric Blayo.

We are heavily involved in the development of NEMOVAR (Variational assimilation for NEMO). For several years now, we built a working group (coordinated by A. Vidard) in order to bring together various NEMOVAR user-groups with diverse scientific interests (ranging from singular vector and sensitivity studies to specific issues in variational assimilation). It has led to the creation of the VODA (Variational Ocean Data Assimilation for multi scales applications) ANR project (ended in 2012). A new project, part of a larger EU-FP7 project (ERA-CLIM2) has been submitted late 2012 and will start early 2014.
The project aims at delivering a common NEMOVAR platform based on NEMO platform for 3D and 4D variational assimilation. Following 2009-11 VODA activities, a fully parallel version of NEMOTAM (Tangent and Adjoint Model for NEMO) is now available for the community in the standard NEMO version. This version is based on the released 3.4.1 version of NEMO.

We are also investigating variational data assimilation methods applied to high resolution ocean numerical models (see figure 3). This part of the project is now well advanced and encouraging preliminary results are available on an idealised numerical configuration of an oceanic basin. Several novative diagnostics have been also developed in this framework as part of P.A. Bouttier's PhD that will be defended early 2014

Lastly, multi resolution algorithms have been developed to solve the variational problem, and preliminary results were presented in two international communications [52], [51].

6.2.2. Ensemble Kalman filtering for large scale ice-sheet models

Participants: Bertrand Bonan, Maëlle Nodet, Catherine Ritz.

In collaboration with C. Ritz (CNRS, Laboratoire de Glaciologie et Geophysique de l'Environnement (LGGE), Grenoble), we aim to develop inverse methods for ice cap models.

In the framework of global warming, the evolution of sea level is a major but ill-known phenomenon. It is difficult to validate the models which are used to predict the sea level elevation, because observations are heterogeneous and sparse.

Data acquisition in polar glaciology is difficult and expensive. Satellite data have a good spatial coverage, but they allow only indirect observation of the interesting data. Moreover, ice dynamics processes are highly non linear and involve many feedback loops, so that classical linear data assimilation give poor results.

B. Bonan defended his PhD [1] in November 2013 on this subject. We implemented the Ensemble Transform Kalman Filter (ETKF) algorithm for a flowline Shallow-Ice model, called Winnie, developed by C. Ritz at LGGE. On twin experiments we got interesting results. Figures 4 show the reconstruction of the bedrock topography for various ensemble sizes. We can see that the obtained bedrock is very close to the true one, even for small ensemble sizes. This is very promising for the future, as we want to implement this method into a full 3D model. A journal paper has been submitted on this subject, and the results have been presented at many conferences [27], [37], [48], [38], [39].

6.2.3. Inverse methods for full-Stokes glaciology models

Participants: Olivier Gagliardini, Maëlle Nodet, Catherine Ritz.

We are investigating the means to apply inverse modeling to another class of glaciology models, called full-Stokes model. Such a model is developed by LGGE and CSC in Finland, called Elmer/Ice. Contrary to large scale models, Elmer/Ice is based on the full Stokes equations, and no assumptions regarding aspect ratio are made, so that this model is well adapted to high resolution small scale modelling, such as glaciers (and more recently the whole Greenland ice-sheet).

../../../../projets/moise/IMG/seabass.jpg

Figure 3. Snapshot of the relative vorticity field (1/s) for an academic oceanic basin model at 1/100° horizontal resolution.

Figure 4.

Left: Bedrock topography after 20 years of the LETKF with inflation. The background (green) is compared to reference (blue) and the analyses for various ensemble sizes: 100 members (purple), 50 members (cyan) and 30 members (orange).
Right: Standard deviation of the errors (compared to the reference) for the bedrock topography after 20 years of the LETKF with inflation. The background (green) is compared to the analyses for various ensemble sizes: 100 members (purple), 50 members (cyan) and 30 members (orange).

In collaboration with O. Gagliardini, F. Gillet-Chaulet and C. Ritz (Laboratoire de Glaciologie et Géophysique de l'Environnement (LGGE), Grenoble), we investigated a new method to solve inverse problems for a Full-Stokes model of Groenland, which consisted in solving iteratively a sequence of Neumann and Dirichlet problems within a gradient descent algorithm. We also compared this method to an approximate variational algorithm, using the fact that the full Stokes equations are almost self-adjoint. These results have been presented at an international conference [44].

6.3. Development of New Methods for Data Assimilation

6.3.1. Variational Data Assimilation with Control of Model Error

Participants: Bénédicte Lemieux-Dudon, Arthur Vidard.

One of the main limitations of the current operational variational data assimilation techniques is that they assume the model to be perfect, mainly because of computing cost issues. Numerous researches have been carried out to reduce the cost of controlling model errors by controlling the correction term only in certain privileged directions or by controlling only the systematic and time correlated part of the error.

Both the above methods consider the model errors as a forcing term in the model equations. Trémolet (2006) describes another approach where the full state vector (4D field: 3D spatial + time) is controlled. Because of computing cost one cannot obviously control the model state at each time step. Therefore, the assimilation window is split into sub-windows, and only the initial conditions of each sub-window are controlled, the junctions between each sub-window being penalized. One interesting property is that, in this case, the computation of the gradients, for the different sub-windows, are independent and therefore can be done in parallel.

This method is now implemented in a realistic oceanic framework using OPAVAR/ NEMOVAR. An extensive documentation has been produced and we are now assessing the improvement over the previous scheme

6.3.2. Direct assimilation of sequences of images

Participants: François-Xavier Le dimet, Maëlle Nodet, Nicolas Papadakis, Arthur Vidard, Vincent Chabot.

At the present time the observation of Earth from space is done by more than thirty satellites. These platforms provide two kinds of observational information:

- Eulerian information as radiance measurements: the radiative properties of the earth and its fluid envelops. These data can be plugged into numerical models by solving some inverse problems.
- Lagrangian information: the movement of fronts and vortices give information on the dynamics of the fluid. Presently this information is scarcely used in meteorology by following small cumulus clouds and using them as Lagrangian tracers, but the selection of these clouds must be done by hand and the altitude of the selected clouds must be known. This is done by using the temperature of the top of the cloud.

MOISE was the leader of the ANR ADDISA project dedicated to the assimilation of images, and is a member of its follow-up GeoFluids (along with EPI FLUMINANCE and CLIME, and LMD, IFREMER and Météo-France) that just ended in 2013.

During the ADDISA project we developed Direct Image Sequences Assimilation (DISA) and proposed a new scheme for the regularization of optical flow problems [101]. Thanks to the nonlinear brightness assumption, we proposed an algorithm to estimate the motion between two images, based on the minimization of a nonlinear cost function. We proved its efficiency and robustness on simulated and experimental geophysical flows [77]. As part of the ANR project GeoFluids, we are investigating new ways to define distance between a couple of images. One idea is to compare the gradient of the images rather than the actual value of the pixels. This leads to promising results. Another idea, currently under investigation, consists in comparing main structures within each image. This can be done using, for example, a wavelet representation of images. Both approaches have been compared, in particular their relative merits in dealing with observation errors, in a submitted paper late 2013 [63] and presented in a national conference [34]

We are also part of TOMMI, another ANR project started mid 2011, where we are investigating the possibility to use optimal transportation based distances for images assimilation.

6.3.3. Image processing, Optimal transport

Participants: Maëlle Nodet, Nicolas Papadakis, Arthur Vidard, Nelson Feyeux.

Within the optimal transport project TOMMI funded by the ANR white program, a new optimization scheme based on proximal splitting method has been proposed to solve the dynamic optimal transport problem. This work allows the computation of generalized optimal transports and will be published in SIAM Journal on Imaging Sciences [96]. We investigate also the use of optimal transport based distances for data assimilation. N. Feyeux just started his PhD on this subject, and his PhD project has been presented in a regional workshop [49].

6.3.4. A Nudging-Based Data Assimilation Method: the Back and Forth Nudging

Participants: Maëlle Nodet, Jacques Blum, Didier Auroux.

The Back and Forth Nudging (BFN) algorithm has been recently introduced for simplicity reasons, as it does not require any linearization, nor adjoint equation, or minimization process in comparison with variational schemes. Nevertheless it provides a new estimation of the initial condition at each iteration.

Previous theoretical results showed that BFN was often ill-posed for viscous partial differential equations. To overcome this problem, we proposed a new version of the algorithm, which we called the Diffusive BFN, and which showed very promising results on one-dimensional viscous equations. Experiments on more sophisticated geophysical models, such as Shallow-Water equations and NEMO ocean model are still in progress, in collaboration with University of Nice, and have been presented at the MAMERN conference [30].

6.3.5. Multigrid methods for Variational Data Assimilation.

Participants: Laurent Debreu, François-Xavier Le Dimet, Arthur Vidard.

In order to lower the computational cost of the variational data assimilation process, we investigate the use of multigrid methods to solve the associated optimal control system. On a linear advection equation, we study the impact of the regularization term on the optimal control and the impact of discretization errors on the efficiency of the coarse grid correction step. We show that even if the optimal control problem leads to the solution of an elliptic system, numerical errors introduced by the discretization can alter the success of the multigrid methods. The view of the multigrid iteration as a preconditioner for a Krylov optimization method leads to a more robust algorithm. A scale dependent weighting of the multigrid preconditioner and the usual background error covariance matrix based preconditioner is proposed and brings significant improvements. This work is presented in a paper submitted to QJRMS ([68]).

6.3.6. Variational Data Assimilation and Control of Boundary Conditions

Participant: Eugène Kazantsev.

A variational data assimilation technique is applied to the identification of the optimal boundary conditions for two configurations of the NEMO model.

The first one is a full-physics low-resolution configuration, known as ORCA-2 model. In this experiment we identify optimal parametrizations of boundary conditions on the lateral boundaries as well as on the bottom and on the surface of the ocean [17]. The influence of boundary conditions on the solution is analyzed as in the assimilation window and beyond the window. It is shown that the influence of the lateral boundaries is not significant in this configuration, while optimal surface and bottom boundary conditions allow us to better represent the jet streams, such as Gulf Stream and Kuroshio. Analyzing the reasons of the jets reinforcement, we notice that data assimilation has a major impact on parametrization of the bottom boundary conditions for u and v [23].

The second configuration of the Nemo model is devoted to the identification of the optimal parametrization of lateral boundary conditions. The model in a rectangular box placed in mid-latitudes and subjected to the classical single or double gyre wind forcing is studied. The model grid can be rotated on a desired angle around the center of the rectangle in order to simulate the boundary approximated by a staircase-like coastlines. The solution of the model on the grid aligned with the box borders was used as a reference solution and as artificial observational data. It is shown that optimal boundary has a rather complicated geometry which is neither a staircase, nor a strait line. The boundary conditions found in the data assimilation procedure brings the solution toward the reference solution allowing to correct the influence of the rotated grid (see fig. 5).

Adjoint models, necessary to variational data assimilation, have been produced by the TAPENADE software, developed by the TROPICS team. This software is shown to be able to produce the adjoint code, that can be used in data assimilation after a memory usage optimization.

Figure 5. Sea surface elevation: reference solution on the aligned grid (left), solutions on the $30°$ *rotated grid with optimal (center) and classical (right) boundary conditions.*

6.4. Quantifying Uncertainty

6.4.1. Sensitivity analysis for forecasting ocean models

Participants: Anestis Antoniadis, Eric Blayo, Gaëlle Chastaing, Céline Helbert, Alexandre Janon, François-Xavier Le Dimet, Simon Nanty, Maëlle Nodet, Clémentine Prieur, Federico Zertuche, Simon Nanty, Laurent Gilquin.

6.4.1.1. Scientific context

Forecasting ocean systems require complex models, which sometimes need to be coupled, and which make use of data assimilation. The objective of this project is, for a given output of such a system, to identify the most influential parameters, and to evaluate the effect of uncertainty in input parameters on model output. Existing stochastic tools are not well suited for high dimension problems (in particular time-dependent problems), while deterministic tools are fully applicable but only provide limited information. So the challenge is to gather expertise on one hand on numerical approximation and control of Partial Differential Equations, and on the other hand on stochastic methods for sensitivity analysis, in order to develop and design innovative stochastic solutions to study high dimension models and to propose new hybrid approaches combining the stochastic and deterministic methods.

6.4.1.2. Estimating sensitivity indices

A first task is to develop tools for estimated sensitivity indices. In variance-based sensitivity analysis, a classical tool is the method of Sobol' [100] which allows to compute Sobol' indices using Monte Carlo integration. One of the main drawbacks of this approach is that the estimation of Sobol' indices requires the use of several samples. For example, in a d-dimensional space, the estimation of all the first-order Sobol' indices requires $d+1$ samples. Some interesting combinatorial results have been introduced to weaken this defect, in particular by Saltelli [98] and more recently by Owen [95] but the quantities they estimate still require $O(d)$ samples. In a recent work [104] we introduce a new approach to estimate for any k all the k-th order Sobol' indices by using only two samples based on replicated latin hypercubes. We establish theoretical properties of such a method for the first-order Sobol' indices and discuss the generalization to higher-order indices. As an illustration, we propose to apply this new approach to a marine ecosystem model of the Ligurian sea (northwestern Mediterranean) in order to study the relative importance of its several parameters. The calibration process of this kind of chemical simulators is well-known to be quite intricate, and a rigorous and robust — i.e. valid without strong regularity assumptions — sensitivity analysis, as the method of Sobol' provides, could be of great help. The computations are performed by using CIGRI, the middleware used on the grid of the Grenoble University High Performance Computing (HPC) center. We are also applying these estimates to calibrate integrated land use transport models. It is the first step in the PhD of Laurent Gilquin (started in October 2013). Laurent Gilquin is supervised by Clémentine Prieur and Elise Arnaud (EPI STEEP) and his PhD is funded by the ANR project CITIES.

We can now wonder what are the asymptotic properties of these new estimators, or also of more classical ones. In [67], the authors deal with asymptotic properties of the estimators. In [70], the authors establish also a multivariate central limit theorem and non asymptotic properties.

6.4.1.3. Intrusive sensitivity analysis, reduced models

Another point developed in the team for sensitivity analysis is model reduction. To be more precise regarding model reduction, the aim is to reduce the number of unknown variables (to be computed by the model), using a well chosen basis. Instead of discretizing the model over a huge grid (with millions of points), the state vector of the model is projected on the subspace spanned by this basis (of a far lesser dimension). The choice of the basis is of course crucial and implies the success or failure of the reduced model. Various model reduction methods offer various choices of basis functions. A well-known method is called “proper orthogonal decomposition" or “principal component analysis". More recent and sophisticated methods also exist and may be studied, depending on the needs raised by the theoretical study. Model reduction is a natural way to overcome difficulties due to huge computational times due to discretizations on fine grids. In [10], the authors present a reduced basis offline/online procedure for viscous Burgers initial boundary value problem, enabling efficient approximate computation of the solutions of this equation for parametrized viscosity and initial and boundary value data. This procedure comes with a fast-evaluated rigorous error bound certifying the approximation procedure. The numerical experiments in the paper show significant computational savings, as well as efficiency of the error bound. When a metamodel is used (for example reduced basis metamodel, but also kriging, regression, ...) for estimating sensitivity indices by Monte Carlo type estimation, a twofold error appears : a sampling error and a metamodel error. Deriving confidence intervals taking into account these two sources of uncertainties is of great interest. We obtained results particularly well fitted for reduced basis metamodels [90]. We are now considering problems related to more general models such as Shallow-Water models. In [15], the authors provide asymptotic confidence intervals in the double limit where the sample size goes to infinity and the metamodel converges to the true model.

Let us come back to the output of interest. Is it possible to get better error certification when the output is specified. A work in this sense has been submitted, dealing with goal oriented uncertainties assessment [89].

6.4.1.4. Sensitivity analysis with dependent inputs

An important challenge for stochastic sensitivity analysis is to develop methodologies which work for dependent inputs. For the moment, there does not exist conclusive results in that direction. Our aim is to define an analogue of Hoeffding decomposition [88] in the case where input parameters are correlated. Clémentine

Prieur supervised Gaëlle Chastaing's PhD thesis on the topic (defended in September 2013) [2]. We obtained first results [81], deriving a general functional ANOVA for dependent inputs, allowing defining new variance based sensitivity indices for correlated inputs. We then adapted various algorithms for the estimation of these new indices. These algorithms make the assumption that among the potential interactions, only few are significant. Two papers have been submitted [64], [66].

Céline Helbert and Clémentine Prieur supervise the PhD thesis of Simon Nanty (funded by CEA Cadarache). The subject of the thesis is the analysis of uncertainties for numerical codes with temporal and spatio-temporal input variables, with application to safety and impact calculation studies. This study implies functional dependent inputs. A first step is the modeling of these inputs.

6.4.1.5. Multy-fidelity modeling for risk analysis

Federico Zertuche's PhD concerns the modeling and prediction of a digital output from a computer code when multiple levels of fidelity of the code are available. A low-fidelity output can be obtained, for example on a coarse mesh. It is cheaper, but also much less accurate than a high-fidelity output obtained on a fine mesh. In this context, we propose new approaches to relieve some restrictive assumptions of existing methods ([91], [97]) : a new estimating method of the classical cokriging model when designs are not nested and a nonparametric modeling of the relationship between low-fidelity and high-fidelity levels. The PhD takes place in the REDICE consortium and in close link with industry. The first year was also dedicated to the development of a case study in fluid mechanics with CEA in the context of the study of a nuclear reactor.

The second year of the thesis was dedicated to the development of a new sequential approach based on a course to fine wavelets algorithm.

6.4.2. Evaluation of a posteriori covariance errors

In the context of data assimilation, taking into account the a priori covariance error on the prediction and on the observations, the model and the observations, an analysis can be obtained followed by a prediction. This one makes sense only if an estimation of the error can be provided. The tendency is to use "ensemble methods" i.e. to realize a large number of predictions and estimate statistics on the results. This method raises two problems: the high computational cost and the weak theoretical justification. We have proposed a new method based on the fact that in the linear case the covariance is the inverse of the Hessian. The principle of our method is to add a correcting term to the Hessian in the non linear case. This work has been published in 2013 [14]. This paper has also be presented at the 6th WMO Symposium on Data Assimilation held in College Park, MD, USA in October 2013 [73].

6.4.3. Second Order Information in Variational Data Assimilation

This theme is centered around sensitivity analysis with respect to the observations. The link between data and models is made only in the Optimality System. Therefore a sensitivity analysis on the observations must be carried out on the Optimality System thus using second order information. This research is done in cooperation with Victor Shutyaev (Institute of Numerical Mathematics, Moscow), Tran Thu Ha (Institute of Mechanics, Ha Noi, Vietnam). One paper is published in the Russ. J. Of Numerical Analysis [18]. Another application to identification of parameters in a hydrological model is submitted [105].

6.5. Tracking of mesoscale convective systems

Participants: Clémentine Prieur, Alexandros Makris.

6.5.1. Scientific context

We are interested in the tracking of mesoscale convective systems.A particular region of interest is West Africa. Data and hydrological expertise is provided by T. Vischel and T. Lebel (LTHE, Grenoble).

6.5.2. Results

A first approach involves adapting the multiple hypothesis tracking (MHT) model originally designed by the NCAR (National Centre for Atmospheric Research) for tracking storms [102] to the data for West Africa. With A. Makris (working on a post-doctoral position), we proposed a Bayesian approach [76], which consists in considering that the state at time t is composed on one hand by the events (birth, death, splitting, merging) and on the other hand by the targets' attributes (positions, velocities, sizes, ...). The model decomposes the state into two sub-states: the events and the targets positions/attributes. The events are updated first and are conditioned to the previous targets sub-state. Then given the new events the target substate is updated. A simulation study allowed to verify that this approach improves the frequentist approach by Storlie et al. (2009). It has been tested on simulations and must now be investigated in the specific context of real data on West Africa. Using PHD (probability hypothesis density) filters adapted to our problem, generalising recent developments in particle filtering for spatio-temporal branching processes (e.g. [80]) could be an interesting alternative to explore. The idea of a dynamic, stochastic tracking model should then provide the base for generating rainfall scenarios over a relatively vast area of West Africa in order to identify the main sources of variability in the monsoon phenomenon.

