Activity Report 2014

Section New Results

<p style="text-align:center">**ARIC Project-Team**</p>

6. New Results

6.1. Arithmetic operators

6.1.1. A table-based method to evaluate trigonometric functions

Linear (order-one) function evaluation schemes, such as bipartite and multipartite tables, are usually effective for low precision approximations. For high output precision, the lookup table size is often too large for practical use. Dong Wang and Milos Ercegovac (UC Los Angeles) and Nicolas Brisebarre and Jean-Michel Muller investigate the so-called (M, p, k) scheme that reduces the range of input argument to a very small interval so that trigonometric functions can be approximated with very small lookup tables and a few additions/subtractions. An optimized hardware architecture is proposed and implemented in both FPGA device and standard cell based technology. Experimental results show that the proposed scheme achieves more than 50% reduction in total chip area compared with the best linear approach for 24-bit evaluation [14].

6.2. Floating-Point arithmetic

6.2.1. On the computation of the reciprocal of floating point expansions using an adapted Newton-Raphson iteration

Many numerical problems require a higher computing precision than that offered by common floating point (FP) formats. One common way of extending the precision is to represent numbers in a *multiple component* format. With so-called *floating point expansions*, numbers are represented as the unevaluated sum of standard machine precision FP numbers. This format offers the simplicity of using directly available and highly optimized FP operations and is used by multiple-precisions libraries such as Bailey's 'Q'D or the analogue Graphics Processing Units tuned version, GQD. Mioara Joldes (LAAS), Jean-Michel Muller, and Valentina Popescu introduced a new algorithm for computing the reciprocal FP expansion a^{-1} of a FP expansion a. Their algorithm is based on using an adapted Newton-Raphson iteration where "truncated" operations (additions, multiplications) involving FP expansions are used. The error analysis given shows that their algorithm allows for computations of very accurate quotients. Precisely, after $i \geq 0$ iterations, the computed FP expansion $x = x_0 + \cdots + x_{2^i-1}$ satisfies the relative error bound $|\frac{x-a^{-1}}{a^{-1}}| \leq 2^{-2^i(p-3)-1}$, where $p > 4$ is the precision of the FP representation used ($p = 24$ for single precision and $p = 53$ for double precision) [19].

6.2.2. Error bounds on complex floating-point multiplication with a fused-multiply add

The accuracy analysis of complex floating-point multiplication done by Brent, Percival, and Zimmermann [*Math. Comp.*, 76:1469–1481, 2007] is extended by Peter Kornerup (Odense Univ. Denmark), Claude-Pierre Jeannerod, Nicolas Louvet, and Jean-Michel Muller [42] to the case where a fused multiply-add (FMA) operation is available. Considering floating-point arithmetic with rounding to nearest and unit roundoff u, they show that the bound $\sqrt{5}\,u$ on the normwise relative error $|\hat{z}/z - 1|$ of a complex product z can be decreased further to $2u$ when using the FMA in the most naive way. Furthermore, they prove that the term $2u$ is asymptotically optimal not only for this naive FMA-based algorithm, but also for two other algorithms, which use the FMA operation as an efficient way of implementing rounding error compensation. Thus, although highly accurate in the componentwise sense, these two compensated algorithms bring no improvement to the normwise accuracy $2u$ already achieved using the FMA naively. Asymptotic optimality is established for each algorithm thanks to the explicit construction of floating-point inputs for which it is proven that the normwise relative error then generated satisfies $|\hat{z}/z - 1| \to 2u$ as $u \to 0$. All these results hold for IEEE floating-point arithmetic, with radix $\beta \geq 2$, precision $p \geq 2$, and rounding to nearest; it is only assumed that underflows and overflows do not occur and, when bounding errors from below, that $\beta^{p-1} \geq 12$.

6.2.3. *Refined error analysis of the Cornea-Harrison-Tang method for* $ab + cd$

In their book *Scientific Computing on Itanium-based Systems*, Cornea, Harrison, and Tang introduced an accurate algorithm for evaluating expressions of the form $ab + cd$ in binary floating-point arithmetic, assuming a fused-multiply add instruction is available. They showed that if p is the precision of the floating-point format and if $u = 2^{-p}$, the relative error of the result is of order u. Jean-Michel Muller improved their proof to show that the relative error is bounded by $2u + 7u^2 + 6u^3$. Furthermore, by building an example for which the relative error is asymptotically (as $p \to \infty$ or, equivalently, as $u \to 0$) equivalent to $2u$, he proved that this error bound is asymptotically optimal [11]. Claude-Pierre Jeannerod then showed in [41] that an error bound of the form $2u + 2u^2 + O(u^3)$ in fact holds for any radix $\beta \geq 2$, with $u = \frac{1}{2}\beta^{1-p}$. He also showed that the possibility of removing the $O(u^2)$ term from this bound depends on the radix parity and the tie-breaking strategy used for rounding: if β is odd or rounding is *to nearest even* then the simpler bound $2u$ is obtained, while if β is even and rounding is *to nearest away*, then there exist floating-point inputs a, b, c, d that lead to a relative error larger than $2u + \frac{1}{\beta}u^2$.

6.2.4. *On the maximum relative error when computing integer powers by iterated multiplications in floating-point arithmetic*

Stef Graillat (Paris 6 University), Vincent Lefèvre and Jean-Michel Muller improved the usual relative error bound for the computation of x^n through iterated multiplications by x in binary floating-point arithmetic. The obtained error bound is only slightly better than the usual one, but it is simpler. They also discussed the more general problem of computing the product of n terms [7].

6.2.5. *Improved error bounds for numerical linear algebra*

When computing matrix factorizations and solving linear systems in floating-point arithmetic, classical rounding error analyses provide backward error bounds whose leading terms have the form $\gamma_n = nu/(1 - nu)$ for suitable values of n and with u the unit roundoff. With Siegfried M. Rump (Hamburg University of Technology), Claude-Pierre Jeannerod showed in [13] that for LU and Cholesky factorizations as well as for triangular system solving, γ_n can be replaced by the $O(u^2)$-free and unconditional constant nu. To get these new bounds the main ingredient is a general framework for bounding expressions of the form $|\rho - s|$, where s is the exact sum of a floating-point number and $n - 1$ real numbers, and where ρ is a real number approximating the computed sum $\hat{s}$.

6.2.6. *On relative errors of floating-point operations*

Rounding error analyses of numerical algorithms are most often carried out via repeated applications of the so-called standard models of floating-point arithmetic. Given a round-to-nearest function RN and barring underflow and overflow, such models bound the relative errors $E_1(t) = |t - \text{RN}(t)|/|t|$ and $E_2(t) = |t - \text{RN}(t)|/|\text{RN}(t)|$ by the unit roundoff u. In [34] Claude-Pierre Jeannerod and Siegfried M. Rump (Hamburg University of Technology) investigated the possibility of refining these bounds, both in the case of an arbitrary real t and in the case where t is the exact result of an arithmetic operation on some floating-point numbers. They provided explicit and attainable bounds on $E_1(t)$, which are all less than or equal to $u/(1 + u)$ and, therefore, smaller than u. For $E_2(t)$ the bound u is attainable whenever $t = x \pm y$ or $t = xy$ or, in base > 2, $t = x/y$ with x, y two floating-point numbers. However, for division in base 2 as well as for square root, smaller bounds are derived, which are also shown to be attainable. This set of sharp bounds was then applied to the rounding error analysis of various numerical algorithms: in all cases, they obtained either much shorter proofs of the best-known error bounds for such algorithms, or improvements on these bounds themselves.

6.2.7. *Comparison between binary and decimal floating-point numbers*

In collaboration with Christoph Lauter and Marc Mezzarobba (LIP6 laboratory, Paris), Nicolas Brisebarre and Jean-Michel Muller introduce an algorithm to compare a binary floating-point (FP) number and a decimal FP number, assuming the "binary encoding" of the decimal formats is used, and with a special emphasis on the basic interchange formats specified by the IEEE 754-2008 standard for FP arithmetic. It is a two-step algorithm: a first pass, based on the exponents only, quickly eliminates most cases, then, when the first pass

does not suffice, a more accurate second pass is performed. They provide an implementation of several variants of our algorithm, and compare them [37].

6.2.8. Correctly rounded sum of floating-point numbers in GNU MPFR

Vincent Lefèvre has designed a new algorithm to compute the correctly rounded sum of several floating-point numbers, each having its own precision and the output having its own precision, as in GNU MPFR. At the same time, the `mpfr_sum` function is being reimplemented (not finished yet). While the old algorithm was just an application of Ziv's method, thus with exponential time and memory complexity in the worst case such as the sum of a huge number and a tiny number, the new algorithm does the sum by blocks (reiterations being needed only in case of cancellations), taking such holes between numbers into account.