6.6. Multivariate risk indicators

Participants: Clémentine Prieur, Patricia Tencaliec.

6.6.1. Scientific context

Studying risks in a spatio-temporal context is a very broad field of research and one that lies at the heart of current concerns at a number of levels (hydrological risk, nuclear risk, financial risk etc.). Stochastic tools for risk analysis must be able to provide a means of determining both the intensity and probability of occurrence of damaging events such as e.g. extreme floods, earthquakes or avalanches. It is important to be able to develop effective methodologies to prevent natural hazards, including e.g. the construction of barrages.

6.6.2. Results

Different risk measures have been proposed in the one-dimensional framework . The most classical ones are the return level (equivalent to the Value at Risk in finance), or the mean excess function (equivalent to the Conditional Tail Expectation CTE). However, most of time there are multiple risk factors, whose dependence structure has to be taken into account when designing suitable risk estimators. Relatively recent regulation (such as Basel II for banks or Solvency II for insurance) has been a strong driver for the development of realistic spatio-temporal dependence models, as well as for the development of multivariate risk measurements that effectively account for these dependencies. We refer to [82] for a review of recent extensions of the notion of return level to the multivariate framework. In the context of environmental risk, [99] proposed a generalization of the concept of return period in dimension greater than or equal to two. Michele et al. proposed in a recent study [83] to take into account the duration and not only the intensity of an event for designing what they call the dynamic return period. However, few studies address the issues of statistical inference in the multivariate context. In [9], [86], we proposed non parametric estimators of a multivariate extension of the CTE. As might be expected, the properties of these estimators deteriorate when considering extreme risk levels. In collaboration with Elena Di Bernardino (CNAM, Paris), Clémentine Prieur is working on the extrapolation of the above results to extreme risk levels.

Elena Di Bernardino, Véronique Maume-Deschamps (Univ. Lyon 1) and Clémentine Prieur also derived an estimator for bivariate tail [10]. The study of tail behavior is of great importance to assess risk.

With Anne-Catherine Favre (LTHE, Grenoble), Clémentine Prieur supervises the PhD thesis of Patricia Tencaliec. We are working on risk assessment, concerning flood data for the Durance drainage basin (France). The PhD thesis started in October.

6.7. Non-parametric estimation for kinetic diffusions

Participant: Clémentine Prieur.

This research is the subject of a collaboration with Venezuela (Professor Jose R. Leon, Caracas Central University) and is partly funded by an ECOS Nord project.

We are focusing our attention on models derived from the linear Fokker-Planck equation. From a probabilistic viewpoint, these models have received particular attention in recent years, since they are a basic example for hypercoercivity. In fact, even though completely degenerated, these models are hypoelliptic and still verify some properties of coercivitiy, in a broad sense of the word. Such models often appear in the fields of mechanics, finance and even biology. For such models we believe it appropriate to build statistical non-parametric estimation tools. Initial results have been obtained for the estimation of invariant density, in conditions guaranteeing its existence and unicity [8] and when only partial observational data are available. A paper on the non parametric estimation of the drift has been submitted recently [62] (see Samson et al., 2012, for results for parametric models). As far as the estimation of the diffusion term is concerned, we obtained promising results, in collaboration with J.R. León (Caracas, Venezuela) and P. Cattiaux (Toulouse). These results should be submitted shortly.

6.8. CO_2 Storage

Participant: Céline Helbert.

In collaboration with Bernard Guy (EMSE, Saint-Etienne) and more specifically in the context the PhD of Joharivola Raveloson (EMSE, Saint-Etienne), we are interested in the study of the water-rock interactions in the case of CO_2 storage in geological environment. This work is following the study of Franck Diedro in the same subject [87]. In this study we focus on the scale of observation of geochemical phenomena while taking into account the heterogeneity of the reservoir. This heterogeneity at small and large scale helps to maintain a local variability of the chemical composition and influence reaction rates at the pore as well at the reservoir scale. To connect the parameters at both scale (pore and reservoir) we use deterministic and stochastic simulations of a reactive transport code developed by IFPEN.

6.9. Land Use and Transport models calibration

Participants: Thomas Capelle, Laurent Gilquin, Clémentine Prieur, Nicolas Papadakis, Arthur Vidard.

Given the complexity of modern urban areas, designing sustainable policies calls for more than sheer expert knowledge. This is especially true of transport or land use policies, because of the strong interplay between the land use and the transportation systems. Land use and transport integrated (LUTI) modelling offers invaluable analysis tools for planners working on transportation and urban projects. Yet, very few local authorities in charge of planning make use of these strategic models. The explanation lies first in the difficulty to calibrate these models, second in the lack of confidence in their results, which itself stems from the absence of any well-defined validation procedure. Our expertise in such matters will probably be valuable for improving the reliability of these models. To that purpose we participated to the building up of the ANR project CITiES lead by the STEEP EPI. This project has just started early 2013 and Two PhD about sensitivity analysis and calibration were launched this fall.

6.10. Mathematical modelling for CFD-environment coupled systems

Participants: Antoine Rousseau, Maëlle Nodet.

6.10.1. Minimal-time bioremediation of natural water resources

The objective of this work is to provide efficient strategies for the bioremediation of natural water resources. The originality of the approach is to couple minimal time strategies that are determined on a simplified model with a faithful numerical model for the hydrodynamics. Based on a previous paper that deals with an implicit representation of the spatial inhomogenity of the resource with a small number of homogeneous compartments (with a system of ODEs), we implement a coupled ODE-PDE system that accounts for the spatial non-homogeneity of pollution in natural resources. The main idea is to implement a Navier-Stokes model in the resource (such as a lake), with boundary conditions that correspond to the output feedback that has been determined to be optimal for the simple ODEs model of a (small) bioreactor. A first mathematical model has been introduced and numerical simulations have been performed in academic situations. We built a reduced model that approximates the reference PDE model thanks to a set of ODEs with parameters. Numerical optimization is performed on these parameters in order to better fit the reference model. In addition, bioremediation algorithms proposed by the authors have been sent to Inria Technology Transfert Services for a patent registration. Two publications (ready for submission) will be sent as soon as the patent submission process is complete.

Finally, A. Rousseau spent 2 weeks in Santiago (April 2013) upon Inria Chile's invitation in order to work on the bioremediation of natural resources. AR and Inria Chile made a common answer to a chilean funding program (by COPEC) that was not chosen.

6.10.2. Mathematical modelling for the confinement of lagoons

This work deals with the concept of confinement of paralic ecosystems. It is based on a recent paper by E. Frénod that presents a modelling procedure in order to compute the confinement field of a lagoon.

A. Rousseau and E. Frénod improve in 2012 the existing model in order to account for tide oscillations in any kind of geometry such as a non-rectangular lagoons with a non-flat bottom. The new model, that relies on PDEs rather than ODEs, is then implemented thanks to the finite element method. Numerical results confirm the feasibility of confinement studies thanks to the introduced model. During the internship of J.-P. Bernard, we implemented the proposed method in a realistic situation, namely the Etang de Thau in Languedoc-Roussillon, France (see Figure 6). This led to two publications in 2013 [5] and [13], plus one accepted paper in 2014 [6].

`../../../../projets/moise/IMG/Thau.png`

Figure 6. Confinement map in the Thau Lagoon (France).

POMDAPI Project-Team (section vide)

SAGE Project-Team

6. New Results

6.1. Parallel numerical algorithms

6.1.1. Parallel Adaptive GMRES with deflated restarting

Participant: Jocelyne Erhel.

Grants and projects: C2S@EXA 8.2.3 , JLPC 8.4.4

Software: DGMRES, AGMRES, GPREMS.

Publications: [17], [26].

Abstract: The GMRES iterative method is widely used as a Krylov subspace technique for solving sparse linear systems when the coefficient matrix is nonsymmetric and indefinite. The Newton basis implementation has been proposed on distributed memory computers as an alternative to the classical approach with the Arnoldi process. The aim of our work here is to introduce a modification based on deflation techniques. This approach builds an augmented subspace in an adaptive way to accelerate the convergence of the restarted formulation. In our numerical experiments, we show the benefits of using this implementation with hybrid direct/iterative methods to solve large linear systems.

6.1.2. Hybrid algebraic solvers for CFD problems

Participant: Jocelyne Erhel.

Grants and projects: C2S@EXA 8.2.3 , JLPC 8.4.4

Software: DGMRES, AGMRES, GPREMS.

Publications: [18].

Abstract: Sparse linear systems arise from design optimization in computational fluid dynamics. In this approach, a linearization of the discretized compressible Navier-Stokes equations is built, in order to evaluate the sensitivity of the entire flow with respect to each design parameter. The goal is to reduce the memory requirements and indirectly, the computational cost at different steps of this scheme. Numerical results are presented with industrial test cases to show the benefits of our methodology.

6.1.3. Algebraic multilevel preconditioning

Participant: Thomas Dufaud.

Grants: C2S@EXA 8.2.3

Publications: [51], [23], [24].

Conferences: [37], [24].

Abstract: The Schwarz domain decomposition method is a very attractive numerical method for parallel computing as it needs only to update the boundary conditions on the artificial interfaces generated by domain decomposition. Thus only local communications between the neighbouring sub-domains are required. We review the use of Aitken's acceleration applied to the Schwarz domain decomposition method.

6.1.4. Counting eigenvalues in domains of the complex field

Participant: Bernard Philippe.

Grants: momappli 8.4.2

Publications: [15], [28].

Conferences: [47], [48], [22].

Abstract: A procedure for counting the number of eigenvalues of a matrix in a region surrounded by a closed curve is presented. It is based on the application of the residual theorem. The quadrature is performed by evaluating the principal argument of the logarithm of a function. A strategy is proposed for selecting a path length that insures that the same branch of the logarithm is followed during the integration. Numerical tests are reported for matrices obtained from conventional matrix test sets.

6.1.5. Sliced-time computation method

Participant: Jocelyne Erhel.

Grants: MODNUM 8.4.5

Publications: [16], [25].

Abstract: We consider the mathematical framework of a sliced-time computation method for explosive solutions to systems of ordinary differential equations. We also derive an Adaptive Parallel-in-Time Method with application to a membrane problem.

6.1.6. Interacting particles systems

Participant: Lionel Lenôtre.

Grants: H2MNO4 8.2.1

Conferences: [31]

Abstract: We consider a variance reduction method for simulations with particles.

6.2. Numerical models and simulations applied to physics

6.2.1. Heat and mass transfer modeling in porous media

Participants: Édouard Canot, Salwa Mansour.

Grants: MODNUM 8.4.5 , HYDRINV 8.4.7

Conferences: [33], [35],[44]

Abstract: The effective thermal conductivity is a key parameter for obtaining good simulations of heat transfer in wet porous media. It is very sensitive to the presence of liquid water, even in very small quantity. Moreover, during the evaporation of water, some changes of geometric configuration of the liquid meniscus lead to hysteresis behaviors. Micro-scale studies help us in understanding the global properties, via numerical simulations.

6.2.2. Heat transfer in soils applied to archaeological fires

Participants: Édouard Canot, Salwa Mansour.

Grants: MODNUM 8.4.5 , ARPHYMAT 8.4.6

Conferences: [34], [36]

Abstract: In order to be validated, the numerical simulations of heat transfer at the surface of the soil are compared to experimental results, because of the complexity of the phenomenon and the great number of physical mechanisms involved. It appears that making good experiments is hard, not to mention the limitations and lacks of the Laloy and Massard method used to obtained the effective thermal conductivity of the granular material. The Laloy and Massard method have been slightly improved; besides a different, new experimental method, based on the mathematical properties of heat transfer, has been proposed.

6.2.3. Granular materials

Participant: Édouard Canot.

Publications: [19].

Abstract: Using the $\mu(I)$ continuum model recently proposed for dense granular flows, we study theoretically steady and fully developed granular flows in two configurations: a plane shear cell and a channel made of two parallel plates (Poiseuille configuration).

6.2.4. Geodesy

Participants: Amine Abdelmoula, Bernard Philippe.

Grants: LIRIMA-EPIC 8.4.3 , joint Ph-D 8.4.9 .

Publications: [12].

Thesis: Ph-D of Amine Abdelmoula, University of Rennes 1 and Tunis, defended in December 2013.

Abstract: We solve a geodetic inverse problem for the determination of a distribution of point masses (characterized by their intensities and positions), such that the potential generated by them best approximates a given potential field.

6.3. Models and simulations for flow and transport in porous media

6.3.1. Flow and transport in highly heterogeneous porous medium

Participants: Jean-Raynald de Dreuzy, Jocelyne Erhel, Géraldine Pichot.

Grants: H2MN04 8.2.1 , H2OGuilde 8.2.4 , HEMERA 8.2.2

Software: PARADIS, H2OLab

Publications: [13]

Abstract: Models of hydrogeology must deal with both heterogeneity and lack of data. We consider a flow and transport model for an inert solute. The conductivity is a random field following a stationary log normal distribution with an exponential or Gaussian covariance function, with a very small correlation length. The quantities of interest studied here are the expectation of the spatial mean velocity, the equivalent permeability and the macro spreading. In particular, the asymptotic behavior of the plume is characterized, leading to large simulation times, consequently to large physical domains. Uncertainty is dealt with a classical Monte Carlo method, which turns out to be very efficient, thanks to the ergodicity of the conductivity field and to the very large domain. These large scale simulations are achieved by means of high performance computing algorithms and tools.

6.3.2. Diffusion processes in porous media

Participants: Lionel Lenôtre, Géraldine Pichot.

Grants: H2MN04 8.2.1

Software: SBM 5.1.7 , PALMTREE 5.2.1

Publications: [21]

Conferences: [41], [43], [42]

Abstract: We present some recent results about Monte Carlo simulations in media with interfaces. By nature, porous media are extremely heterogeneous. We consider a one-dimensional advection-diffusion equation with piecewise constant coefficients. Without drift term, the Skew Brownian Motion permits to develop several exact algorithms with constant time step. We aim at adding the drift term and dealing with higher dimensional problems.

6.3.3. Adaptive stochastic model for flow and transport with random data

Participants: Jocelyne Erhel, Mestapha Oumouni.

Grants: HYDRINV 8.4.7 , joint Ph-D 8.4.8

Publications: [27].

Conferences: [46].

Thesis:[11].

Abstract: This work presents a development and an analysis of an effective approach for partial differential equation with random coefficients and data. We are interesting in the steady flow equation with stochastic input data.

A projection method in the one-dimensional case is presented to compute efficiently the average of the solution.

An anisotropic sparse grid collocation method is also used to solve the flow problem. First, we introduce an indicator of the error satisfying an upper bound of the error, it allows us to compute the anisotropy weights of the method. We demonstrate an improvement of the error estimation of the method which confirms the efficiency of the method compared with Monte Carlo and will be used to accelerate this method by the Richardson extrapolation technique.

We also present a numerical analysis of a probabilistic method to quantify the migration of a contaminant in random media. We consider the previous flow problem coupled with the advection-diffusion equation, where we are interested in the computation of the mean extension and the mean dispersion of the solute. The flow model is discretized by a mixed finite elements method and the concentration of the solute is the density of the solution of a stochastic differential equation, which is discretized by an Euler scheme. We present an explicit formula of the dispersion and optimal a priori error estimates.

6.3.4. Reactive transport

Participants: Édouard Canot, Jocelyne Erhel, Souhila Sabit.

Grants: H2MN04 8.2.1 , ANDRA 7.1 , MOMAS 8.2.7 , C2SEXA 8.2.3

Software: GRT3D.

Publications: [52],[30].

Conferences: [20],[40],[50], [32].

Abstract: Numerical simulations are essential for studying the fate of contaminants in aquifers, for risk assessment and resources management. In this study, we deal with reactive transport models and show how a Newton method can be used efficiently. Numerical experiments illustrate the efficiency of a substitution technique. Moreover, it appears that using logarithms in the chemistry equations lead to ill conditioned matrices and increase the computational cost.

6.4. Models and simulations for flow in porous fractured media

6.4.1. Synthetic benchmark for modeling flow in 3D fractured media

Participants: Jean-Raynald de Dreuzy, Jocelyne Erhel, Géraldine Pichot.

Grants: GEOFRAC 8.2.5 , FRACINI 8.1.1

Software: MPFRAC

Publications: [14]

Abstract: Intensity and localization of flows in fractured media have promoted the development of a large range of different modeling approaches including Discrete Fracture Networks, pipe networks and equivalent continuous media. While benchmarked usually within site studies, we propose an alternative numerical benchmark based on highly-resolved Discrete Fracture Networks (DFNs) and on a stochastic approach. Test cases are built on fractures of different lengths, orientations, aspect ratios and hydraulic apertures, issuing the broad ranges of topological structures and hydraulic properties classically observed. We present 18 DFN cases, with 10 random simulations by case.

6.4.2. Robust numerical methods for solving flow in stochastic fracture networks

Participants: Jean-Raynald de Dreuzy, Jocelyne Erhel, Géraldine Pichot.

Grants: GEOFRAC 8.2.5 , FRACINI 8.1.1

Software: MPFRAC, H2OLab.

Publications: [29].

Conferences: [49].

Abstract: In this work, flow in Discrete Fracture Networks (DFN) is solved using a Mortar Mixed Hybrid Finite Element Method. To solve large linear systems derived from a nonconforming discretization of stochastic fractured networks, a Balancing Domain Decomposition is used. Tests on three stochastically generated DFN are proposed to show the ability of the iterative solver SIDNUR to solve the flow problem.

6.4.3. Flow in complex 3D geological fractured porous media

Participants: Jean-Raynald de Dreuzy, Thomas Dufaud, Jocelyne Erhel, Géraldine Pichot.

Grants: GEOFRAC 8.2.5 , FRACINI 8.1.1

Software: MPFRAC, H2OLab

Conferences: [38], [39]

Abstract: Taking into account water and solute exchanges between porous and fractured media is of great interest in geological applications. The coupled porous-fractured flow equations and their discretization by a Mixed Hybrid Finite Element Method are presented as well as the derived linear system. An appropriate mesh generation is proposed to deal with the complexity involved by randomly generated fracture networks. Numerical experiments are shown, that provide flow fields for forthcoming transport simulations.

STEEP Team

6. New Results

6.1. Data Mining for Material Flow Analysis: Application in the Territorial Breakdown of French Regions

One of the major issues for assessment of the long-term sustainability of urban areas is related to the concept of "imported sustainability". In order to produce such an assessment for a given territory, one must first identify and quantify the types of materials used, and the impacts associated to these uses. Material Flow Analysis (MFA) is directly related to how the material circulates and how it is transformed within a territory. In most cases this analysis is performed at national and regional levels, where the statistical data is available. The challenge is to establish such an analysis at smaller scales, e.g. in the case of France, at the department or city level.

We have explored the possibility of applying data analysis at the regional level by generating a mathematical model that can fit well the data at regional scale and estimate well the departmental one. The downscaling procedure relies on the assumption that the obtained model at level 'n' (for example region) will be also true at level 'n+1' (for example department), such that it could properly estimate the unknown data based on a set of chosen drivers (socio-economic data). We have designed and implemented techniques based on parameter optimization and model selection as well as robust estimation, in order to estimate the best drivers for a given set of territories, i.e. the socio-economic data (e.g. employees per type of manufacturing industry, population data, etc.) that best correlate with the production of various types of agricultural or other products [19].