6.3. Certified computing and computer algebra

6.3.1. Standardization of interval arithmetic

The IEEE 1788 working group is devoted to the standardization of interval arithmetic. V. Lefèvre and N. Revol are very active in this group. This year has been devoted to a ballot on the whole text of the standard [28], and to editorial work to make it compliant with IEEE rules. The final, remaining step, is the so-called "Sponsor ballot" and it should be completed in 2015.

6.3.2. Interval linear algebra on multi-core processors

For the product of matrices with interval coefficients, fast approximate algorithms have been developed by Philippe Théveny: they compute an enclosure of the exact product. These algorithms rely on the representation of intervals by their midpoints and radii. This representation allows one to use optimized routines for the multiplication of matrices with floating-point coefficients. In [4], the quality of the approximation of several algorithms is established, which accounts for roundoff errors and not only method's errors. A new algorithm is proposed, which requires even less (only 2) calls to a floating-point routine and still offers a good approximation quality, for a well specified type of input matrices. Three of the studied algorithms are implemented on a multi-core architecture. To avoid problems listed in [12] and to offer good performances, Philippe Théveny developed optimizations. The resulting implementations exhibit good performances: guaranteed results are obtained with an overhead less than 3, high numerical intensity and good scalability.

6.3.3. Numerical reproducibility

What is called *numerical reproducibility* is the problem of getting the same result when the scientific computation is run several times, either on the same machine or on different machines. In [12], the focus is on interval computations using floating-point arithmetic: Nathalie Revol and Philippe Théveny identified implementation issues that may invalidate the inclusion property, and presented several ways to preserve this inclusion property. This work has also been replaced in the larger context of numerical validation [15].

6.3.4. Faster multivariate interpolation with multiplicities

Muhammad Chowdhury (U. Western Ontario), Claude-Pierre Jeannerod, Vincent Neiger (ENS de Lyon), Éric Schost (U. Western Ontario), and Gilles Villard proposed in [38] a fast algorithm for interpolating multivariate polynomials with multiplicities. This algorithm relies on the reduction to a problem of simultaneous polynomial approximations, which is then solved using fast structured linear algebra techniques. This algorithm leads to the best known complexity bounds for the interpolation step of the list-decoding of Reed-Solomon codes, Parvaresh-Vardy codes or folded Reed-Solomon codes. In the special case of Reed-Solomon codes, it allows to accelerate the interpolation step of Guruswami and Sudan's list-decoding by a factor (list size)/(multiplicity).

6.3.5. Polynomial system solving

M. Bardet (U. Rouen), J.-C. Faugère (PolSys team) and B. Salvy studied the complexity of Gröbner bases computation, in particular in the generic situation where the variables are in simultaneous Noether position with respect to the system. They gave a bound on the number of polynomials of each degree in a Gröbner basis computed by Faugère's F_5 algorithm in this generic case for the grevlex ordering (which is also a bound on the number of polynomials for a reduced Gröbner basis, independently of the algorithm used) and used it to bound the complexity of the F_5 algorithm [5].

6.3.6. Linear differential equations

In [6], A. Bostan (SpecFun team), K. Raschel (U. Tours) and B. Salvy proved that the sequence $(e_n^{\mathfrak{S}})_{n \geq 0}$ of excursions in the quarter plane corresponding to a nonsingular step set $\mathfrak{S} \subseteq \{0, \pm 1\}^2$ with infinite group does not satisfy any nontrivial linear recurrence with polynomial coefficients. Accordingly, in those cases, the trivariate generating function of the numbers of walks with given length and prescribed ending point is not D-finite. Moreover, they displayed the asymptotics of $e_n^{\mathfrak{S}}$. This completes the classification of these walks.

Colleagues from the LAAS (Toulouse) and B. Salvy provided a new method for computing the probability of collision between two spherical space objects involved in a short-term encounter. In this specific framework of conjunction, classical assumptions reduce the probability of collision to the integral of a 2-D normal distribution over a disk shifted from the peak of the corresponding Gaussian function. Both integrand and domain of integration directly depend on the nature of the short-term encounter. Thus the inputs are the combined sphere radius, the mean relative position in the encounter plane at reference time as well as the relative position covariance matrix representing the uncertainties. The method they presented is based on an analytical expression for the integral. It has the form of a convergent power series whose coefficients verify a linear recurrence. It is derived using Laplace transform and properties of D-finite functions. The new method has been intensively tested on a series of test-cases and compares favorably to other existing works [29].