6.2. Calculating indices for urban sprawl

Urban sprawl is a complex concept, that is generally associated with auto-oriented, low-density development. It is the subject of a wide range of research efforts, aiming at understanding and characterizing the underlying driving factors. We have followed up on an effort by Burchfield et al. who proposed a simple measure for urban sprawl, a so-called sprawl index. We proposed several variants of this index with the aim of achieving richer and/or more flexible characterizations of urban sprawl [16]. We also proposed ways of determining metropolitan areas that have similar patterns of urban sprawl, using clustering techniques.

6.3. Computing environmental accounts from the consumer's viewpoint using Input Output Analysis

The Russian-American economist Wassily Leontief introduced Input-Output Analysis (IOA) in the 30's and was awarded the Nobel prize in economy in 1973 for this contribution. IOA is a macro-economic tool which investigates the links and retroactions between the sectors of an economy. It makes it possible to allocate production factors (labor, capital etc.) and environmental externalities of production processes (depletion on resources, emission of pollutants etc.) to final consumption. Our first task was to reproduce the results from various studies on the carbon footprint of France. We couldn't reproduce the results from *Analyse des impacts environnementaux de la consommation des ménages et des marges de manœuvre pour réduire ces impacts* (Ademe, 2012). We underlined a mistake in the mathematical formulas presented in the annex of the paper but couldn't concur it was indeed the source of discrepancy because we were not granted access to the raw data of the study. Our results are however in line with the papers originating from the statistical service of the ministry of ecology (SOeS, J-L Pasquier).

Because working on excel sheets, although widespread in a large number of agencies, proves very inefficient, we started to work on the development of a software called Wassily (see "new results" section) that would automatize the critical calculations. We prepared databases on input-output tables and air emissions of several European countries based on Eurostat data and started to work on the architecture of the software with Julien Alapetite. In parallel, we reviewed the downscaling and regionalizing techniques in the IO literature and looked for the necessary information concerning the Rhône-Alpes region. We concluded that enough data was available in order to carry out downscaling but that data was still too scarce for the finer levels of regionalizing.

6.4. Mapping and land use and land cover change for the ESNET project

The ESNET project (EcoSystem services NETworks) is a collaboration lead by LECA (Laboratoire d'ECologie Alpine, UJF) that aims at characterizing the ecosystem services of the Grenoble urban region (about 2/3 of the Isere département) at the 2030/2040 horizon under various constraints of urban policy planning, changes in agricultural and forest management, and climate change impact on ecosystems. A preliminary task in this research program was the elaboration of very detailed maps (both in terms of land use and of resolution) of the study area at three different dates (1998, 2003 and 2009) based on available satellite and IGN data, in order to characterize past land use patterns as well as agricultural rotation patterns. These have been made and completed at Inria with the hiring of specialized engineers in these tasks, funded by the ESNET program. This exercice informs the next task (land use and land cover change – LUCC – modelling). Hosting this work at Inria was not only logical in terms of the available computer environment, but also useful in terms of visibility of Inria from outside planning agencies.

The LUCC model itself is developed partly at Inria (for modelling expertise) and partly at LECA (for expertise on ecological change drivers). The model development is still underway but in a rather advanced stage. Relevant drivers for urban development have been identified and statistically characterized. The so-called "transition potentials" (which characterize change of land use over a given period of time) are in the process of being calibrated. The next steps involve the completion of this calibration task, the development of relevant scenarios (underway by the whole ESNET collaboration) and projections of land use into the future. Some sensitivity analysis will also be performed in order to characterize the robustness of the model.

6.5. Benchmarking tools for the climate negotiation of GHG emission reduction trajectories

Climate negotiations related to global warming are another important issue of sustainable development. In this framework that is place at international scale we propose a benchmarking tool that is designed to avoid the main limitations of actual negotiation schemes. Our approach is based on the original Soft Landing proposition, made by Criqui and Kouvaritakis in the early 2000. We develop an up to date solution which improves the original idea mainly by introducing common but differentiated emission reduction profiles and by developing a dedicated algorithm for that purpose (called REDEM). To be compatible with global objectives, it is commonly accepted that for most developing regions, the national emission curves should admit a maximum and then should progressively decline. Similarly, we emphasize the fact that, in order to achieve the global objectives, all states will have to entail mitigation efforts, the intensity which may be measured by the rate of variation of the national emissions. At one point, the effort will reach a maximum, when the rate of variation in absolute value is at its maximum, and then decrease. In other words, there will also be a peak in the effort. Then we propose to base the benchmark on this peak of effort. This work has been done in collaboration with EDDEN Laboratory, in particular Patrick Criqui and Constantin Ilasca.

BIOCORE Project-Team

6. New Results

6.1. Mathematical methods and methodological approach to biology

6.1.1. Mathematical analysis of biological models

6.1.1.1. Mathematical study of semi-discrete models

Participants: Jean-Luc Gouzé, Frédéric Grognard, Ludovic Mailleret, Pierre Bernhard, Elsa Rousseau, Nicolas Bajeux.

Semi-discrete models have shown their relevance in the modeling of biological phenomena whose nature presents abrupt changes over the course of their evolution [99]. We used such models and analysed their properties in several practical situations that are developed in Section 6.2.2 , most of them requiring such a modeling in order to take seasonality into account. Such is the case when the year is divided into a cropping season and a 'winter' season, where the crop is absent, as in our analysis of the sustainable management of crop resistance to pathogens [53] or in the co-existence analysis of epidemiological strains [21]. Seasonality also plays a big role in the semi-discrete modeling required for the analysis of consumers' adaptive behavior in seasonal consumer-resource dynamics, where only dormant offspring survive the 'winter' [61].

6.1.1.2. Model design, identification and validation

Participants: Olivier Bernard, Francis Mairet.

One of the main families of biological systems that we have studied involves mass transfer between compartments, whether these compartments are microorganisms or chemical species in a bioreactor, or species populations in an ecosystem. We have developed methods to estimate the models of such systems [79]. These systems can be represented by models having the general structure popularized by [78], [84], and based on an underlying reaction network:

$$\frac{d\xi}{dt} = K\,r(\xi,\psi) + D(\xi_{in} - \xi) - Q(\xi)$$

We address two problems: the determination of the pseudo-stoichiometric matrix K and the modelling of the reaction rates $r(\xi,\psi)$.

In order to identify K, a two-step procedure has been proposed. The first step is the identification of the minimum number of reactions to be taken into account to explain a set of data. If additional information on the process structure is available, we showed how to apply the second step: the estimation of the pseudo-stoichiometric coefficients.

This approach has been applied to various bioproduction processes, among which activated sludge processes [77], anaerobic digestion [92], [106] and anaerobic digestion of microalgae [100]. Recently it was also used to reduce the ADM1 model in the case of winery effluent wastewater [24].

6.1.2. Metabolic and genomic models

Participants: Jean-Luc Gouzé, Madalena Chaves, Alfonso Carta, Ismail Belgacem, Olivier Bernard, Caroline Baroukh, Rafael Muñoz-Tamayo, Jean-Philippe Steyer.

Global stability for metabolic models and full Michaelis-Menten equations

With techniques of monotone and compartmental systems, we studied full (i.e. not reduced by any time-scale argument) Michelis-Menten reactions or chains of reactions: we prove global stability when the equilibrium exists, and show that it may not exist. This fact has important consequences for reduction of metabolic systems in a coupled genetic-metabolic system [17].

Structural principles for the existence of limit cycles in two-dimensional piecewise affine models

Using concavity and continuity properties of Poincaré maps, we have derived some structural principles which link the topology of the transition graph to the existence, number and stability of limit cycles in a class of two-dimensional piecewise affine biological models [13].

Probabilistic approach for predicting periodic orbits in piecewise affine models

In the state transition graph, a *transition probability* between two nodes can be defined in terms of the parameters of the piecewise affine models. For a cyclic transition graph, this approach can be used to predict the most likely periodic orbit for a given a set of parameters [22].

Growth rate models in bacteria: piecewise affine systems with a dilution term

We have extended the class of piecewise affine systems to deal with dynamics dependent on dilution due to cell growth rate. Considering that growth rate is determined by two limiting factors (RNA polymerase and ribosomes), in [42] we propose and analyze a switched system with two piecewise quadratic modes. This is part of the PhD thesis of Alfonso Carta, and done in collaboration with IBIS project-team.

Transcription and translation models in bacteria

We study detailed models of transcription and translation for genes in a bacterium. With techniques of monotone systems, and time scale hypotheses, we can show the stability of the fast part of these systems, and reduce them to much smaller models [40], [39]. We also study other models of the global cellular machinery. This is part of the PhD theses of Ismael Belgacem, Alfonso Carta, and done in collaboration with IBIS project-team. Moreover, in collaboration with IBIS, we studied and experimentally validated the time scale reduction of the classical two-step model for gene expression [51]

Analysis of circadian rhythms in cyanobacteria

A hierarchy of models (from Boolean to continuous) was used in [23] to successively characterize the wiring structure, qualitative dynamical properties, and then perform parameter estimation on a model describing the system responsible for the circadian rhythm of cyanobacteria.

Interconnections of Boolean modules: asymptotic and transient behaviour

The asymptotic dynamics of high-dimensional networks (e.g., genetic) can be obtained from the interconnection of two input/output Boolean subnetworks, and the analysis of their attractors. This computational cost reducing method is described in [34]. Some extensions include the characterization of the attractors of the interconnected system in terms of invariant sets.

Structure estimation for Boolean models of gene regulation networks

The problem of estimating Boolean models of gene networks from few and noisy measurements is addressed in [41], joint work with C. Breindl and F. Allgöwer from the University of Stuttgart. The class of unate or canalizing Boolean functions is considered and represented by multi-affine polynomials, leading to a reformulation of the estimation problem as a mixed integer linear program.

Analysis of dynamical systems by combining discrete and continuous formalisms

The work reviewed in the HDR of M. Chaves [11] highlights methods of analysis that use and combine techniques from discrete and piecewise affine modeling formalisms, such as construction of the transition graph and its association with the parameters of the system. Some basic methods for generating a discrete transition graph from a given continuous system are described in the internship project of F. Todoran [75].

State estimation for gene networks

We address state estimation for gene regulatory networks with intrinsic and extrinsic noise at the level of single cells. We take the Chemical Master Equation (CME) with random parameters as a reference modeling approach, and investigate the use of stochastic differential model approximations for the construction of practical real-time filters (based on non-linear Kalman filtering) [43]. This is a collaboration with Ibis team.

Modelling the metabolic network in non balanced growth conditions

On the basis of the knowledge of the metabolic network, we propose a new methodology to go beyond the “balanced growth paradigm”(assuming that there is no storage within the cell). We have therefore a tool to represent the possible storage of some key biochemical compounds. This approach was applied to describe the effect of both a light cycle and a nitrogen starvation on the lipid accumulation [37]. The first stage of the approach consists in splitting the metabolic network into sub-networks, which are assumed to satisfy balanced growth condition. The left metabolites interconnecting the sub-networks are allowed to behave dynamically. Then, thanks to Elementary Flux Mode analysis, each sub-network is reduced to macroscopic reactions, for which simple kinetics are assumed. This approach was applied to the accumulation of lipids and carbohydrates of the microalgae *Tisochrysis lutea* under day/night cycles. The resulting model described accurately experimental data obtained in day/night conditions; it efficiently predicts the accumulation and consumption of lipids and carbohydrates.

6.2. Fields of application

6.2.1. Bioenergy

6.2.1.1. Modelling of microalgae production

Participants: Olivier Bernard, Antoine Sciandra, Frédéric Grognard, Philipp Hartmann, Rafael Muñoz-Tamayo, Ghjuvan Grimaud, David Demory, Frédéric Chazalon, Hubert Bonnefond, Jean-Philippe Steyer, Francis Mairet.

Experimental developments

Experiments have been carried out to study the effects of nitrogen limitation on the lipid production in microalgae and support model development. These experiments have been carried out in the Lagrangian simulator, under constant or periodic light and temperature, varying the total amount of light dose in the day. The response in terms of storage carbon (triglycerides and carbohydrates) has been observed.

Other experiments were carried out to reproduce the light percept by a cell in a raceway pond [74], that is a large-scale raceway-track shaped open-air photobioreactor with circulating medium. An electronic platform was developed to reproduce the flashing light which, from the hydrodynamical studies, is likely to happen in a raceway at the cell scale. The experiments show that the microalgae adapt their pigments to the average light that they have received.

The effect in the cell cycle of both the light periodic signal and a nitrogen limitation were studied. The strong interactions of the interactions between the different phases of the cell cycle through checkpoints was highlighted [104].

Finally, we have tested the effect of cement flue gas on microalgae growth and demonstrated that this CO_2 source can be used to feed microalgal industrial cultures [33].

These works have been carried out in collaboration with A. Talec, S. Rabouille, E. Pruvost and C. Combe (CNRS/UPMC -Océanographic Laboratory of Villefranchesur-Mer).

In collaboration with the IFREMER-PBA team (Nantes) we contributed to a study (within the Symbiose project) of the possible associations between microalgae and bacteria to enhance overall productivity [27].

Metabolism of carbon storage and lipid production

A macroscopic model for lipid production by oleaginous microalgae [10] has been previously proposed. This model describes the accumulation of neutral lipids (which can be turned into biofuel), carbohydrates and structural carbon. We now start to progressively dig deeper in the metabolism, with the objective to better predict carbohydrate and lipid accumulation[37], [64].

Modeling the coupling between hydrodynamics and biology

In collaboration with the Inria ANGE team, a model coupling the hydrodynamics of the raceway (based on multilayer Saint Venant system) with microalgae growth was developed [86]. This model is supported by the work of ANGE aiming at reproducing the hydrodynamics of the raceway, with a specific attention to the effect of the paddle wheel on the fluid.

Modeling the photosynthesis response to fast fluctuating light

The impact of the hydrodynamics on the light percept by a single cell was studied thanks to fluid dynamics simulations of a raceway pond [26]. The light signals that a cell experiences at the Lagrangian scale, depending on the fluid velocity, were then estimated. A Droop-Han model was used to assess the impact of light fluctuation on photosynthesis. A new model accounting for photoacclimation was also proposed [46]. Single cell trajectories were simulated by this tool, and the effect on photosynthesis efficiency was assessed using models of photosynthesis [94]. These results were compared to experimental measurements where the high frequency light was reproduced [74].

Modeling a microalgae production process

The integration of different models developed in the group [81], [101], [10] was performed to represent the dynamics of microalgae growth and lipid production in raceway systems, on the basis of the dynamical model developed to describe microalgal growth in a photobioreactor under light and nitrogen limitations. The strength of this model is that it takes into account the strong interactions between the biological phenomena (effects of light and nitrogen on growth, photoacclimation ...), temperature effect [85],[31] and the radiative transfer in the culture (light attenuation due to the microalgae).

Using these approaches, we have developed a model which predicts lipid production in raceway systems under varying light, nutrients and temperature [30]. This model is used to predict lipid production in the perspective of large scale biofuel production.

Finally, we provide guidelines for the design of experiments with high informative content that allows an accurate parameter estimation of this model, concerning the effect of temperature and light on microalgae growth. The optimal experiment design problem was solved as an optimal control problem. E-optimal experiments were obtained by using two discretization approaches namely sequential and simultaneous. Simulation results showed the relevance of determining optimal experimental inputs for achieving an accurate parameter estimation [50].

Nitrogen fixation by nitrogenotrophs

The fixation of nitrogen by *Croccosphera watsonii* was represented with a macro metabolic model [44]. The main fluxes of carbon and nitrogen are represented in the cell. The accumulation of starch during the day to fuel the nitrogenase working in the absence of oxygen during the night was the key process to explain the nitrogen fixation. The strong influence of the cell cycle was also included in the model. Finally, the model was calibrated and validated with the data of 3 experiments carried out with different duration of the light period and daily dose. The model succeded to efficiently reproduce the experimental data.

This work is done in collaboration with Sophie Rabouille (CNRS-Océanographic Laboratory of Villefranche-sur-Mer).

Including phytoplankton photoadaptation into biogeochemical models

The complexity of the marine ecosystem models and the representation of biological processes, such as photoadaptation, is very challenging to tackle so that their representation remains an open question. We compared several marine ecosystem models with increasing complexity in the phytoplankton physiology representation in order to assess the consequences of the complexity of photoadaptation models in biogeochemical model predictions. Three models of increasing complexity were considered, and the models were calibrated to reproduce ocean data acquired at the Bermuda Atlantic Time-series Study (BATS) from in situ JGOFS (Joint Global Ocean Flux Study) data. It turns out that the more complex models are trickier to calibrate and that intermediate complexity models, with an adapted calibration procedure, have a better prediction capability [15].

This work is done in collaboration with Sakina Ayata (UPMC-Océanographic Laboratory of Villefranche-sur-Mer).

6.2.1.2. Control and Optimization of microalgae production

On-line monitoring

Interval observers give an interval estimation of the state variables, provided that intervals for the unknown quantities (initial conditions, parameters, inputs) are known [7]. Several developments were carried out in this direction to improve the design and performances of interval observers. The approach has been applied to estimation of the microalgae growth and lipid production within a production process [28].

Optimization of the bioenergy production systems

Based on simple microalgae models, analytical optimization strategies were proposed. We first focused on the optimal operating conditions for the biomass productivity under day/night cycles using Pontryiagin's maximum principle (assuming a periodic working mode) [25].
On the other hand, we assessed strategies for optimal operation in continuous mode using the detailed model for raceways [49], [30]. Two strategies were developed. The first one resides in solving numerically an optimal control problem in which the input flow rate of the raceway is calculated such that the productivity in microalgae biomass is maximized on a finite time horizon. In the second strategy, we aimed at translating the optimization problem into a regulation problem. We proposed a simple operational criterion that when integrated in a strategy of closed-loop control allows to attain biomass productivities very near to the maximal productivities obtained with the optimal control. We demonstrated that the practical advantages for real implementation makes our proposed controller a suitable control strategy for optimizing microalgae production in raceways.
We also propose a nonlinear adaptive controller for light-limited microalgae culture, which regulates the light absorption factor (defined by the ratio between the incident light and the light at the bottom of the reactor). We show by numerical simulation that this adaptive controller can be used to obtain near optimal productivity under day-night cycles [47].

Interactions between species

Large scale culture of microalgae for bioenergy involves a huge biodiversity (different mutants, invasion, growth-promoting bacteria [96]...). Control of such system requires to consider the interactions between the different species.

In the framework of the ANR Facteur 4 project, we propose to drive this competition exploring different strategies in order to select species of interest.

We have proposed an adaptive controller which regulates the light at the bottom of the reactor [48]. When applied for a culture with n species, the control law allows the selection of the strain with the maximum growth rate for a given range of light intensity. This is of particular interest for optimizing biomass production as species adapted to high light levels (with low photoinhibition) can be selected.
Other strategies (e.g. periodic temperature stress) are now under investigation through simulations (in order to design selection experiments that will be performed at LOV) and model analysis.
Finally, in a more theoretical framework, we studied how to select as fast as possible a given species in a chemostat with two species at the initial instant. Using the Pontryagin maximum principle, we have shown that the optimal strategy is to maintain the substrate concentration to the value maximizing the difference between the growth rates of two species [66].

6.2.2. Design of ecologically friendly plant production systems

6.2.2.1. Controlling plant pests

Participants: Frédéric Grognard, Ludovic Mailleret, Suzanne Touzeau, Mickaël Teixeira-Alves, Nicolas Bajeux.

Optimization of biological control agent introductions

The question of how many and how frequently natural enemies should be introduced into crops to most efficiently fight a pest species is an important issue of integrated pest management. The topic of natural enemies introductions optimization has been investigated for several years [9] [105], unveiling the crucial influence of within-predator density dependent processes. Because contrarily to predatory biocontrol agents, parasitoids may be more prone to exhibit positive density dependent dynamics rather than negative ones, the current modeling effort concentrates on studying the impact of positive predator-predator interactions on the optimal introduction strategies [72]. Connected experimental research is also being pursued in the laboratory on *trichogramma spp.* which tends to show positive density dependence because of demographic stochasticity [35], and the PhD thesis of Thibaut Morel Journel (UMR ISA) has just started on this topic. Non-impulsive positive feedback control of predator-prey systems in that framework was also addressed in [45].