6.4. Lattices and cryptography

6.4.1. Worst-Case to Average-Case Reductions for Module Lattices

Most lattice-based cryptographic schemes are built upon the assumed hardness of the Short Integer Solution (SIS) and Learning With Errors (LWE) problems. Their efficiencies can be drastically improved by switching the hardness assumptions to the more compact Ring-SIS and RingLWE problems. However, this change of hardness assumptions comes along with a possible security weakening: SIS and LWE are known to be at least as hard as standard (worst-case) problems on euclidean lattices, whereas Ring-SIS and Ring-LWE are only known to be as hard as their restrictions to special classes of ideal lattices, corresponding to ideals of some polynomial rings. Adeline Langlois and Damien Stehlé defined the Module-SIS and Module-LWE problems, which bridge SIS with Ring-SIS, and LWE with Ring-LWE, respectively. They proved that these average-case problems are at least as hard as standard lattice problems restricted to module lattices (which themselves bridge arbitrary and ideal lattices). As these new problems enlarge the toolbox of the lattice-based cryptographer, they could prove useful for designing new schemes. Importantly, the worst-case to average-case reductions for the module problems are (qualitatively) sharp, in the sense that there exist converse reductions. This property is not known to hold in the context of Ring-SIS/Ring-LWE: Ideal lattice problems could reveal easy without impacting the hardness of Ring-SIS/Ring-LWE [8].

6.4.2. Semantically Secure Lattice Codes for the Gaussian Wiretap Channel

Cong Ling (Imperial College, UK), Laura Luzzi (ENSEA), Jean-Claude Belfiore (Telecom ParisTech) and Damien Stehlé proposed a new scheme of wiretap lattice coding that achieves semantic security and strong secrecy over the Gaussian wiretap channel. The key tool in their security proof is the flatness factor which characterizes the convergence of the conditional output distributions corresponding to different messages and leads to an upper bound on the information leakage. They not only introduced the notion of secrecy-good lattices, but also proposed the flatness factor as a design criterion of such lattices. Both the modulo-lattice

Gaussian channel and the genuine Gaussian channel are considered. In the latter case, they proposed a novel secrecy coding scheme based on the discrete Gaussian distribution over a lattice, which achieves the secrecy capacity to within a half nat under mild conditions. No a priori distribution of the message is assumed, and no dither is used in their proposed schemes [9].

6.4.3. GGHLite: More Efficient Multilinear Maps from Ideal Lattices

The Garg-Gentry-Halevi (GGH) Graded Encoding Scheme, based on ideal lattices, is the first plausible approximation to a cryptographic multilinear map. Unfortunately, the scheme requires very large parameters to provide security for its underlying encoding re-randomization process. Adeline Langlois, Damien Stehlé and Ron Steinfeld (Monash University, Australia) formalized, simplified and improved the efficiency and the security analysis of the re-randomization process in the GGH construction. This results in a new construction that they called GGHLite. In particular, they first lowered the size of a standard deviation parameter of the GGH re-randomization process from exponential to polynomial in the security parameter. This first improvement is obtained via a finer security analysis of the so-called drowning step of re-randomization, in which they applied the Rényi divergence instead of the conventional statistical distance as a measure of distance between distributions. Their second improvement is to reduce the number of randomizers needed to 2, independently of the dimension of the underlying ideal lattices. These two contributions allowed them to decrease the bit size of the public parameters to $O(\lambda \log^2 \lambda)$ in GGHLite, with respect to the security parameter λ (for a constant multilinearity parameter κ) [22].

6.4.4. LLL reducing with the most significant bits

Let B be a basis of a Euclidean lattice, and $\widetilde{B}$ an approximation thereof. Saruchi (IIT Delhi, India), Ivan Morel, Damien Stehlé and Gilles Villard gave a sufficient condition on the closeness between $\widetilde{B}$ and B so that an LLL-reducing transformation U for $\widetilde{B}$ remains valid for B. Further, they analysed an efficient reduction algorithm when B is itself a small deformation of an LLL-reduced basis. Applications include speeding-up reduction by keeping only the most significant bits of B, reducing a basis that is only approximately known, and efficiently batching LLL reductions for closely related inputs [30].

6.4.5. Hardness of k-LWE and Applications in Traitor Tracing

San Ling (NTU, Singapore), Duong Hieu Phan (LAGA), Damien Stehlé and Ron Steinfeld (Monash University, Australia) introduced the k-LWE problem, a Learning With Errors variant of the k-SIS problem. The Boneh-Freeman reduction from SIS to k-SIS suffers from an exponential loss in k. Ling *et al.* improved and extended it to an LWE to k-LWE reduction with a polynomial loss in k, by relying on a new technique involving trapdoors for random integer kernel lattices. Based on this hardness result, they presented the first algebraic construction of a traitor tracing scheme whose security relies on the worstcase hardness of standard lattice problems. The proposed LWE traitor tracing is almost as efficient as the LWE encryption. Further, it achieves public traceability, i.e., allows the authority to delegate the tracing capability to untrusted parties. To this aim, Ling *et al.* introduced the notion of projective sampling family in which each sampling function is keyed and, with a projection of the key on a well chosen space, one can simulate the sampling function in a computationally indistinguishable way. The construction of a projective sampling family from k-LWE allows us to achieve public traceability, by publishing the projected keys of the users [27].

6.4.6. Lattice-Based Group Signatures Scheme with Verifier-local Revocation

Support of membership revocation is a desirable functionality for any group signature scheme. Among the known revocation approaches, verifier-local revocation (VLR) seems to be the most flexible one, because it only requires the verifiers to possess some up-to-date revocation information, but not the signers. All of the contemporary VLR group signatures operate in the bilinear map setting, and all of them will be insecure once quantum computers become a reality. Adeline Langlois, San Ling, Khoa Nguyen and Huaxiong Wang (NTU, Singapore) introduced the first lattice-based VLR group signature [21], and thus, the first such scheme that is believed to be quantum-resistant. In comparison with existing lattice-based group signatures, this scheme has several noticeable advantages: support of membership revocation, logarithmic-size signatures, and weaker security assumption. In the random oracle model, our scheme is proved to be secure based on the hardness of

the Shortest Independent Vector Problem with approximation factor $\gamma = \widetilde{O}(n^{1.5})$ - an assumption that is as weak as those of state-of-the-art lattice-based standard signatures. Moreover, this construction works without relying on encryption schemes, which is an intriguing feature for group signatures.

6.4.7. Proxy Re-Encryption Scheme Supporting a Selection of Delegatees

Julien Devigne (Orange Labs), Eleonora Guerrini (Univ. Montpellier 2, LIRMM) and Fabien Laguillaumie adapt the primitive of proxy re-encryption which allows a user to decide that in case of unavailability, one (or several) particular user, the delegatee, will be able to read his confidential messages. They modify it so that a sender can choose who among many potential delegatees will be able to decrypt his messages, and propose a simple and efficient scheme which is secure under chosen plaintext attack under standard algorithmic assumption in a bilinear setting. They also investigate the possibility to add a traceability of the proxy so that one can detect if it has leaked some re-encryption keys [17].

6.4.8. Practical validation of several fault attacks against the Miller algorithm

Ronan Lashermes (SAS-ENSMSE, PRISM), Marie Paindavoine, Nadia El Mrabet (Univ. P8, LIASD), Jacques Fournier (SAS-ENSMSE) and Louis Goubin (UVSQ, PRISM) describe practical implementations of fault attacks against the Miller algorithm, which computes pairing evaluations on algebraic curves. These implementations validate common fault models used against pairings. In the light of the implemented fault attacks, they show that some blinding techniques proposed to protect the algorithm against Side-Channels Analyses cannot be used as countermeasures against the implemented fault attacks [23].

6.4.9. Non-Malleability from Malleability: Simulation-Sound Quasi-Adaptive NIZK Proofs and CCA2-Secure Encryption from Homomorphic Signatures