Food source diversity and classical biological control efficiency using generalist natural enemies

Because generalist biocontrol agents can feed on different food sources like, e.g. a given pest and pollen, they are capable of surviving pest absence within crops and, when supplied with different food types, generalist biocontrol agents are expected to thrive. However, feeding on different food sources means that a given individual cannot feed on each food source at the same moment, which thus potentially reduces the overall predation pressure imposed by the natural enemy population. We exhibited conditions under which the predator distraction effects can dominate the demographic response of the predator populations, potentially disrupting pest control [12]. Such results were at the center of Mickaël Teixeira Alves's PhD thesis.

Plant compensation, pest control and plant-pest dynamics

Introducing a plant compartment into our models, we first focused on plant-insect interactions and showed how the level and timing of the pest invasion and pests control interventions could have important effects on the plant's growth pattern and its final biomass. We then modelled plant compensation, which is the process by which some plants respond positively to recover from the effects of pest injury. We have shown that depending on plants and pests characteristics, as well as the level of pest attack, plant overcompensation may or may not happen [97].

This work is part of the PhD thesis of Audrey Lebon (Cirad), and done in collaboration with Yves Dumont (Cirad).

6.2.2.2. Controlling plant pathogens

Participants: Frédéric Grognard, Ludovic Mailleret, Suzanne Touzeau, Elsa Rousseau.

Sustainable management of plant resistance

Because in addition to being eaten, plants can also get sick, we studied other forms of biological control dedicated to fight plant pathogens. One such method is the introduction of plant strains that are resistant to one pathogen. This often leads to the appearance of virulent pathogenic strains that are capable of infecting the resistant plants. It is therefore necessary to develop ways of introducing such resistance into crop production without jeopardizing its future efficiency. Considering plant viruses, we computed the proportion of resistant plants that should be cropped together with the non-resistant ones in a seasonal model, in order to optimize the resistance for production or patrimonial objectives [53]. The study of factors influencing resistance breakdown from the within-plant to the landscape level is the topic of Elsa Rousseau's PhD thesis, with emphasis both on experimental and modelling approaches. Experiments have been held in Avignon to determine the respective impacts of selection and genetic drift on resistance breakdown.

This work is done in collaboration with Frédéric Fabre and Benoit Moury (INRA Avignon).

Eco-evolutionary dynamics of plant pathogens in seasonal environments

Understanding better pathogen evolution also requires to understand how closely related plant parasites may coexist. Indeed, such coexistence is widespread and is hardly explained through resource specialization. We showed that, in agricultural systems in temperate environments, the seasonal character of agrosystems can induce complex plant-pathogens dynamics [98] and is an important force promoting evolutionary diversification of plant pathogens [93]. Plant parasites reproduction mode may also strongly interact with seasonality. In this context, we investigated the influence of cyclical parthenogenesis, i.e. the alternation of sexual and asexual reproduction phases, on the eco-evolutionary dynamics of plant parasites [59], [60], [21].

This work is part of the PhD thesis of Magda Castel (Agrocampus Ouest) and is done in collaboration with Frédéric Hamelin (Agrocampus Ouest).

6.2.3. Biological depollution

6.2.3.1. Coupling microalgae to anaerobic digestion

Participants: Olivier Bernard, Antoine Sciandra, Jean-Philippe Steyer, Frédéric Grognard, Philipp Hartmann, Francis Mairet.

The coupling between a microalgal pond and an anaerobic digester is a promising alternative for sustainable energy production and wastewater treatment by transforming carbon dioxide into methane using light energy. The ANR Symbiose project is aiming at evaluating the potential of this process [108], [107].

In a first stage, we developed models for anaerobic digestion of microalgae. Two approaches were used: first, a dynamic model has been developed trying to keep a low level of complexity so that it can be mathematically tractable for optimisation [100] . Considering three main reactions, this model fits adequately the experimental data of an anaerobic digester fed with *Chlorella vulgaris* (data from INRA LBE). On the other hand, we have tested the ability of ADM1 [109] (a reference model which considers 19 biochemical reactions) to represent the same dataset. This model, after modification of the hydrolysis step [102] has then been used to evaluate process performances (methane yield, productivity...) and stability though numerical simulations.

6.2.3.2. Life Cycle Assessment

Participants: Olivier Bernard, Jean-Philippe Steyer.

This work is the result of a collaboration with Laurent Lardon and Arnaud Helias of INRA-LBE through the co-supervision of Pierre Collet's PhD thesis [88].

An analysis of the potential environmental impacts of biodiesel production from microalgae has been carried out using the life cycle assessment (LCA) methodology [95]. This study has allowed to identify the obstacles and limitations which should receive specific research efforts to make this process environmentally sustainable. This study has been updated and the effects of technological improvements (leading to higher productivities) have been compared to the source of electricity. It turns out that the overall environmental balance can much more easily be improved when renewable electricty is produced on the plant [91], [90]. As a consequence, a new paradigm to transform solar energy (in the large) into transportation biofuel is proposed, including a simultaneous energy production stage. This motivated the design of the purple sun ANR-project.

These studies have allowed to identify the obstacles and limitations which should receive specific research efforts to make this process environmentally sustainable [65].

A LCA has been carried out to assess the environmental impact of methane production by coupling microalgae and anaerobic digestion. The study highlights the limitation derived by the low biodegradability of the considered microalgae [89] which induces a large digester design and thus more energy to mix and heat it.

These works have been carried out in collaboration with E. Latrille and B. Sialve (INRA - Laboratory of Environmental Biotechnology, Narbonne).

6.2.4. Models of ecosystems

6.2.4.1. Optimality/games in population dynamics

Participants: Frédéric Grognard, Ludovic Mailleret, Pierre Bernhard.

Adaptive behavior in seasonal consumer-resource dynamics

In this work we studied the evolution of a consumer-resource (or predator-prey) system with seasonal character of the dynamics. We specified two main parts of the process. First, we considered the system during one season with a fixed length: the prey lay eggs continuously and the predators lay eggs or hunt the prey (choose their behavior) according to the solution of an optimal control problem [76]. We then showed that, in most situations, mutants can take advantage of their low frequency and fare better than the residents. Over the course of a large number of seasons, the mutants replace the residents, only to find themselves applying the original resident behavior [61].

Optimal foraging and residence times variations

Charnov's marginal value theorem (MVT) [87] is a central tenet of ecological theory. In fragmented environments, the MVT connects the quality and distribution of patches to the optimal time an individual should spend on any patch, and thus the rate of movement in the habitat. Unfortunately, it does not offer explicit predictions regarding how changing habitat quality would affect residence times. In this work, we answer that question in a very general setting, for habitats with homogeneous or heterogeneous patches and with general fitness functions. We then particularize it to the resource consumption framework and indicate how the residence times variations relate to the curvatures of the functional responses,[20].

This last work is done in collaboration with Vincent Calcagno and Eric Wajnberg (INRA Sophia Antipolis)

The handicap paradox

We have investigated the “handicap paradox” of sexual selection, and more specifically revisited Grafen's mathematical models of Zahavi's “handicap principle”. The paradox is that in many species, male secondary sexual characters that clearly attract the females are so developed as to be a handicap to the male's viability, and therefore should be counter-selected by evolution. Zahavi's explanation, made mathematical by Grafen, is that if this secondary sexual character is a signal to the female of the male's quality that she cannot observe otherwise, if this signal were costless, it could be cheated, a low quality male being induced to mimic the signal of a high quality one. We have cast this problem into a signaling game, using the bayesian equilibrium of game theory. This easily shows that indeed, under mild conditions, at equilibrium the signal should be “costly”. We have developed several models inspired by Grafen, and to a lesser extent Getty, with explicit solutions, and explained why an undesirable feature appeared in Grafen's model (as well as in one of ours) and proposed a model free of this artifact [19].

6.3. Software design

6.3.1. Odin

Participants: Olivier Bernard, Mélaine Gautier.

Over the years, BIOCORE has been developing a software framework for bioprocess control and supervision called ODIN [80]. This C++ application (working under Windows and Linux) enables researchers and industrials to easily develop and deploy advanced control algorithms through the use of a Scilab interpreter [82], [83]. It also contains a Scilab-based process simulator which can be harnessed for experimentation and training purposes. ODIN is primarily developed in the C++ programming language and uses CORBA to define component interfaces and provide component isolation. ODIN is a distributed platform, enabling remote monitoring of the controlled processes as well as remote data acquisition. Recently, a software development effort has been directed to the graphical user interface, a synoptic view component, new drivers for the experimental hardware and integration of the PlantML data exchange format. ODIN has been tested on four different processes and has been set up with Eric Latrille to supervise the 66m2 high rate pond at the LBE, INRA Narbonne.

6.3.2. In@lgae

Participants: Etienne Delclaux, Francis Mairet, Olivier Bernard.

The simulation platform In@lgae is jointly developed with the Inria Ange team. Its objective is to simulate the productivity of a microalgae production system, taking into account both the process type and its location and time of the year. A first module (Freshkiss) developed by Ange computes the hydrodynamics, and reconstructs the Lagrangian trajectories percept by the cells. Coupled with the Han model, it results in the computation of an overall photosynthesis yield. A second module is coupled with a GIS (geographic information system) to take into account the meteorology of the considered area (any location on earth). The evolution of the temperature in the culture medium together with the solar flux is then computed. Finally, the productivity in terms of biomass, lipids, pigments together with CO_2, nutrients, water consumption, ... are assessed. The productivity map which is produced can then be coupled with a resource map describing the availability in CO_2 nutrients and land.

CARMEN Team

6. New Results

6.1. Mathematical models

- *Mathematical derivation of a bilayer surface model of the atria using asymptotic analysis methods [28], [16]*

  We derived rigorously, by using asymptotic analysis tools, a bilayer model of atrial electrophysiology. Starting with a 3D model of atrial tissue that includes two layers with distinct electrophysiological characteristics and with an aspect ratio of ϵ, we obtained an asymptotic equivalent model when $\epsilon \longrightarrow 0$ made up of two surface models coupled by a coupling term. The bilayer model discribes the evolution of the mean in the thickness of the 3D potential in each layer. This approach is an improvement of the classical surface model of cardiac electrophysiology, because it guaranties a higher convergence speed, and allows to take into account transmural heterogeneities. We numerically implemented the 3D and bilayer models and compared it to the classical surface model. We observed a second order accuracy of the bilayer model and drastically reduced computational times respectively to the 3D model.

- *Formal derivation of a macroscopic model of propagation that includes the non linear behavior of gap junctions [16]*

  A macroscopic model of electrical propagation that take into account the non linear conductivity of the gap junction is obtained by a formal homogenization method. We derived a one dimensional macroscopic model which diffusive tensor varies in time. We compared this macroscopic model with a cell-to-cell propagation. This is an very important improvement of existing models that only consider a linear cell-to-cell coupling. The introduction of this non linear phenomenon in homogenized models gives a simulation tool to investigate the impact of the microscopic nonlinear mechanisms on the macroscopic propagation.

- *Influence of periodic diffusive inclusions on the bidomain model [29]*

  We present a new mathematical model of the electric activity of the heart. In the standard bidomain model we can distinguish the intra- and the extracellular space with different conductivities for excitable cells and the fibrotic tissue around them. The main drawback is that it assumes the existence of excitable cells everywhere in the heart, while it is known that there exist non small regions where fibroblasts take place. The fibroblasts are equally distributed and since they are non excitable cells, they can be considered as a diffusive part. Hence we extend the standard bidomain model as follows: we assume that we have periodic alternation of the healthy tissue (linear bidomain model) and fibrotic extracellular space (diffusive part). We use homogenisation techniques to derive our macroscopic partial differential equations. Interestingly, we obtain again a bidomain type model with modified conductivities that involve the volume fraction of the diffusive domain. Preliminary numerical experiments will conclude on the influence of these diffusive inclusions.

6.2. Construction of numerical models

- *Implementation of an accurate bilayer model of human atria, including realistic geometry and qualitative fibre direction [19], [21], [16]*

  We introduce a bilayer model of the human atria. We set a specific mathematical model based on two surface monodomain problems coupled by a coupling term. We recalled convergence results of the bilayer model towards a 3D model for thin tissues, we formalized an optimization method to set the coupling coefficient and we present two different asymptotically equivalent numerical implementation of the model. We then present a geometrically and electrophysiologically accurate model of the atrial heterogeneities, including two layers of fibre directions and ionic function heterogeneities based on histological and modelling works. We assess the physiological relevance of the model during a sinus wave and we check the occurrence of three-dimensional electrical behaviour such as slight electrical dissociation. This bilayer model is able to take into account transmural heterogeneities only accessible since then with full 3D models, while keeping the low computational load associated with surface models. It is then a light and relevant tool for long-lasting simulations designed to investigate atrial arrhythmia.

- *Personalization method of the bilayer model to registrate the geometry to a patient dependant geometry [16]*

  If the generic atrial bilayer model developed in [21] allows to conduct general experiments, greater customization of the model is necessary to carry out more specific studies on a given patient. We present a methodology to obtain a patient-dependant model containing the geometry of the patient, a generic fibrous organization and an image of the patient's fibrosis obtained by late-enhancement MRI. This is a common work with the clinical team of the CHU du Haut-Lévêque (H. Cochet and P. Jaïs) and the Asclepios Inria team.

  The methodology is based on a registration method developed by Durrleman et al. [34] that allows to register surfaces : the generic model is registrated towards a patient-specific geometry (work by M. Sermseant and R. Cabrera-Lozoya, Asclepios Team). The fibre organisation is transported by the same linear local transformations. The late-enhancement is projected on the model to take into account the complex patient-specific fibrotic repartition. A similar methodology was presented by McDowell et al. [37]. However, the authors took as a starting point a three-dimensional geometry and a different methodology to registrate the geometry. The work presented here is therefore innovative.

- *Faster solvers for cardiac electro-physiology problems [27]*

  There are many applications in cardiac electro-physiology where computational time is the main requirement to fulfill, even by sacrificing accuracy. Some techniques were investigated in this direction, in order to obtain a break-even point between accuracy and speed. The complete problem involves solving some ODEs on each mesh node and inverting large sparse matrices, often ill-conditionned.

  We first designed a method based on the Proper Orthogonal Decomposition (POD) technique: we project the linear system onto a well-chosen orthogonal basis of smaller dimension while still solving the ODES. We tried the method on both the bidomain and monodomain equations, and extended the tests on an HPC machine, in order to observe scalability performances. There is no improvment for the monodomain equations because its linear systems are well-conditionned. For the bidomain equations, the CPU time decreases by a factor of 10 between the full and reduced models, and better scalability performances.

  We secondly developped an eikonal model, in view of serious games applications for the Medic Activ project. The Dijkstra algorithm is used to solve the eikonal equation and the transmembrane potential is determined by the solution of a Mitchell-Schaeffer model on each mesh node. Some modifications where introduced to take into account re-excitability and allox re-entrant waves. Compared by the algorithm proposed by [38], the transmembrane potential comes from the solution of an underlying model, not through an approximation. This represent an innovation, to our knowledge not present in literature.

6.3. Medical applications of numerical models

- *Influence of Transmural Slow-Conduction Zones on the Long-Time Behaviour Of Atrial Arrhythmia. A Numerical Study with a Human Bilayer Atrial Model. [20], [31], [16]*

  Atrial fibrosis is known to be a factor in the perpetuation of atrial arrhythmia. Despite the thinness of atrial tissue, the fibrosis distribution may not be homogeneous through the entire thickness of the atria. The aim of this study is twofold. 1) We want to elucidate the respective influences of a transmural and a non-transmural distribution of fibrosis, described as a slow conduction zone, on the perpetuation of a rotor-like arrhythmic episode, compared to a control situation. 2) We aim to assess which is the more efficient ablation protocol between a) a lesion-box ablation, b) an ablation line connecting the fibrotic zone to the closest anatomical obstacle, c) ablation spots.

We used a bilayer monodomain representation of the atria that included transmural heterogeneities of fibre organisation, and an arrhythmic scenario composed of a rotor initiated near the pulmonary veins. This model allowed long simulations for a sustainable computational load. We observed that when the fibrosis was transmural, the centre of the rotor was anchored in the slow conduction zone and was stable during a 10 seconds simulation, whereas the other simulations showed meandering rotors that disappeared after a few seconds. In our model framework, only a transmural fibrosis distribution had a stabilizing effect on reentrant circuits. Furthermore, the lesion-box ablation and the line ablation were able to stop the arrhythmia, unlike the spot lesions. The bilayer model proved to be a good trade-off between accuracy and speed for observing the influence of transmural heterogeneities on atrial arrhythmia over long periods.

- *Effects of L-type Calcium channel and hERG blockers on the electrical activity of the human heart: A simulation study.*

  Class III and IV drugs affect cardiac hERG (IKr) and L-type calcium (ICaL) channels, resulting in complex alterations in repolarization with both anti and pro-arrhythmic consequences. Interpretation of their effects on cellular and ECG-based biomarkers for risk stratification is challenging. As pharmaceutical compounds often exhibit multiple ion channel effects, our goal is to investigate the simultaneous effect of ICaL and IKr block on human ventricular electrophysiology from ionic to ECG level. ECG simulations show that ICaL block results in shortening of the QT interval, ST elevation and reduced T wave amplitude, caused by reduction in APD and AP amplitude during the plateau phase, and in repolarization times. In contrast, IKr block results in QT prolongation and reduced T wave amplitude. Combined ICaL and IKr block are combined, the degree of ICaL block strongly determines QT interval whereas the effect of IKr block is more pronounced on the T wave amplitude.

6.4. Inverse problems

- *A Steklov-Poincaré approch to solve the inverse problem in electrocardiography [23]*

  In the cardiac electrophysiology imaging commu- nity the most widely used approach to solve the inverse prob- lem is the least square formulation with different Thikhonov regularizations. Clinicians are not yet fully satisfied by the technology that solves the inverse problem. Reformulating the inverse problem could bring new techniques to solve it. In this paper we use the Steklov-Poincare ´ formulation of the Cauchy problem in order to solve the inverse problem in electrocardiography imaging. We present in this work the technique and an algorithm of gradient descent. We also show numerical results based on simulated synthetical data.

- *A machine learning regularization of the inverse problem in electrocardiography imaging [22]*

  Radio-frequency ablation is one of the most ef- ficient treatments of atrial fibrillation. The idea behind it is to stop the propagation of ectopic beats coming from the pulmonary vein and the abnormal conduction pathways. Medical doctors need to use invasive catheters to localize the position of the triggers and they have to decide where to ablate during the intervention. ElectroCardioGraphy Imaging (ECGI) provides the opportunity to reconstruct the electrical potential and activation maps on the heart surface and analyze data prior to the intervention. The mathematical problem behind the reconstruction of heart potential is known to be ill posed. In this study we propose to regularize the inverse problem with a statistically reconstructed heart potential, and we test the method on synthetically data produced using an ECG simulator.

- *Inverse problem in electrocardiography via factorization method of boundary value problems : How to reconstruct epicardial potential maps from measurements on the torso ? [26]*

We are working on a new approach for solving the inverse problem of electrocardiography. This approach is based on an invariant embedding method: the factorization method of boundary values problems [35]. The idea is to embed the initial problem into a family of similar problems on subdomains bounded by a moving boundary from the torso skin to the epicardium surface. For the direct problem this method provides an equivalent formulation with two Cauchy problems evolving on this moving boundary and which have to be solved successively in opposite directions. This method calculates Neumman-Dirichlet and Dirichlet-Neumann operators on this moving boundary that satisfy Riccati equations. Regarding the inverse problem, mathematical analysis allows to write an optimal estimation of the epicardial potential based on a quadratic criterion. Then, the ill-posed behaviour of the inverse problem can be analyzed and a better regularization and discretization of the problem can be proposed. One of the advantages of this method is the computation of the potential at different times during cardial cycle: it is not necessary to repeat the resolution of all the equations at every time. In a first time the simplar case of a cylinder is considered. In a second time the method is applied to the 3D model of concentric spheres. The next step will be to use 3D deformed surfaces.