Verifiability is central to building protocols and systems with integrity. Initially, efficient methods employed the Fiat-Shamir heuristics. Since 2008, the Groth-Sahai techniques have been the most efficient in constructing non-interactive witness indistinguishable and zero-knowledge proofs for algebraic relations in the standard model. For the important task of proving membership in linear subspaces, Jutla and Roy (Asiacrypt 2013) gave significantly more efficient proofs in the quasi-adaptive setting (QA-NIZK). For membership of the row space of a $t \times n$ matrix, their QA-NIZK proofs save $\Omega(t)$ group elements compared to Groth-Sahai. In [26], Benoît Libert, Thomas Peters (UCL, Belgique), Marc Joye (Technicolor, USA) and Moti Yung (Google and Columbia U, USA) gave QA-NIZK proofs made of a *constant* number group elements – regardless of the number of equations or the number of variables – and additionally proved them *unbounded* simulation-sound. Unlike previous unbounded simulation-sound Groth-Sahai-based proofs, their construction does not involve quadratic pairing product equations and does not rely on a chosen-ciphertext-secure encryption scheme. Instead, they built on structure-preserving signatures with homomorphic properties. They applied their methods to design new and improved CCA2-secure encryption schemes. In particular, they built the first efficient threshold CCA-secure keyed-homomorphic encryption scheme (*i.e.*, where homomorphic operations can only be carried out using a dedicated evaluation key) with publicly verifiable ciphertexts.

6.4.10. Born and Raised Distributively: Fully Distributed Non-Interactive Adaptively-Secure Threshold Signatures with Short Shares

Threshold cryptography is a fundamental distributed computational paradigm for enhancing the availability and the security of cryptographic public-key schemes. It does it by dividing private keys into n shares handed out to distinct servers. In threshold signature schemes, a set of at least $t + 1 \leq n$ servers is needed to produce a valid digital signature. Availability is assured by the fact that any subset of $t + 1$ servers can produce a signature when authorized. At the same time, the scheme should remain robust (in the fault tolerance sense) and unforgeable (cryptographically) against up to t corrupted servers; *i.e.*, it adds quorum control to traditional cryptographic services and introduces redundancy. Originally, most practical threshold signatures have a number of demerits: They have been analyzed in a static corruption model (where the set of corrupted servers is fixed at the very beginning of the attack), they require interaction, they assume a trusted dealer in the key generation phase (so that the system is not fully distributed), or they suffer from certain overheads in terms of storage (large share sizes).

In [24], Benoît Libert, Marc Joye (Technicolor, USA) and Moti Yung (Google and Columbia U, USA) constructed practical *fully distributed* (the private key is born distributed), non-interactive schemes – where the servers can compute their partial signatures without communication with other servers – with adaptive security (*i.e.*, the adversary corrupts servers dynamically based on its full view of the history of the system). Their schemes are very efficient in terms of computation, communication, and scalable storage (with private key shares of size $O(1)$, where certain solutions incur $O(n)$ storage costs at each server). Unlike other adaptively secure schemes, their schemes are erasure-free (reliable erasure is a hard to assure and hard to administer property in actual systems). Such a fully distributed highly constrained scheme has been an open problem in the area. In particular, and of special interest, is the fact that Pedersen's traditional distributed key generation (DKG) protocol can be safely employed in the initial key generation phase when the system is born – although it is well-known not to ensure uniformly distributed public keys. An advantage of this is that this protocol only takes one round optimistically (in the absence of faulty player).

6.4.11. Concise Multi-challenge CCA-Secure Encryption and Signatures with Almost Tight Security

To gain strong confidence in the security of a public-key scheme, it is most desirable for the security proof to feature a tight reduction between the adversary and the algorithm solving the under-lying hard problem. Recently, Chen and Wee (Crypto '13) described the first Identity-Based Encryption scheme with almost tight security under a standard assumption. Here, "almost tight" means that the security reduction only loses a factor $O(\lambda)$ – where λ is the security parameter – instead of a factor proportional to the number of adversarial queries. Chen and Wee also gave the shortest signatures whose security almost tightly relates to a simple assumption in the standard model. Also recently, Hofheinz and Jager (Crypto '12) constructed the first CCA-secure public-key encryption scheme in the multi-user setting with tight security. These constructions give schemes that are significantly less efficient in length (and thus, processing) when compared with the earlier schemes with loose reductions in their proof of security. Hofheinz and Jager's scheme has a ciphertext of a few hundreds of group elements, and they left open the problem of finding truly efficient constructions. Likewise, Chen and Wee's signatures and IBE schemes are somewhat less efficient than previous constructions with loose reductions from the same assumptions.