- *Reconstruction of 3D depolarization wavefronts from surface optical mapping images [33]*

  Starting from the diffusion-absorption equation of light in a tissue we solved the forward problem for excitation light using the `FreeFem++` software (www.freefem.org/ff++). We first considered a spherical wave front expanding in time: the tissue is depolarized inside the sphere. This choice allowed us to locate the position of the excitation. Using this representation of the wavefront, we obtained in silico data. We defined a functional to minimize and implemented the BFGS method to solve the inverse problem. We tested our method on in silico data and obtained good results. We next compared our results with an approach developed by Khait [36] and found that our method is more accurate and that we have less restrictions for the convergence of the method. We modified the wave front into ellipsoid in order to start working on experimental data.

DRACULA Project-Team

6. New Results

6.1. Mathematical modeling in chronobiology

Circadian clocks are autonomous oscillators entrained by external Zeitgebers such as light-dark and temperature cycles. On the cellular level, rhythms are generated by negative transcriptional feedback loops. In mammals, the suprachiasmatic necleus (SCN) in the anterior part of the hypothalamus plays the role of the central circadian pacemaker. Coupling between individual neurons in the SCN leads to specify self-sustained oscillations even in the absence of external signals. These neuronal rhythms orchestrate the phasing of circadian oscillations in peripheral organs. Altogether, the mammalian circadian system can be regarded as a network of coupled oscillators. In order to understand the dynamic complexity of these rhythms, mathematical models successfully complement experimental investigations. In [19], we discuss basic ideas of modeling on three different levels: (i) rhythm generation in single cells by delayed negative feedbacks, (ii) synchronization of cells via external stimuli or cell-cell coupling, and (iii) optimization of chronotherapy.

6.2. Hybrid Models of Cell Population

The paper [20] is devoted to hybrid discrete-continuous models of cell populations dynamics. Cells are considered as individual objects which can divide, die by apoptosis, differentiate and move under external forces. Intra-cellular regulatory networks are described by ordinary differential equations while extracellular species by partial differential equations. We illustrate the application of this approach to some model examples and to the problem of tumor growth. Hybrid models of cell populations present an interesting nonlinear dynamics which is not observed for the conventional continuous models.

6.3. Multiscale Model in Biology

Biological processes span several scales in space, from the single molecules to organisms and ecosystems. Multiscale modelling approaches in biology are useful to take into account the complex interactions between different organisation levels in those systems. We present in [6] several single- and multiscale models, from the most simple to the complex ones, and discuss their properties from a multiscale point of view. Approaches based on master equations for stochastic processes, individual-based models, hybrid continuous-discrete models and structured PDE models are presented.

6.4. Model of hematopoiesis

In [2], a model of blood cell production in the bone marrow (hematopoiesis), has been investigated. It describes both the evolution of primitive hematopoietic stem cells and the maturation of these cells as they differentiate to form the three kinds of progenitors and mature blood cells (red blood cells, white cells and platelets). The three types of progenitors and mature cells are coupled to each other via their common origin in primitive hematopoietic stem cells compartment. The resulting system is composed by eleven age-structured partial differential equations. To analyze this model, we don't take into account cell age-dependence of coefficients, that prevents a usual reduction of the structured system to an unstructured delay differential system. We study the existence of stationary solutions: trivial, axial and positive steady states. Then we give conditions for the local asymptotic stability of the trivial steady state and by using a Lyapunov function, we obtain a sufficient condition for its global asymptotic stability. In some particular cases, we analyze the local asymptotic stability of the positive steady state by using the characteristic equation. Finally, by numerical simulations, we illustrate our results and we show that a change in the duration of cell cycle can cause oscillations.

6.5. Analysis of radiocarbon to facilitate identification of unknown decedents

The characterization of unidentified bodies or suspected human remains is a frequent and important task for forensic investigators. However, any identification method requires clues to the person's identity to allow for comparisons with missing persons. If such clues are lacking, information about the year of birth, sex and geographic origin of the victim, is particularly helpful to aid in the identification casework and limit the search for possible matches. We present in [4] results of stable isotope analysis of $(13)C$ and $(18)O$, and bomb-pulse $(14)C$ analyses that can help in the casework. The $(14)C$ analysis of enamel provided information of the year of birth with an average absolute error of 1.8 ± 1.3 years. We also found that analysis of enamel and root from the same tooth can be used to determine if the $(14)C$ values match the rising or falling part of the bomb-curve. Enamel laydown times can be used to estimate the date of birth of individuals, but here we show that this detour is unnecessary when using a large set of crude $(14)C$ data of tooth enamel as a reference. The levels of $(13)C$ in tooth enamel were higher in North America than in teeth from Europe and Asia, and Mexican teeth showed even higher levels than those from USA. DNA analysis was performed on 28 teeth, and provided individual-specific profiles in most cases and sex determination in all cases. In conclusion, these analyses can dramatically limit the number of possible matches and hence facilitate person identification work.

M3DISIM Team

6. New Results

6.1. Modeling

6.1.1. Collective effect in molecular motors assembly

Participants: Matthieu Caruel, Jean-Marc Allain [Ecole Polytechnique], Lev Truskinovsky [Ecole Polytechnique].

Skeletal muscles consist of active material capable of producing force. At the microscale, force is the result of complex interactions between two types of proteins named actin and myosin that work coherently in very large assemblies ($\sim 10^9$). The passive mechanical response of skeletal muscles at fast time scales is dominated by long range interactions inducing cooperative behavior without breaking the detailed balance. This leads to such unusual “material properties” as negative equilibrium stiffness and different behavior in force and displacement controlled loading conditions. Our fitting of experimental data suggests that “muscle material” is finely tuned to perform close to a critical point which explains large fluctuations observed in muscles close to the stall force. See paper [28].

6.1.2. Dimensional reductions of a cardiac model for effective validation and calibration

Participants: Matthieu Caruel, Alexandre Imperiale, Radomir Chabiniok [King’s College London], Philippe Moireau, Dominique Chapelle, Yves Lecarpentier [Institut du Cœur].

Complex 3D beating heart models are now available, but their complexity makes calibration and validation very difficult tasks. We thus propose a systematic approach of deriving simplified reduced-dimensional models, in “0D”— typically, to represent a cardiac cavity, or several coupled cavities and in “1D”—to model elongated structures such as muscle samples or myocytes. We apply this approach with an earlier-proposed 3D cardiac model designed to capture length-dependence effects in contraction, which we here complement by an additional modeling component devised to represent length-dependent relaxation. We then present experimental data produced with rat papillary muscle samples when varying preload and afterload conditions, and we achieve some detailed validations of the 1D model with these data, including for the length-dependence effects that are accurately captured. Finally, when running simulations of the 0D model pre-calibrated with the 1D model parameters, we obtain pressure–volume indicators of the left ventricle in good agreement with some important features of cardiac physiology, including the so-called Frank–Starling mechanism, the End-Systolic Pressure–Volume Relationship, as well as varying elastance properties. This integrated multi-dimensional modeling approach thus sheds new light on the relations between the phenomena observed at different scales and at the local versus organ levels. See papers [13], [22].

6.1.2.1. Surface-based electrophysiology modeling and assessment of physiological simulations in atria

Participants: Dominique Chapelle, Annabelle Collin, Jean-Frédéric Gerbeau [Reo Project-Team], Mélèze Hocini [Institut LIRYC - IHU Bordeaux], Michel Haïssaguerre [Institut LIRYC - IHU Bordeaux].

The objective of this work is to assess a previously-proposed surface-based electrophysiology model with detailed atrial simulations. This model – derived and substantiated by mathematical arguments – is specifically designed to address thin structures such as atria, and to take into account strong anisotropy effects related to fiber directions with possibly rapid variations across the wall thickness. The simulation results are in excellent adequacy with previous studies, and confirm the importance of anisotropy effects and variations thereof, see Figure 1 . Furthermore, this surface-based model provides dramatic computational benefits over 3D models with preserved accuracy. See paper [23].

6.1.3. Strong convergence results in the asymptotic behavior of the 3D-shell model

Participants: Dominique Chapelle, Annabelle Collin.

`../../../../projets/m3disim/IMG/atria.png`

Figure 1. Surface-based modeling of atrial electrophysiology

The objective of this work is to revisit the asymptotic convergence properties – with respect to the thickness parameter – of the earlier-proposed 3D-shell model. This shell model is very attractive for engineering applications, in particular due to the possibility of directly using a general 3D constitutive law in the corresponding finite element formulations. We establish strong convergence results for the 3D-shell model in the two main types of asymptotic regimes, namely, bending- and membrane-dominated behavior. This is an important achievement, as it completely substantiates the asymptotic consistency of the 3D-shell model with 3D linearized isotropic elasticity. See paper [14].

6.1.4. Mechanical modeling and numerical methods for poromechanics: Applications to cardiac perfusion

Participants: Bruno Burtschell, Dominique Chapelle, Philippe Moireau.

We have previously formulated a rather general modeling framework of poromechanics – formulations that combine solid and fluid components to represent the behavior of a porous medium – to take into account large deformations and rapid fluid flows, see [29]. This allows to consider, in particular, the application of blood perfusion within the cardiac tissue, which – indeed – features these specific complex phenomena, out of the scope of classical poromechanical models. One of our major objectives now, within the PhD of Bruno Burtschell, is to propose and assess some associated relevant numerical schemes, which requires special care regarding both space and time discretizations. In a first stage, in order to ease numerical prototyping and assessments, an axisymmetric reduction of the model has been formulated, and some existing algorithms of fluid-structure interaction have been implemented within this axisymmetric framework (in FreeFEM++). The rationale is that our poromechanics formulations feature some rather deep similarities to so-called Arbitrary-Lagrangian-Eulerian (ALE) fluid-structure formulations, hence the latter are considered as a natural starting point for further extensions.

6.1.5. Personalized modeling for cardiac amyloidosis diagnosis

Participants: Alessandro Felder, Dominique Chapelle, Philippe Moireau, Jean-François Deux [Hôpital Henri Mondor], Thibault Damy [Hôpital Henri Mondor].

Cardiac amyloidosis is a condition induced by pathological deposition of amyloid proteins within muscle tissue and nerves, thus severely impairing the cardiac function and often requiring cardiac transplantation as the only available treatment. Our objective here in a first stage is to use our previously developed patient-specific modeling methodologies to analyse some clinical cases – based on actual patient data – to better apprehend the impact of the pathology on biomechanical properties. Further perspectives include the modeling of the protein deposition and associated tissue remodeling in order to predict the disease evolution in a patient-specific context. This work is performed in collaboration with medical doctors from Hôpital Henri Mondor (Créteil).

6.2. Model-Data Interaction

6.2.1. State observers of a vascular fluid-structure interaction model through measurements in the solid

Participants: Cristobal Bertoglio [Reo Project-Team], Dominique Chapelle, Miguel Angel Fernández [Reo Project-Team], Jean-Frédéric Gerbeau [Reo Project-Team], Philippe Moireau.

We analyze the performances of two types of Luenberger observers – namely, the so-called Direct Velocity Feedback and Schur Displacement Feedback procedures, originally devised for elasto-dynamics – to estimate the state of a fluid-structure interaction model for hemodynamics, when the measurements are assumed to be restricted to displacements or velocities in the solid. We first assess the observers using hemodynamics-inspired test problems with the complete model, including the Navier-Stokes equations in Arbitrary Lagrangian-Eulerian formulation, in particular. Then, in order to obtain more detailed insight we consider several well-chosen simplified models, each of which allowing a thorough analysis – emphasizing spectral considerations – while illustrating a major phenomenon of interest for the observer performance, namely, the added mass effect for the structure, the coupling with a lumped-parameter boundary condition model for the fluid flow, and the fluid dynamics effect *per se*. Whereas improvements can be sought when additional measurements are available in the fluid domain in order to more effectively deal with strong uncertainties in the fluid state, in the present framework this establishes Luenberger observer methods as very attractive strategies – compared, e.g., to classical variational techniques – to perform state estimation, and more generally for uncertainty estimation since other observer procedures can be conveniently combined to estimate uncertain parameters. See paper [11].

6.2.2. Improving Efficiency of Data Assimilation Procedure for a Biomechanical Heart Model by Representing Surfaces as Currents

Participants: Alexandre Imperiale, Alexandre Routier [Aramis Team], Philippe Moireau, Stanley Durrleman [Aramis Team].

We adapt the formalism of currents to compare data surfaces and surfaces of a mechanical model and we use this discrepancy measure to feed a data assimilation procedure. We apply our methodology to perform parameter estimation in a biomechanical model of the heart using synthetic observations of the endo- and epicardium surfaces of an infarcted left ventricle. We compare this formalism with a more classical signed distance operator between surfaces and we numerically show that we have improved the efficiency of our estimation justifying the use of state-of-the-art computational geometry formalism in the data assimilation measurements processing. See paper [24].

6.2.3. Optimal observer for parabolic problems

Participants: Karine Mauffrey, Philippe Moireau.

../../../../projets/m3disim/IMG/cropped_mesh_img_4.png

Figure 2. Heart model immersed in MR image data

We aim at proposing optimal observers strategies for reconstructing the solution of general systems of PDEs using avalaible observations, including both wave-type equations and heat-like equations. The main objective of this work is to present a self-contained analysis. For a general parabolic system, we have established the exponential stability of the operator occurring in the equation satisfied by the error between the target and the optimal observer. The proof relies on two major hypotheses: an observability inequality satisfied by the observation operator and a controllability property for the modeling error operator by which model noises enter the dynamics (controllability property which is related to the invertibility of the solution of the associated infinite dimensional Riccati equation). The next questions we want to tackle are the discretisation of the model and the construction of a reduced Kalman filter.

MASAIE Project-Team

5. New Results

5.1. Release of Wolbachia as a preventive action against dengue

We have designed a model of infection by *Wolbachia* of a *Aedes aegypti* population, to take into account the biology of this infection and also the data that can be obtained. The objective is to use this model for predicting the sustainable introduction of this bacteria. We provide a complete mathematical analysis of the model proposed and give the basic reproduction ratio $\mathcal{R}_0$ for *Wolbachia* . We observe a bistability phenomenon. Two equilibria are asymptotically stable : an equilibrium where all the population is uninfected and an equilibria where all the population is infected. A third unstable equilibrium exists. We are in a backward bifurcation situation. The bistable situations occurs with natural biological values for the parameters. Our model is an example of an epidemiological model with only vertical transmission.

This infection model is then connected with a classical dengue model. We prove that for the complete model the equilibrium with *Wolbachia* for the mosquitoes and without dengue for the human is asymptotically stable. We prove that, if a sufficiently great population of infected mosquitoes is introduced, dengue will disappear.

We use the data of a real trial of releases of infected mosquitoes in Cairns (Australia) to calibrate our model. Our model behave remarkably well versus the observed data. We use then the calibrated model to simulate different scenarii of appearance of dengue. We use a pessimistic situation where the basic reproduction ratio $\mathcal{R}_0$ of dengue is 24.5. The simulations confirm our findings, dengue epidemics does not occur, and show that the introduction of *Wolbachia* is a promising way of control dengue.

../../../../projets/masaie/IMG/GordonYorkpercent-vert

Figure 1. Frequencies observed and predicted. The red squares are the frequencies of infection given by the model. The blue circles are the frequencies observed

../../../../projets/masaie/IMG/dengueUetW30.png

Figure 2. In green the infected human with Wolbachia present. In blue when mosquitoes have not been infected by Wolbachia.

5.2. Arboviruses on urban environments

We investigate the influence of *human movement* for the onset of an arboviral (mosquito-borne) epidemics (such as Dengue, Chikungunya, West Nile or Yellow fever) on an urban environment. The metapopulation model has a standard SIR (human)/SI (mosquito) model as the basic dynamics on the patches. The nodes consist of notification districts used by public health authorities. The subsystems are coupled by human movement. Our main result provides quantitative relations between three reproduction numbers: *local* - at each isolated subsystem, *uniform* or *mixing* - aggregating the data of the whole region, and the *network* reproduction number - for the coupled dynamics. We observe that the epidemics can spread among the patches as a consequence solely of human movement: while all nodes may have, if isolated, local reproduction ratio less than one and, moreover, the uniform reproduction number being also less than one, however, the network reproduction number can be greater than one. An estimate is provided on the overall effect of vector control on a chosen patch [13].

5.3. Analysis of the dynamics of some models for vector-borne diseases with host circulation

In this work we study the dynamics of a vector borne disease on a metapopulation model that accounts for host circulation. For such models, the movement network topology gives rise to a contact network topology, corresponding to a bipartite graph. Under the assumption that the contact network is strongly connected, we can define the basic reproductive number R_0 and show that this system has only two equilibria: the so called disease free equilibrium (DFE); and a unique interior equilibrium that exists if, and only if, the basic reproduction number, R_0, is greater that unity. We are also able to show that the DFE is globally asymptotically stable, if $R_0 \leq 1$. If $R_0 > 1$, the dynamics is uniformly persistent and, with further assumptions on the contact network structure, we also show that the endemic equilibrium (EE) is globally asymptotically stable [17].

5.4. Analysis and observer design for a schistosomiasis model

Human schistosomiasis is a behavioral and occupational disease associated with poor human hygiene, insanitary animal husbandry and economic activities. Among human parasitic diseases, schistosomiasis ranks second behind malaria as far as the socio-economic and public health importance in tropical and subtropical areas are concerned. The spread and persistence of schistosomiasis have made of it one of the most complex host-parasite process to model mathematically because of the different steps of growth of larvae assumed by the parasite and the requirement of two host elements (definitive human host and intermediate snail hosts) during their life cycle.

An efficient method to control the schistosomiasis infection that may require relatively little funding is a biological control. Particularly, trematode parasites or competitive snails of the intermediate snail hosts have been proved to be effective in controlling schistosomiasis in the Caribbean area.

We have studied a schistosomiasis infection model that involves human and intermediate snail hosts as well as an additional mammalian host and a competitor snail species. This mathematical analysis of the model gives insight about the epidemiological consequences of the introduction of a competitor resistant snail species [15].

We have also proposed a solution to the state estimation problem for a schistosomiasis infection dynamical model described by a continuous non linear system when only the infected human population is measured. We have constructed an estimator that is able to give dynamical estimates of the variables that can not be measured [14].

MODEMIC Project-Team

6. New Results

6.1. Macroscopic models

6.1.1. About species coexistence

Participants: Fabien Campillo, Jérôme Harmand, Claude Lobry, Alain Rapaport, Tewfik Sari.

The so called “Principle of Competitive Exclusion” states that in the chemostat model, in presence of p substrates only p species can coexist. By contrast, in a bioreactor used for decontamination, hundreds to thousand different species are observed in presence of just very few substrates. Actually the classical chemostat models rely on assumptions: perfect mixing, substrate-dependent growth rate, constant environment, only asymptotic results are considered, deterministic continuous models...

A long term objective since Mere Inria project-team is to revisit the chemostat model in the absence of one or more of these hypotheses having in view the question of coexistence. In our “major publications” we proved coexistence in absence of the second hypothesis [6] or during long transient [9]. In [57], we consider the case where the environment (in some sense) is periodic in time. Our results concerning non continuous and/or stochastic models (see Section 6.2.3) are also a first step in avoiding the fifth hypothesis.

6.1.2. Modeling and analysis of bioprocesses

Participants: Boumédiène Benyahia, Radhouane Fekih-Salem, Jérôme Harmand, Claude Lobry, Guilherme Pimentel, Alain Rapaport, Tewfik Sari.

Within the supervision of the PhD thesis of R. Fekih-Salem, we have studied a chemostat model where the species are present in two forms, isolated and aggregated individuals, such as attached bacteria in biofilm or bacteria in flocks [22]. We show that our general model contains a lot of models that were previously considered in the literature [77]. Assuming that flocculation and deflocculation dynamics are fast compared to the growth of the species, we construct and analyse a reduced chemostat-like model in which both the growth functions and the apparent dilution rate depend on the density of the species.

Within the framework of the PhD thesis of B. Benyahia, we have included the fouling dynamics of membranes into the AM2 (or AMOCO) model and we have analyzed the resulting model (called the AM2b) [15]. In particular, we have integrated into this model the production and the degradation of Soluble Microbial Products (SMP), which are known to play an important role in the membrane fouling phenomenon. We show that under some general assumptions, the AM2b model has the same number of equilibria as the AM2 model and can exhibit bi-stability. However, under certain operating conditions or if biological parameters values are slightly modified, the AM2b model exhibits equilibria bifurcations and multi-stability properties.

The available anaerobic digestion models used for control purposes do usually only consider soluble matter. In fact, part of the pollutants are not soluble but are under a particulate form. In order to establish whether adding the dynamics of such matter into the models is important for the system behavior or not, we have studied new anaerobic models and established that depending on the kinetics of this additional reaction step, the qualitative behavior of the process may be significantly modified [44].

This year, G. Pimentel as started a PhD co-supervised with the University of Mons, about modeling of the membrane fouling in bioreactors in view of control. The objective is to represent cake formation and air cross-flow as a manipulated variable in the models, in view of future studies of control strategies for improving the efficiency of MBR processes [35], [47], [52], [53].

6.1.3. Ecosystem functioning in heterogeneous environments

Participants: Céline Casenave, Jérôme Harmand, Alain Rapaport.

This year, we have carried out a study of particular spatial interconnections such as “buffered” configurations, and its ecological impacts in terms of setup of a species in environments that are unfavorable when perfectly mixed. We have extended our previous results about the design of configurations for obtaining a global stability [28]. New conditions have been obtained for a species to setup when it is impossible in a perfectly mixed environment. At the opposite, we have characterized configurations that could destabilize bioprocesses.

With UMR Géosciences (Univ. of Rennes 1), we have carried on our analysis of the equivalence of two soil fracture models in terms of transfer functions [19]: the MINC (Multiple INteractiong Continua) and MRMT (Multiple Rate Mass Transfer) models that are quite popular in soil hydrodynamics. We have shown that a strict equivalence can be obtained if one considers different volumes in the discretization of the MINC model. For the moment, this study concerns the transfer of abiotic substances only.

In soil ecosystems, it often happens that several functional groups can be detected to operate concomitantly. We have investigated the mathematical properties of a relatively simple model that has been proposed by the UREP lab (Inra Clermont), that distinguishes explicitly two functional groups of micro-organisms: the decomposers of SOM (soil organic mater) and the producers of SOM, and compared it with a single microbial compartment model in terms of prediction of the so-called “priming effect” [27].

Together with agronomists of the UMR Eco & Sols (Cirad, Inra, IRD, Montpellier SupAgro) and the supervision of the MSc thesis of C. Droin, we have proposed and started to study a new model of consumer/resource for soil microbial ecosystems, in which we explicitly distinguish available and recalcitrant resources [71].

6.2. Stochastic and hybrid models

6.2.1. Stochastic macroscopic models

Participants: Fabien Campillo, Marc Joannides.

We continued our study of stochastic modeling of the chemostat. In a first study we establish the Fokker-Planck equation of the law of the diffusion process. This equation features relevant boundary condition for the washout. We propose specific finite difference schemes to account for this feature [18]. In a second work we adopt the same approach to more accurately study the logistic model [64] which allowed us to propose estimation procedures to take into account the extinction (see Section 6.3.2).

6.2.2. From microscopic models to macroscopic laws

Participants: Fabien Campillo, Coralie Fritsch, Jérôme Harmand, Claude Lobry.

We proposed a chemostat model where the bacterial population is individually-based (IBM), each bacterium is explicitly represented and has a mass evolving continuously over time, and where the substrate concentration is represented as a conventional ordinary differential equation. These two components are coupled with the bacterial consumption. Mechanisms acting on the bacteria are explicitly described (growth, division and uptake). Bacteria interact via consumption. We set the exact Monte Carlo simulation algorithm of this model and its mathematical representation as a stochastic process. We prove the convergence of this process to the solution of an integro-differential equation (IDE) when the population size tends to infinity. The IDE is discretized with the help of finite differences, with simulation as well as the IBM are developed in `Python` with the help of the Gamma-Team (UMR Mistea) [63].

Finally with O. Ovaskainen (Univ. of Helsinki) we developed an evolution model for the chemostat.

6.2.3. Simulation and analysis of hybrid models and the atto-fox problem

Participants: Fabien Campillo, Claude Lobry, Alain Rapaport.

We proposed a new “hybrid” model for the simulation of biofilm growth in a plug flow bioreactor, that combines information from three scales: a microscopic one for the individual bacteria, a mesoscopic or “coarse-grained” one that homogenizes at an intermediate scale the quantities relevant to the attachment/detachment process, and a macroscopic one in terms of substrate concentration. In contrast to existing partial differential equations models, this approach is based on a description of biological mechanisms at the individual scale, thus bringing in a biological justification of the attachment/detachment process responsible for the macroscopic behavior [20].

We pursue our study of the “atto-fox” question in the classical Rosenzweig-MacArthur model for a resource-consumer relationship: for certain values of parameters the system has a limit cycle such that the smallest value reached by the resource on this cycle is so small that the model validity is jeopardized [65].

6.3. Identification and control

6.3.1. Reconstruction methods of kinetics functions

Participant: Alain Rapaport.

A collaboration with Sisyphe Inria project-team has led to the development of a new identification method of the kinetics function in the chemostat model, without any a priori on the monotonicity of the function (thus allowing the consideration of bio-processes that are unstable in open loop) [29]. An extension of this method, that is based on singular perturbations, has been proposed for the extremum seeking problem with only two times scale (instead of three for the usual extremum seeking techniques [75]) [50].

6.3.2. Parameter estimation and particle filtering

Participants: Amine Boutoub, Fabien Campillo, Jérôme Harmand, Marc Joannides.

We consider a stochastic logistic growth model involving both birth and death rates in the drift and diffusion coefficients for which extinction eventually occurs almost surely. We then use the numerical integration of the Fokker-Planck equation presented in Section 6.2.1 to build a likelihood function for the unknown model parameters, when discretely sampled data is available. The existing estimation methods need adaptation in order to deal with the extinction problem. We propose such adaptations, based on the particular form of the Fokker-Planck equation, and we evaluate their performances with numerical simulations [64].

We develop particle approximation methods for the nonlinear filtering and parameter estimation with the help of the chemostat model [70].

6.3.3. Functional assignments methods

Participants: Jérôme Harmand, Alain Rapaport.

Following the philosophy of the work that was achieved within the framework of the former PhD thesis of M. Dumont [3], we have applied part of the proposed methodology for a better understanding of the dynamics of specific species of the anaerobic digestion [30], with Chilean collaborators (see Sections 7.4.1.1 and 7.4.2).

Using a combinatorial approach, we have also developed together with UMR Eco & Sols (Cirad, Inra, IRD, SupAgro – Montpellier) a new method to study the role of the interactions within bacterial species on the performance of an ecosystem. More precisely, based on the specific characteristics of the species of a community and the way they interact between each other, we propose a method to predict the behavior of the ecosystem with respect to its biodiversity [34], [25].

6.3.4. Stabilizing strategies for bioprocesses

Participants: Céline Casenave, Jérôme Harmand, Guilherme Pimentel, Alain Rapaport.

We have carrying on developments of stabilizing strategies for bio-processes, with specific characteristics:

- In [48], it has been shown how the buffered configuration of two chemostat models studied in Section 6.1.3 provides an efficient way to stabilize bioprocesses with inhibition. In the same spirit, its has been shown how the consideration of a “passive” buffer (i.e. without biological reaction) can play the role of a delay and enlarge the attraction basin of stable equilibria [49].
- If often happens in bio-processes that measurements are delayed. In [46], a new stabilizing strategy have been proposed to robustly cope with such delays for single chemostats with inhibition.
- For the stabilization of a series of reactors with multiple inputs, a control strategy based on a linearizing control law coupled with a state observer and an anti windup component has been proposed [66], [67], in view of its application in wine fermentation processes. The originality and difficulty of this multi-inputs problem are due to the inputs constraint that imposes that the manipulated dilution rate of each tank has to be less or equal than the one of the previous tank.

6.3.5. Optimal syntheses for bioprocesses control

Participants: Térence Bayen, Amel Ghouali, Jérôme Harmand, Claude Lobry, Alain Rapaport, Tewfik Sari.

We have continued our activities related to the development of optimal control laws for the optimization of bio-processes, notably in taking advantage of the presence of T. Bayen in the team in 2013. Three kinds of results, depending on the kind of processes under interest, were obtained.

a. Control of batch processes. Sequencing Batch bioReactors can be used to efficiently treat water containing both carbonaceous and nitrogenous pollutants. In such a case, an efficient control that can be used is the oxygen concentration. In such systems, oxic and anoxic bacterial are in competition for certain substances. For a simplified version of this complex situation, we have investigated the optimal strategies in order to minimize the energy to be introduced into the system under performance constraints. The originality of the approach lies in the fact that the original problem is transformed into a very general form. Thus, the optimal control problem is formulated and solved for a very general class of systems of ecological relevance [16].

b. Control of fed-batch processes. References [13], [36], [60] are devoted to the study of a bioreactor which is operated in fed-batch mode. We aim at finding an optimal control in feedback form (i.e. depending of the state) that steers the system in a minimal amount of time to a target (which typically has several interests in wastewater treatment). Finding an optimal control in feedback form is crucial from a practical point of view. In [13], previous works on the subject are extended to the case where the growth function depends on an additional product of the reaction. In the references [36], [60], we provide an optimal control in feedback form whenever mortality and recycling rates are considered, and in the case where the maximum dilution rate is not large enough to compete the growth of the species (in the latter case, this implies that the singular arc is non-necessary controllable implying difficulties in determining optimal controls). References [58], [61] are devoted to the study of optimal control problem governed by a chemostat-type model. In [58], an optimal feedback control law is provided in order to optimize the selection of a species in a chemostat model with one limiting substrate and two species. This brings an interesting issue in order to extend this result to the case where the number of species is larger than 3.

c. Coupled dynamics. References [59], [60] give the results of the study of an optimal control problem of a system coupling a culture of micro-algae limited by light and an anaerobic digester. The mathematical model for the dynamics of the reactors takes into account a periodic day-night model of the light in the culture of micro-algae and a chemostat model for the digester. Our aim is to optimize the production of methane in the digester during a certain number of days with respect to the dilution rate. In [59], some preliminary results on this problem are given for an optimal control problem governed by a one-dimensional Kolmogorov equation. In [60], the full system is analyzed by combining direct methods and indirect methods based on Pontryagin's Principle. In [62], we provide a complete characterization of optimal controls for a minimal time control problem where the system describes a two tanks gradostat model under a cascade inputs constraint. This model allows to create a gradient of resources that is expected to be more realistic to mimic real environment for studying micro-organisms growth.

6.4. Distributed delay systems

Participant: Céline Casenave.

In microbial ecosystems, time delays are often present. For a long time (especially with V. Volterra), distributed delay models (or integro-differential equations) have been proposed to take into account these delays in population models. Some dynamic problems dealing with integro-differential models can be tackled in an original way by using the methodology called "diffusive representation". Some works, which began during the PhD thesis of Céline Casenave, are still under development.

In [26], a new formulation of an integro-differential model of a porous media is proposed, based on this methodology. From this formulation, the dissipative and passive features of the porous wall are established, and numerical simulations are performed. A reduced order model is also proposed which summarizes the boundary behavior of the porous wall ([5]).

This work is done in collaboration with LAAS (Univ. Toulouse III) and the Gipsa-lab (CNRS, Grenoble-INP, Univ. Joseph Fourier, Univ. Stendhal). In the future, these works could be adapted to the case of microbial ecosystems.

6.5. Applications to wastewater treatment

Participants: Térence Bayen, Fabien Campillo, Radhouane Fekih-Salem, Amel Ghouali, Jérôme Harmand, Claude Lobry, Alain Rapaport, Tewfik Sari.

If an important part of our work has been done with the final objective of confronting models to data, the studies realized this year are rather theoretical (cf. research achieved within the framework of PhD theses by B. Benyahia, R. Fekih Salem, S. Hassam and G. Araujo Pimentel). In fact, they can be considered as prerequisites before being applied to real systems which, as for most Anaerobic MBRs, are still often found only at pilot scale and not yet applied on real sites.

Concerning the study of membrane fouling, we collaborate with our colleagues of the University Montpellier 2 within the framework of A. Charfi to characterize membrane fouling [43], [54], [42].

In association with the “Laboratoire d’Automatique de Tlemcen” (Univ. Aboubekr Belkaid) and the Gamma Team (UMR Mistea), Modemic launch the NuWat project (Numerics for water treatment research) in the Lirima network (see Section 7.4.2). The first visit of colleagues from the LAT allowed to make choices on the establishment of teachings for the Master in Tlemcen of general trainings and to define research priorities.

The collaboration with Moise Inria project-team has led to a patent application about an algorithm for “intelligent” pumps for the efficient treatment of large water resources [69]. The method relies on an extension of a former work [4] coupled with faithful simulations of the hydrodynamics of the resource and the pollutant dispersion. This typically applied for Chilean lakes, an application that we plan to launch within a common project with CIRIC Inria-Chile.

6.6. Applications to environmental microbiology

Participants: Céline Casenave, Jérôme Harmand, Alain Rapaport.

We have several ongoing works on the modeling of soil microbial ecosystems. The main characteristics of the models we develop with these partners, compared to aquatic microbial ecosystems, concern the availability of the resources, in terms of:

- spatial distribution and transfer of resources, using simple space representations, with the UMR Eco & Sols (Cirad, Inra, IRD, SupAgro – Montpellier) and the UMR Géosciences (Rennes). See Section 6.1.3 and references [28], [19];
- consideration of recalcitrant forms and recycling of nutrients, with UMR Eco & Sols (Cirad, Inra, IRD, SupAgro – Montpellier) and Inra UREP (Clermont). See Section 6.1.3 and [27], [71].

We have also proposed, together with the UMR Eco & Sols (Cirad, Inra, IRD, SupAgro – Montpellier), a new methodology to deduce from the observation of the performances of several assembling of reconstituted ecosystems, the number and the nature of species interactions (see Section 6.3.3 and [34], [25]).

The organization of a “research school” dedicated to the biologists of the marine research station of Banyuls (see Section 8.1) has led to a primary work about variable yield in marine microbial populations.

6.7. Applications to wine fermentation

Participants: Térence Bayen, Céline Casenave, Jérôme Harmand, Alain Rapaport.

---

[5] Two journal articles dealing with the identification of integro-differential models, and the controllability of some SISO Volterra models are still under revision.

We study the problem of the control of a Multi-State Continuous Fermentor (MSCF) composed of 4 reactors connected in series (the experimental pilot plant is located at Montpellier, in the UMR SPO (INRA, Montpellier SupAgro, Univ. Montpellier 1)). The goal is to control the sugar concentration of the four reactors with, as control inputs, the input flow rates of the four reactors. The originality of the problem comes from the cascade structure of the device which leads to a constraint on the control inputs. Two control strategies have been studied.

An output stabilizing control strategy. The linearizing control law proposed in [66], [67] (see Section 6.3.4) has been validated on numerical simulations, and then has been implemented (Labview-Matlab interface) on the experimental process. The obtained results are convincing; others experiments are scheduled in 2014 to refine the control law.

A minimal time state feedback strategy. The optimal state feedback studied in [62] (see Section 6.3.5) is of completely different nature, as it relies on bang-bang controls and singular arcs. We plan to couple this control law when far from the target with the previous stabilizing law when close from the target, in order to provide a practical sub-optimal strategy.

The first part of this work was conducted as a part of the European CAFE project (Computer-Aided Food processes for control Engineering) described in Section 7.3.1 , in collaboration with CESAME (Univ. Catholique de Louvain-la-neuve), and UMR SPO.

A new project, see Section 7.2 , in which the UMR SPO and the Unit Mistea are involved has begun in 2013. Preliminary work has been performed during the MSc thesis of S. Sekkat [74] about the modeling of the fermentation with addition of nitrogen in the MSCF.

6.8. Applications to micro-algae

Participants: Térence Bayen, Matthieu Sebbah.

An originality developed within the Biocore Inria project-team is to couple a bioreactor that cultivate micro-algae with an anaerobic digester, that uses micro-algae as a substrate that is then converted into valuable bio-gaz (methane). As micro-algae are micro-organisms whose growth is limited by light, one has to take into account periodic day-night model of the light. In [59], [61], control laws that maximize the biogaz production in this periodic framework have been proposed (see Section 6.3.5). In the framework of the Inria Project Lab “Algae in Silico” (see Section 7.2) and the Inria-CIRIC Center in Chile (see Section 7.4), several extensions and collaborations with Biocore Inris project-team are scheduled for the coming year.

6.9. Other results

This section contains some theoretical as well as applied results, that are not directly connected to the main field of the team.

Theoretical ecology
Participant: Tewfik Sari.

In [24], ecological trade-offs between species are studied to explain species coexistence in ecological communities. In our model, plant species compete for sites where each site has a fixed stress condition. Species differ both in stress tolerance and competitive ability. We derive the deterministic discrete-time dynamical system for the species abundances. We prove the conditions under which plant species can coexist in a stable equilibrium. We compare our model with a recently proposed, continuous-time dynamical system for a tolerance-fecundity trade-off in plant communities, and we show that this model is a special case of the continuous-time version of our model.

Calculus of variations
Participant: Térence Bayen.

The work [37] is devoted to the study of necessary and sufficient optimality conditions for weak and strong minima for optimal control problems governed by semi-linear parabolic equations; whereas in the field of calculus of variation, these conditions (such as Euler-Lagrange equation, Legendre's condition, Weierstrass's condition) have been deeply investigated, the study of strong solutions for optimal control problems of partial differential equations is new.

Ice cream crystallization
Participant: Céline Casenave.

We study the problem of the control of an ice cream crystallization process (the experimental pilot plant is located at Irstea Antony). The goal is to control the viscosity of the ice cream at the outlet of the continuous crystallizer. The problem has been studied in two steps.

Modeling, model reduction and parameter identification. On the basis of a population balance equation describing the evolution of the crystal size distribution (CSD) of the ice cream, and an energy balance equation, we have proposed an input-output reduced order model of the process, which is based on physical assumptions. The parameters of the model have been identified and the model has been validated from experimental data [68].

Design of the control law. Based on the reduced order model, a nonlinear control strategy based on an adaptive linearizing control law coupled with a Smith predictor to account for the measurement delay has been proposed [41], [66], [67]. The control has been implemented (Labview-Matlab interface) and then validated on the experimental pilot plant. During the industrial conference which has been organized in February 2013 by the European CAFE project (see Section 7.3.1), and which representatives of several industries in food processing attended, a live demonstration of the designed control law has been performed.

This work was conducted as a part of the European CAFE project, described in Section 7.3.1 , in collaboration with CESAME (Univ. Catholique de Louvain-la-neuve), Irstea Antony and AgroParisTech.

Semi-Markov land use dynamic
Participant: Fabien Campillo, Angelo Raherinirina.

We pursued the development of semi-Markov model for the inference of land use dynamic from data proposed by IRD. The thesis of A. Raherinirina was defended in August 2013 [12]. Later in the year during the stay of A. Raherinirina in Montpellier we completed an article accepted by the journal ARIMA and that will be published in 2014 [17].

NUMED Project-Team (section vide)

REO Project-Team

6. New Results

6.1. Mathematical and numerical analysis of fluid-structure interaction problems

Participants: Muriel Boulakia, Miguel Ángel Fernández Varela, Jean-Frédéric Gerbeau, Céline Grandmont, Mikel Landajuela Larma, Jimmy Mullaert, Marina Vidrascu.

- In [17] we analyze the performances of two types of Luenberger observers – namely, the so-called Direct Velocity Feedback and Schur Displacement Feedback procedures, originally devised for elasto-dynamics – to estimate the state of a fluid-structure interaction model for hemodynamics, when the measurements are assumed to be restricted to displacements or velocities in the solid. We first assess the observers using hemodynamics-inspired test problems with the complete model, including the Navier-Stokes equations in Arbitrary Lagrangian-Eulerian formulation, in particular. Then, in order to obtain more detailed insight we consider several well-chosen simplified models, each of which allowing a thorough analysis – emphasizing spectral considerations – while illustrating a major phenomenon of interest for the observer performance, namely, the added mass effect for the structure, the coupling with a lumped-parameter boundary condition model for the fluid flow, and the fluid dynamics effect per se. Whereas improvements can be sought for when additional measurements are available in the fluid domain, the present framework this establishes Luenberger observer methods as very attractive strategies – compared, e.g., to classical variational techniques – to perform state estimation, and more generally for uncertainty estimation since other observer procedures can be conveniently combined to estimate uncertain parameters.
- In [28] we introduce a class of incremental displacement-correction schemes for the explicit coupling of a thin-structure with an incompressible fluid. We provide a general stability and convergence analysis that covers both the incremental and the non-incremental variants. Their stability properties are independent of the added-mass effect. The superior accuracy of the incremental schemes (with respect to the original non-incremental variant) is highlighted by the error estimates, and then confirmed in a benchmark by numerical experiments.
- In [29], [62] we introduce a class of fully decoupled time-marching schemes (velocity-pressure-displacement splitting) for the coupling of an incompressible fluid with a thin-walled viscoelastic structure. A priori energy estimates guaranteeing unconditional stability are established for the variants without extrapolation and with first-order extrapolation. The accuracy and performance of the methods proposed are discussed in several numerical examples.
- In [30] we introduce a class of explicit coupling schemes for the numerical solution of fluid-structure interaction problems involving a viscous incompressible fluid and a general thin-walled structure (e.g., including damping and non-linear behavior). The fundamental ingredient in these methods is the (parameter free) explicit Robin interface condition for the fluid, which enables the fluid-solid splitting through appropriate extrapolations of the solid velocity and fluid stress on the interface. The resulting solution procedures are genuinely partitioned. Stability and error estimates are provided for all the variants (depending on the extrapolations), using energy arguments within a representative linear setting. In particular, we show that one of them yields added-mass free unconditional stability and optimal (firs-order) time accuracy. A comprehensive numerical study, involving different examples from the literature, supports the theory.
- In [63] we introduce a new class of explicit coupling schemes for the numerical solution of fluid-structure interaction problems involving a viscous incompressible fluid and an elastic structure. These methods generalize the arguments reported in [28], [30] to the case of the coupling with thick-walled structures. The basic idea lies in the derivation of an intrinsic interface Robin consistency at

the space semi-discrete level, using a lumped-mass approximation in the structure. The fluid-solid splitting is then performed through appropriate extrapolations of the solid velocity and stress on the interface. Based on these methods, a new, parameter-free, Robin-Neumann iterative procedure is also proposed for the partitioned solution of implicit coupling. A priori energy estimates, guaranteeing the (added-mass free) stability of the schemes and the convergence of the iterative procedure, are established. The accuracy and robustness of the methods are illustrated in several numerical examples.

- In [22] we discuss explicit coupling schemes for fluid-structure interaction problems where the added mass effect is important. We show the close relation between coupling schemes using Nitsche's method and a Robin-Robin type coupling. In the latter case the method may be implemented either using boundary integrals of the stresses or the more conventional discrete lifting operators. We also make the observation that these scheme are stable under a hyperbolic type CFL condition, but that optimal accuracy imposes a parabolic type CFL conditions due to the splitting error. Two strategies to enhance the accuracy of the coupling scheme under the hyperbolic CFL-condition are suggested, one using extrapolation and defect-correction and one using a penalty-free non-symmetric Nitsche method. Finally we illustrate the performance of the proposed schemes on some numerical examples in two and three space dimensions.
- In [60] we consider the extension of the Nitsche method to the case of fluid-structure interaction problems on unfitted meshes. We give a stability analysis for the space semi-discretized problem and show how this estimate may be used to derive optimal error estimates for smooth solutions, irrespective of the mesh/interface intersection. Some numerical examples illustrate the theoretical discussion.
- In [21] we are interested by the three-dimensional coupling between an incompressible fluid and a rigid body. The fluid is modeled by the Navier-Stokes equations, while the solid satisfies the Newton's laws. In the main result of the paper we prove that, with the help of a distributed control, we can drive the fluid and structure velocities to zero and the solid to a reference position provided that the initial velocities are small enough and the initial position of the structure is close to the reference position. This is done without any condition on the geometry of the rigid body.
- In the book chapter [57] we deal with some specific existence and numerical results applied to a 2D/1D fluid-structure coupled model, for an incompressible fluid and a thin elastic structure. We underline some of the mathematical and numerical difficulties that one may face when studying this kind of problems such as the geometrical nonlinearities or the added mass effect. In particular we underline the link between the strategies of proof of weak or strong solutions and the possible algorithms to discretize these types of coupled problems.

6.2. Numerical methods for fluid mechanics and application to blood flows

Participants: Grégory Arbia, Benoit Fabrèges, Jean-Frédéric Gerbeau, Sanjay Pant, Saverio Smaldone, Marc Thiriet, Irène Vignon-Clementel.

- In [61] we propose a new approach to the loosely coupled time-marching of a fluid-fluid interaction problems involving the incompressible Navier-Stokes equations. The methods combine a specific explicit Robin-Robin treatment of the interface coupling with a weakly consistent interface pressure stabilization in time. A priori energy estimates guaranteeing stability of the splitting are obtained for a total pressure formulation of the coupled problem. The performance of the proposed schemes is illustrated on several numerical experiments related to simulation of aortic blood flow.
- In [49] we present our strategy to meet the MICCAI Challenge 2013: the goal was to recover a measured (but unrevealed) pressure drop across a coarctation of the aorta through 3D simulation. A filtering-based strategy is devised to perform parameter estimation and subsequent multiscale CFD simulations of arterial blood flow. The method is applied to the patient-specific case in the two physiological states of rest and stress. In both cases, the method is shown to be effective in closely matching the available clinically measured data. Pressure drop across the coarctation is predicted

for both states. At the time of [47], these measurements were available: the computed pressure drop across the coarctation for the stress case appears to be very close to the measured one, while the one for the rest case is not as good. One should note that no participant of the challenge managed to recover the measured pressure drop for the rest case.

- In [35], we aim to reduce the complexity of patient-specific simulations by combining image analysis, computational fluid dynamics and model order reduction techniques. The proposed method makes use of a reference geometry estimated as an average of the population, within an efficient statistical framework based on the currents representation of shapes. Snapshots of blood flow simulations performed in the reference geometry are used to build a POD (Proper Orthogonal Decomposition) basis, which can then be mapped on new patients to perform reduced order blood flow simulations with patient specific boundary conditions. This approach is applied to a data-set of 17 tetralogy of Fallot patients to simulate blood flow through the pulmonary artery under normal (healthy or synthetic valves with almost no backflow) and pathological (leaky or absent valve with backflow) conditions to better understand the impact of regurgitated blood on pressure and velocity at the outflow tracts. The model reduction approach is further tested by performing patient simulations under exercise and varying degrees of pathophysiological conditions based on reduction of reference solutions (rest and medium backflow conditions respectively).
- In [16], we analyze two 3D-0D coupling approaches in which a fractional-step projection scheme is used in the fluid. Our analysis shows that explicit approaches might yield numerical instabilities, particularly in the case of realistic geometries with multiple outlets. We introduce and analyze an implicitly 3D-0D coupled formulation with enhanced stability properties and which requires a negligible additional computational cost. Furthermore, we also address the extension of these methods to fluid-structure interaction problems. The theoretical stability results are confirmed by meaningful numerical experiments in patient specific geometries coming from medical imaging.
- In [36], we developed two multi-scale models, each including the 3D model of the surgical junction constructed from MRI, and a closed-loop LPN derived from pre-operative data obtained from two patients prior to Stage 2 Fontan palliation of single ventricle congenital heart disease. "Virtual" surgeries were performed and a corresponding multi-scale simulation predicted the patient's post-operative hemodynamic conditions, tested under different physiological conditions. The impact of the surgical junction geometry on the global circulation was contrasted with variations of key physiological parameters.
- A novel Y-shaped baffle was proposed for the Stage 3 Fontan operation achieving overall superior hemodynamic performance compared with traditional designs. Previously, we investigated if and how the inferior vena cava flow (which contains an important biological hepatic factor) could be best distributed among both lungs. In [43] we proposed a multi-step method for patient-specific optimization of such surgeries to study the effects of boundary conditions and geometry on hepatic factor distribution (HFD). The resulting optimal Y-graft geometry largely depended on the patient left/right pulmonary flow split. Unequal branch size and constrained optimization on energy efficiency were explored. Two patient-specific examples showed that optimization-derived Y-grafts effectively improved HFD.
- The use of elaborate closed-loop lumped parameter network (LPN) models of the heart and the circulatory system as boundary conditions for 3D simulations can provide valuable global dynamic information, particularly for patient specific simulations. In [27], we have developed and tested a numerical method to couple a 3D Navier-Stokes finite-element formulation and a reduced model of the rest of the circulation, keeping the coupling robust but modular. For Neumann boundaries, implicit, semi-implicit, and explicit quasi-Newton formulations are compared within the time-implicit coupling scheme. The requirements for coupling Dirichlet boundary conditions are also discussed and compared to that of the Neumann coupled boundaries. Both these works were key for applications where blood flows in different directions during the cardiac cycle and where coupling with the rest of the circulation is instrumental (see the shunt optimization application [75]).

- In [26] we propose the first patient-specific predictive modeling of stage 2 palliation for congenital heart disease by using virtual surgery and closed-loop multi-scale modeling. We present a workflow to perform post-operative simulations from pre-operative clinical data. Two surgical options (bi-directional Glenn and hemi-Fontan operations) are virtually performed and coupled to the pre-operative LPM, with the hemodynamics of both options reported. Results are validated against postoperative clinical data.
- In [14] we study the influence of solvers and test case setup on the result of numerical simulations. The need for detailed construction of the numerical model depends on the precision needed to answer the biomedical question at hand and should be assessed for each problem on a combination of clinically relevant patient-specific geometry and physiological conditions. Methods and results between a commercial code and an in-house research code are illustrated on three congenital heart disease examples of increasing complexity. This publication is designed as a tool to better communicate with clinical researchers interested in simulations.

6.3. Numerical methods for cardiac electrophysiology

Participants: Muriel Boulakia, Jean-Frédéric Gerbeau, Annabelle Collin, Elisa Schenone.

- In [25], Computational electrophysiology is a very active field with tremendous potential in medical applications, albeit leads to highly intensive simulations. We here propose a surface-based electrophysiology formulation, motivated by the modeling of thin structures such as cardiac atria, which greatly reduces the size of the computational models. Moreover, our model is specifically devised to retain the key features associated with the anisotropy in the diffusion effects induced by the fiber architecture, with rapid variations across the thickness which cannot be adequately represented by naive averaging strategies. Our proposed model relies on a detailed asymptotic analysis in which we identify a limit model and establish strong convergence results. We also provide detailed numerical assessments which confirm an excellent accuracy of the surface-based model – compared with the reference 3D model – including in the representation of a complex phenomenon, namely, spiral waves.
- In [45] we assess a previously-proposed surface-based electrophysiology model with detailed atrial simulations. This model - derived and substantiated by mathematical arguments - is specifically designed to address thin structures such as atria, and to take into account strong anisotropy effects related to fiber directions with possibly rapid variations across the wall thickness. The simulation results are in excellent adequacy with previous studies, and confirm the importance of anisotropy effects and variations thereof. Furthermore, this surface-based model provides dramatic computational benefits over 3D models with preserved accuracy.

6.4. Lung and respiration modeling

Participants: Grégory Arbia, Laurent Boudin, Muriel Boulakia, Benoit Fabrèges, Miguel Ángel Fernández Varela, Jean-Frédéric Gerbeau, Céline Grandmont, Stéphane Liwarek, Jessica Oakes, Ayman Moussa, Irène Vignon-Clementel.

- In [66] we are interested in the mathematical modeling of the deformation of the human lung tissue, called the lung parenchyma, during the respiration process. The parenchyma is a foam-like elastic material containing millions of air-filled alveoli connected by a tree- shaped network of airways. In this study, the parenchyma is governed by the linearized elasticity equations and the air movement in the tree by the Poiseuille law in each airway. The geometric arrangement of the alveoli is assumed to be periodic with a small period. We use the two-scale convergence theory to study the asymptotic behavior as the period goes to zero. The effect of the network of airways is described by a nonlocal operator and we propose a simple geometrical setting for which we show that this operator converges. We identify in the limit the equations modeling the homogenized behavior under an abstract convergence condition on this nonlocal operator. We derive some mechanical properties of the limit material by studying the homogenized equations: the limit model is nonlocal both in space and time if the parenchyma material is considered compressible, but only in space if it is incompressible. Finally, we propose a numerical method to solve the homogenized equations and we study numerically a few properties of the homogenized parenchyma model.

- In [31] we present a calculation of the functioning of an acinus at exercise. We show that, given the geometry and the breathing dynamics of real acini, respiration can be correlated to a single equivalent parameter that we call the integrative permeability. We find that both V_{O2max} and PAO2 depend on this permeability in a non-linear manner.
- In [19], In this paper, we consider the Stokes equations and we are concerned with the inverse problem of identifying a Robin coefficient on some non accessible part of the boundary from available data on the other part of the boundary. We first study the identifiability of the Robin coefficient and then we establish a stability estimate of logarithm type thanks to a Carleman inequality due to A. L. Bukhgeim and under the assumption that the velocity of a given reference solution stays far from 0 on a part of the boundary where Robin conditions are prescribed.
- In [18],In this work, we investigate the asymptotic behaviour of the solutions to the non-reactive fully elastic Boltzmann equations for mixtures in the diffusive scaling. We deal with cross sections such as hard spheres or cut-off power law potentials. We use Hilbert expansions near the common thermodynamic equilibrium granted by the H-theorem. The lower-order non trivial equality obtained from the Boltzmann equations leads to a linear functional equation in the velocity variable which is solved thanks to the Fredholm alternative. Since we consider multicomponent mixtures, the classical techniques introduced by Grad cannot be applied, and we propose a new method to treat the terms involving particles with different masses. The next-order equality in the Hilbert expansion then allows to write the macroscopic continuity equations for each component of the mixture.
- In [39], In this paper we introduce a PDE system which aims at describing the dynamics of a dispersed phase of particles moving into an incompressible perfect fluid, in two space dimensions. The system couples a Vlasov-type equation and an Euler-type equation: the fluid acts on the dispersed phase through a gyroscopic force whereas the latter contributes to the vorticity of the former. First we give a Dobrushin type derivation of the system as a mean-field limit of a PDE system which describes the dynamics of a finite number of massive pointwise particles moving into an incompressible perfect fluid. This last system is itself inferred from a joint work of the second author with O. Glass and C. Lacave, where the system for one massive pointwise particle was derived as the limit of the motion of a solid body when the body shrinks to a point with fixed mass and circulation. Then we deal with the well-posedness issues including the existence of weak solutions. Next we exhibit the Hamiltonian structure of the system and finally, we study the behavior of the system in the limit where the mass of the particles vanishes.
- In [40] we solved for the airflow and aerosol particle distribution in healthy and emphysematous rat lungs. Following our preliminary work in [79], we first estimated the respiratory resistance and compliance parameters from pressure measurements taken during ventilation experiments performed in healthy and emphysematous rats. Next, the 3D Navier-Stokes equations were solved in a Magnetic Resonance derived airway geometry coupled to 0D models at the boundaries leading to the five rat lobes. The multiscale 3D-0D simulations enabled consistent pressure and airflow results, unlike what was found when a constant pressure was described at the boundaries. Aerosolized particles were tracked throughout inspiration and the effects of particle size and gravity were studied. Healthy, homogeneous and heterogeneous disease cases were assessed. Once available, these in-silico predictions may be compared to experimental deposition data.

6.5. Miscellaneous

Participants: Grégory Arbia, Laurent Boudin, Jean-Frédéric Gerbeau, Damiano Lombardi, Marina Vidrascu, Irène Vignon-Clementel.

- In [13] we analyse the solution of the linear advection equation on a uniform mesh by a non dissipative second order scheme for discontinuous initial condition. We focus on the case of advection of a step function by the leapfrog scheme. We derive closed form exact and approximate solutions for the scheme that accurately predict oscillations of the numerical scheme.
- In [41] The recent biomechanical theory of cancer growth considers solid tumors as liquid-like materials comprising elastic components. In this fluid mechanical view, the expansion ability of a solid tumor into a host tissue is mainly driven by either the cell diffusion constant or the cell division rate, the latter depending either on the local cell density (contact inhibition), on mechanical stress in the tumor, or both. For the two by two degenerate parabolic/elliptic reaction-diffusion system that results from this modeling, we prove there are always traveling waves above a minimal speed and we analyse their shapes. They appear to be complex with composite shapes and discontinuities. Several small parameters allow for analytical solutions; in particular the incompressible cells limit is very singular and related to the Hele-Shaw equation. These singular traveling waves are recovered numerically.
- In [68] This paper is devoted to the use of the entropy and duality methods for the existence theory of reaction-cross diffusion systems consisting of two equations, in any dimension of space. Those systems appear in population dynamics when the diffusion rates of individuals of two species depend on the concentration of individuals of the same species (self-diffusion), or of the other species (cross diffusion).
- In [65] We consider in this paper a spray constituted of an incompressible viscous gas and of small droplets which can breakup. This spray is modeled by the coupling (through a drag force term) of the incom- pressible Navier-Stokes equation and of the Vlasov-Boltzmann equation, together with a fragmentation kernel. We first show at the formal level that if the droplets are very small after the breakup, then the solutions of this system converge towards the solution of a simplified system in which the small droplets produced by the breakup are treated as part of the fluid. Then, existence of global weak solutions for this last system is shown to hold, thanks to the use of the DiPerna-Lions theory for singular transport equations.
- In [42],We propose a method of modelling sail type structures which captures the wrinkling behaviour of such structures. The method is validated through experimental and analytical test cases, particularly in terms of wrinkling prediction. An enhanced wrinkling index is proposed as a valuable measure characterizing the global wrinkling development on the deformed structure. The method is based on a pseudo-dynamic finite element procedure involving non-linear MITC shell elements. The major advantage compared to membrane models generally used for this type of analysis is that no ad hoc wrinkling model is required to control the stability of the structure. We demonstrate our approach to analyse the behaviour of various structures with spherical and cylindrical shapes, characteristic of downwind sails over a rather wide range of shape and constitutive parameters. In all cases convergence is reached and the overall flying shape is most adequately represented, which shows that our approach is a most valuable alternative to standard techniques to provide deeper insight into the physical behaviour. Limitations appear only in some very special instances in which local wrinkling-related instabilities are extremely high and would require specific additional treatments, out of the scope of the present study.
- In [34],Since the pioneering work by Treloar, many models based on polymer chain statistics have been proposed to describe rubber elasticity. Recently, Alicandro, Cicalese, and the first author have rigorously derived a continuum theory of rubber elasticity from a discrete model by variational convergence. The aim of this paper is twofold. First we further physically motivate this model, and complete the analysis by numerical simulations. Second, in order to compare this model to the literature, we present in a common language two other representative types of models, specify their underlying assumptions, check their mathematical properties, and compare them to Treloar's experiments.
- In [48] We apply the domain decomposition method to linear elasticity problems for multi-materials

where the heterogeneities are concentrated in a thin internal layer. In the first case the heterogeneities are small, identical and periodically distributed on an internal surface and in the second one all the thin, curved internal layer is made of an elastic material much more strong than the surrounding one. In the first case the domain decomposition is used to efficiently solve the non-standard transmission problems obtained by the asymptotic expansion method. In the second case a non-standard membrane transmission problem originates from a surface shell like energy.

- In [32] : Our aim is to numerically validate the effectiveness of a matched asymptotic expansion formal method introduced in a pioneering paper by Nguetseng and Sànchez Palencia and extended in [76], [33]. Using this method a simplified model for the influence of small identical heterogeneities periodically distributed on an internal surface to the overall response of a linearly elastic body is derived. In order to validate this formal method a careful numerical study compares the solution obtained by a standard method on a fine mesh to the one obtained by asymptotic expansion. We compute both the zero and the first order terms in the expansion. To efficiently compute the first order term we introduce a suitable domain decomposition method.
- In [39] we introduce a PDE system which aims at describing the dynamics of a dispersed phase of particles moving into an incompressible perfect fluid, in two space dimensions. The system couples a Vlasov-type equation and an Euler-type equation: the fluid acts on the dispersed phase through a gyroscopic force whereas the latter contributes to the vorticity of the former. First we give a Dobrushin type derivation of the system as a mean-field limit of a PDE system which describes the dynamics of a finite number of massive pointwise particles moving into an incompressible perfect fluid. This last system is itself inferred from a joint work of the second author with O. Glass and C. Lacave, where the system for one massive pointwise particle was derived as the limit of the motion of a solid body when the body shrinks to a point with fixed mass and circulation. Then we deal with the well-posedness issues including the existence of weak solutions. Next we exhibit the Hamiltonian structure of the system and finally, we study the behavior of the system in the limit where the mass of the particles vanishes.
- In [65] we consider a spray constituted of an incompressible viscous gas and of small droplets which can breakup. This spray is modeled by the coupling (through a drag force term) of the incom- pressible Navier-Stokes equation and of the Vlasov-Boltzmann equation, together with a fragmentation kernel. We first show at the formal level that if the droplets are very small after the breakup, then the solutions of this system converge towards the solution of a simplified system in which the small droplets produced by the breakup are treated as part of the fluid. Then, existence of global weak solutions for this last system is shown to hold, thanks to the use of the DiPerna-Lions theory for singular transport equations.

SISYPHE Project-Team

6. New Results

6.1. Neuroscience & Neuroendocrinology: Regulation of the Gonadotrope axis

6.1.1. A numerical method for transport equations with discontinuous flux functions: application to mathematical modeling of cell dynamics

Participants: Benjamin Aymard, Frédérique Clément, Frédéric Coquel, Marie Postel.

We have proposed a numerical method to handle discontinuous fluxes arising in transport-like equations [35]. More precisely, we have studied hyperbolic PDEs with flux transmission conditions at interfaces between subdomains where coefficients are discontinuous. A dedicated finite volume scheme with a limited high order enhancement has been adapted to treat the discontinuities arising at interfaces. The validation of the method has been done on one- and two-dimensional toy problems for which exact solutions are available, allowing us to do a thorough convergence study. We have then applied the method to a biological model focusing on complex cell dynamics [40] that initially motivated this study and illustrates the full potentialities of the scheme.

6.1.2. Adaptive mesh refinement strategy for a nonconservative transport problem

Participants: Benjamin Aymard, Frédérique Clément, Marie Postel.

In the framework of transport equations it is usual to need long time simulations, and therefore large physical domains to cover a phenomenon. On the other hand it can happen that only a small time varying portion of the domain is interesting. This motivates the use of adaptivity for the spatial discretization. Biological models involving cell development are often nonconservative to account for cell division. In that case the threshold controlling the spatial adaptivity may have to be time-dependent in order to keep up with the progression of the solution. We have tackled the difficulties arising when applying a multiresolution method to a transport equation with discontinuous fluxes modeling localized mitosis [76]. The analysis of the numerical method is performed on a simplified model and numerical scheme. An original threshold strategy is proposed and validated thanks to extensive numerical tests. It is then applied to a biological model in both cases of distributed and localized mitosis.

6.1.3. Coupled Somatic Cell Kinetics and Germ Cell Growth: Multiscale Model-Based Insight on Ovarian Follicular Development

Participants: Frédérique Clément, Philippe Michel, Danielle Monniaux, Thomas Stiehl.

We have designed a stochastic individual-based model describing the first stages of follicular development, where the cell population is structured with respect to age (progression within the cell cycle) and space (radial distance from the oocyte) [39]. The model accounts for the molecular dialogue existing between the oocyte and granulosa cells. Three dynamically interacting scales are considered in the model: (i) a microscopic, local scale corresponding to an individual cell embedded in its immediate environment, (ii) a mesoscopic, semi-local scale corresponding to anatomical or functional areas of follicles and (iii) a macroscopic, global scale corresponding to the morphology of the follicle. Numerical simulations were performed to reproduce the 3D morphogenesis of follicles and follow simultaneously the detailed spatial distribution of individual granulosa cells, their organization as concentric layers or functional cell clones and the increase in the follicle size. Detailed quantitative simulation results have been provided in the ovine species, in which well characterized genetic mutations lead to a variety of phenotypic follicle morphogenesis. The model can help to explain pathological situations of imbalance between oocyte growth and follicular cell proliferation.

6.1.4. Innovative computational and theoretical tools for slow-fast dynamics

Participants: Mathieu Desroches, Maciej Krupa.

Mixed-Mode Bursting Oscillations: Dynamics created by a slow passage through spike-adding canard explosion in a square-wave burster [44]. This work concerns the phenomenon of Mixed-Mode Bursting Oscillations (MMBOs). These are solutions of fast-slow systems of ordinary differential equations that exhibit both small-amplitude oscillations (SAOs) and bursts consisting of one or multiple large-amplitude oscillations (LAOs). The name MMBO is given in analogy to Mixed-Mode Oscillations, which consist of alternating SAOs and LAOs, without the LAOs being organized into burst events. In this article, we show how MMBOs are created naturally in systems that have a spike-adding bifurcation or spike-adding mechanism, and in which the dynamics of one (or more) of the slow variables causes the system to pass slowly through that bifurcation. Canards are central to the dynamics of MMBOs, and their role in shaping the MMBOs is two-fold: saddle-type canards are involved in the spike-adding mechanism of the underlying burster and permit one to understand the number of LAOs in each burst event, and folded-node canards arise due to the slow passage effect and control the number of SAOs. The analysis is carried out for a prototypical fourth-order system of this type, which consists of the third-order Hindmarsh-Rose system, known to have the spike-adding mechanism, and in which one of the key bifurcation parameters also varies slowly. We also include a discussion of the MMBO phenomenon for the Morris-Lecar-Terman system. Finally, we discuss the role of the MMBOs to a biological modeling of secreting neurons.

Canards in piecewise-linear systems: explosions and super-explosions [43]. We show that a planar slow-fast piecewise-linear (PWL) system with three zones admits limit cycles that share a lot of similarity with van der Pol canards, in particular an explosive growth. Using phase-space compactification, we show that these quasi-canard cycles are strongly related to a bifurcation at infinity. Furthermore, we investigate a limiting case in which we show the existence of a continuum of canard homoclinic connections that coexist for a single-parameter value and with amplitude ranging from an order of ε to an order of 1, a phenomenon truly associated with the non-smooth character of this system and which we call super-explosion.

Some results have been obtained concerning numerical continuation techniques for planar slow-fast systems [42] and short-term synaptic plasticity in the deterministic Tsodyks-Markram model that leads to unpredictable network dynamics [41].

6.2. Quantum engineering: controlled quantum systems

Participants: Joachim Cohen, Loïc Herviou, Mazyar Mirrahimi, Pierre Rouchon, Pierre Six.

6.2.1. Schrödinger cat states and hardware efficient quantum error correction

We introduce a new gate that transfers an arbitrary state of a qubit into a superposition of two quasi-orthogonal coherent states of a cavity mode, with opposite phases [52]. Such a highly non-classical state is often called a Schrödinger cat state. This qcMAP gate is based on conditional qubit and cavity operations exploiting the energy level dispersive shifts, in the regime where they are much stronger than the cavity and qubit linewidths. The generation of multi-component superpositions of quasi-orthogonal coherent states, non-local entangled states of two resonators and multi-qubit GHz states can be efficiently achieved by this gate.

In a second contribution [53], we propose to use an encoding of a quantum bit of information in a four-component Schrödinger cat state to ensure its protection against the photon loss, being the major source of decoherence for such a quantum harmonic oscillator. This protection is ensured by an efficient quantum error correction scheme employing the nonlinearity provided by a single physical qubit coupled to the cavity. We describe in detail how to implement these operations in a circuit quantum electrodynamics system. This directly addresses the task of building a hardware-efficient quantum memory and can lead to important shortcuts in quantum computing architectures.

As an important step towards the realization of such a protected quantum memory, in a collaboration with the team of Robert J. Schoelkopf at Yale university, we have successfully realized the encoding protocol of [52] using a 3D transom qubit coupled to a waveguide cavity resonator with a highly ideal off-resonant coupling [60]. This dispersive interaction is much greater than decoherence rates and higher-order nonlinearities to allow simultaneous manipulation of hundreds of photons. We created cat states as large as 111 photons and created superpositions of up to four coherent states. This control creates a powerful interface between discrete and continuous variable quantum computation and could enable applications in metrology and quantum information processing. This important achievement was published in Science and was also highlighted in Science Perspectives [103].

6.2.2. Quantum reservoir (dissipation) engineering

We have studied the application of dissipation engineering techniques to perform a high-performance and fast qubit reset [64]. Qubit reset is crucial at the start of and during quantum information algorithms. In a collaboration with the team of Michel H. Devoret at Yale university, our protocol, nicknamed DDROP (Double Drive Reset of Population) was experimentally tested on a superconducting transmon qubit and achieves a ground state preparation of at least 99.5% in times less than $3\mu s$; faster and higher fidelity are predicted upon parameter optimization [46].

Next, we proposed a dissipation engineering scheme that prepares and protects a maximally entangled state of a pair of superconducting qubits [54]. This is done by off-resonantly coupling the two qubits to a low-Q cavity mode playing the role of a dissipative reservoir. We engineer this coupling by applying six continuous-wave microwave drives with appropriate frequencies. The two qubits need not be identical. We show that our approach does not require any fine-tuning of the parameters and requires only that certain ratios between them be large. This protocol was experimentally realized in a collaboration with the team of M. H. Devoret at Yale university [57]. Unlike conventional, measurement-based schemes, this autonomous approach uses engineered dissipation to counteract decoherence, obviating the need for a complicated external feedback loop to correct errors. Instead, the feedback loop is built into the Hamiltonian such that the steady state of the system in the presence of drives and dissipation is a Bell state, an essential building block for quantum information processing. Such autonomous schemes, which are broadly applicable to a variety of physical systems, will be an essential tool for the implementation of quantum error correction. This important result appeared in Nature back-to-back to another paper by the group of D.J. Wineland (2012 Nobel prize winner) at NIST implementing similar ideas on another physical system consisting of trapped ion qubits [105].

6.2.3. Quantum measurement and measurement-based feedback

Measuring a quantum system can randomly perturb its state. The strength and nature of this back-action depend on the quantity that is measured. In a partial measurement performed by an ideal apparatus, quantum physics predicts that the system remains in a pure state whose evolution can be tracked perfectly from the measurement record. This property was proved in a collaboration with the group of Michel H. Devoret (Yale university) using a superconducting qubit dispersively coupled to a cavity traversed by a microwave signal [47]. The back-action on the qubit state of a single measurement of both signal quadratures was observed and shown to produce a stochastic operation whose action is determined by the measurement result. This accurate monitoring of a qubit state is an essential prerequisite for measurement-based feedback control of quantum systems. Indeed, in another experiment performed by our collaborators at ENS (team of Benjamin Huard and François Mallet), we demonstrated stabilization of an arbitrary trajectory of a superconducting qubit by such a measurement-based feedback [37]. The protocol benefits from the long coherence time ($T_2 > 10\mu$s) of the 3D transmon qubit, the high efficiency (82%) of the phase preserving Josephson amplifier, and fast electronics ensuring less than 500 ns delay. At discrete time intervals, the state of the qubit is measured and corrected in case an error is detected. For Rabi oscillations, where the discrete measurements occur when the qubit is supposed to be in the measurement pointer states, we demonstrate an average fidelity of 85% to the targeted trajectory. Incidentally, we demonstrate a fast reset protocol allowing to cool a 3D transmon qubit down to 0.6% in the excited state.

6.3. Classical engineering: Monitoring and control of complex systems

6.3.1. Modeling, signal analysis and control with medical applications

Participants: Alexandre Guerrini, Lisa Guigue, Claire Médigue, Michel Sorine, Serge Steer.

Reduced order cardiac modeling and applications. See Section 4.3 for complements. We consider two topics:
- Personalized medecine: a first validation on clinical data of our model of controlled contraction of cardiac muscle has been obtained [55].
- Heart Failure with preserved Ejection Fraction (HFpEF): this work is done in collaboration with Bijan Gahleh (INSERM U955). Our objective is to define markers of HFpEF identifiable from noninvasive measurements. After having assembled a high precision ECG acquisition and post-processing system, we have measured multi-lead ECG on pigs treated to induce HFpEF, cf. B. Gahleh et al [109]. The analysis of the diastolic electric interval (e.g. P-wave, PR interval etc.) is ongoing.

Semiclassical analysis of cardiovascular signals. A summary of the theory is now published [51].

CGAO-REA: Computerized Glucose Control in Critically Ill Patients. The version CGAO_v1 of our controller (see Sections 4.3), has been used in a large multi-center study, CGAO-REA (35 active ICUs, more than 3500 included patients). Mortality has not been changed [70], [48] but the protocol is now formalized and tunable. CGAO-REA has proved that our controller is robust in the real life context and comparable to human control with its present tuning. Improving the tuning (in particular the glycemic target) seems possible.

6.3.2. Diagnosis of inhomogeneous insulation degradation in electric cables by distributed shunt conductance estimation

Participant: Qinghua Zhang.

For the diagnosis of inhomogeneous insulation degradation in electric cables, the estimation of distributed shunt conductance is studied in this work. Gradual growth of the shunt conductance is a consequence of degradation of the dielectric properties of the insulator. The proposed estimation method is based on voltage and current measurements at a single end of the cable. After the linearization of the bilinear term of the telegrapher's equations through a perturbation approach, the Kalman filter is applied to transform the problem of dynamic system parameter estimation to a simple linear regression problem. Numerical simulations are made to demonstrate the feasibility of the proposed method. In particular, it is shown that the weak sensitivity of the available measurements to the shunt conductance can be compensated by long time data samples. See [61] for more details.

6.3.3. Feasibility of reflectometry techniques for non destructive evaluation of external post-tensioned cables

Participants: Michel Sorine, Qinghua Zhang.

Nowadays a considerable number of bridges is reaching an age when renovating operations become necessary. For some bridges, external post-tension is realized with cables protected in ducts, with the residual internal space imperfectly filled with a fluid cement grout. Detecting the problems of injection in the ducts is visually impossible from the outside. In collaboration with IFSTTAR (Institut Français des Sciences et Technologies des Transports, de l'Aménagement et des Réseaux) through the I4S team common to Inria and IFSTTAR, the feasibility of reflectometry techniques for cable health monitoring is investigated via numerical simulations and laboratory experiments. The main idea consists in adding electrically conductive tapes along a duct so that the duct and the added tapes can be treated as an electrical transmission line. It is then possible to apply advanced reflectometry methods developed by the SISYPHE project-team, initially for true electric cables.

6.3.4. Nonlinear system identification

Participants: Boyi Ni, Michel Sorine, Qinghua Zhang.

In the framework of the joint Franco-Chinese ANR-NSFC EBONSI project (see Section 8.1.1), the topics studied this year on nonlinear system identification are mainly on the detection of asymmetric control valve stiction from oscillatory data based on a method for extended Hammerstein system identification, and on the identification of Wiener systems.

The study on control valve stiction is motivated by the detection of control valves with asymmetric stiction resulting in oscillations in feedback control loops. The joint characterization of the control valve and the controlled process is formulated as the identification of a class of extended Hammerstein systems. The input nonlinearity is described by a point-slope-based hysteretic model with two possibly asymmetric ascent and descent paths. An iterative identification method is proposed, based on the idea of separating the ascent and descent paths subject to the oscillatory input and output. The structure of the formulated extended Hammerstein system is shown to be identifiable, and the oscillatory signals in feedback control loops are proved to be informative by exploiting the cyclo-stationarity of these oscillatory signals. Numerical, experimental and industrial examples confirm the effectiveness of the proposed identification method.

Wiener system identification has been investigated this year by focusing on the estimation of the finite impulse response (FIR) of the linear subsystem. Under the assumption of Gaussian input distribution, this work mainly aims at addressing a deficiency of the well-known correlation-based method for Wiener system identification: it fails when the nonlinearity of the Wiener system is an even function. This method is, in the considered Gaussian input case, equivalent to the best linear approximation (BLA), which exhibits the same deficiency. A new method is developed this year, based on a weighted principal component analysis (wPCA). Its consistency is proved for Wiener systems with either even or non even nonlinearities. Its computational cost is almost the same as that of a standard PCA. Numerical simulations are made to compare the new wPCA-based method to the correlation-based method for different Wiener systems with nonlinearities more or less close to an even function.

6.3.5. Model-based fault diagnosis for descriptor systems

Participants: Abdouramane Moussa Ali, Qinghua Zhang.

This work is about fault diagnosis for linear time varying descriptor systems, the discrete time counterpart of dynamic systems described by differential-algebraic equations. The Kalman filter for descriptor systems is first revisited by completing existing results about its properties that are essential for the purpose of fault diagnosis. Based on the analysis of the effects of the considered actuator and sensor faults on the innovation of the Kalman filter, it is shown that the considered fault diagnosis problem in linear time varying descriptor systems can be transformed to a classical linear regression problem formulated by appropriately filtering the input-output data. Following this result, algorithms for fault diagnosis through maximum likelihood estimation are then developed.

In the framework of the ITEA2 MODRIO project (see Section 8.2.1), this work is in preparation for studying hybrid system monitoring, aiming at extending existing results from state-space systems to descriptor systems in the modes of a hybrid system.

6.3.6. Analysis of the Behavior of Networks of Dynamical Systems

Participant: Pierre-Alexandre Bliman.

We have established convergence results for some continuous-time dynamics which are analogs to ant colony optimization algorithms that solve shortest path problems. Global asymptotic stability has been shown, and the speed of convergence has been calculated explicitly and shown to be proportional to the difference between the reciprocals of the second shortest and the shortest paths. Such precise results are missing in the context of ant colony optimization algorithms (which are discrete-time dynamical systems). The systems studied are special instances of networks of dynamical systems which represent the evolution of some state variable on each path, coupled in a competitive way through global macroscopic quantity. Such models are related to simple forms of models studied in mathematical epidemiology, which will be the subject of further work. This work is done in cooperation with Amit Bhaya from COPPE, Universidade Federal de Rio de Janeiro. Papers have been submitted [77].