In [25], Benoît Libert, Thomas Peters (UCL, Belgique), Marc Joye (Technicolor, USA) and Moti Yung (Google and Columbia U, USA) considered space-efficient schemes with security almost tightly related to standard assumptions. As a step in solving the open question by Hofheinz and Jager, they constructed an efficient CCA-secure public-key encryption scheme whose chosen-ciphertext security in the multi-challenge, multi-user setting almost tightly relates to the DLIN assumption (in the standard model). Quite remarkably, the ciphertext size decreases to 69 group elements under the DLIN assumption whereas the best previous solution required about 400 group elements. Their scheme is obtained by taking advantage of a new almost tightly secure signature scheme (in the standard model) they developed and which is based on the recent concise proofs of linear subspace membership in the quasi-adaptive non-interactive zero-knowledge setting (QA-NIZK) defined by Jutla and Roy (Asiacrypt '13). The new signature scheme reduces the length of the previous such signatures (by Chen and Wee) by 37% under the Decision Linear assumption, by almost 50% under the K-LIN assumption, and it becomes only 3 group elements long under the Symmetric eXternal Diffie-Hellman assumption. Our signatures are obtained by carefully combining the proof technique of Chen and Wee and the above mentioned QA-NIZK proofs.

COMPSYS Project-Team

6. New Results

6.1. Highlights of the Year

For 2014, from the point of view of organization, funding, collaborations, the main points to highlight are:

- Christophe Alias and Alexandru Plesco have co-founded the XTREMLOGIC start-up in January 2014 (see Section 7.2), following the incubation of Zettice. XTREMLOGIC recently won the "concours région rhône-alpes" grant in November 2014 (40k).

- Tomofumi Yuki was hired as an Inria researcher and became a permanent member of Compsys.

- The 1988 "Array Expansion" seminal paper of Paul Feautrier has been selected for the 25th Anniversary Volume of the ACM International Conference on Supercomputing (ICS) together with 34 other papers selected from the 1800 papers published from 1987 to 2011. A short "reminescence" paper [13] was written for the occasion.

- The team was evaluated in Nov. 2014 by the HCERES (new name of AERES), as part of the LIP lab evaluation. The report has not been received yet.

From a scientific point of view, the shift, in Compsys III, towards the analysis of parallel programs and the extensions of the polyhedral model, both in terms of techniques and applications, is continuing, see the section "New Results", in particular:

- The design (by Christophe Alias and Alexandru Plesco) of a HLS compiler technology (see Section 6.2), patented by Inria [12] and transferred to XTREMLOGIC under an Inria licence (see Section 5.5).

- Two new static analyses: a more precise array bound check analysis [9] (see Section 6.3) and a more scalable termination algorithm for C programs (see Section 6.4).

- A novel equivalence-checking algorithm [7] modulo associativity/commutativity, which is a first step towards semantic program transformations (see Section 6.5).

- A groundbreaking introduction of polyhedral techniques for the analysis of parallel programs, in particular X10 (see [29] and [6]) and OpenStream (see Section 6.6).

- A seminal paper [5] introducing polynomial techniques in program analysis and compilation (see Section 6.7).

- Innovative contributions on parametric tiling [8], [3], [4] as extensions of the polyhedral model (see Sections 6.8 and 6.9).

6.2. Data-Aware Process Networks

Participants: Christophe Alias, Alexandru Plesco [XTREMLOGIC start-up].

Process networks are execution models expressing naturally the parallelism of a computation. They are a natural intermediate representation for high-level synthesis tools, where the front-end extracts the parallelism and produces a process network and the back-end compiles the process network to the target architecture.

In that context, we have defined a new model of process network that fits HLS-specific constraints, the data-aware process network (DPN). Our model makes explicit the communications with the central memory and the parallel access to channels, and is close enough to the hardware constraints to be translated directly to a circuit. We show how to compile an imperative program to a DPN, so as to optimize both the I/O and the parallelism, while using the polyhedral model.

DPNs are used as the intermediate representation for the HLS compiler suite of the XTREMLOGIC start-up. They are generated from C programs by the Dcc compiler (see Section 5.5). The apparatus underlying the DPN synchronizations has been patented by Inria [12].

6.3. Preventing from Out-of-Bound Memory Accesses

Participants: Laure Gonnord, Fernando Pereira [Univ. Mineas Gerais, Brasil].

The C programming language does not prevent out-of-bounds memory accesses. There exist several techniques to secure C programs; however, these methods tend to slow down these programs substantially, because they populate the binary code with runtime checks. To deal with this problem, we designed and tested two static analyses (symbolic region and range analysis), which we combine to remove the majority of these guards.

In addition to the analyses themselves, we brought two other contributions:

- First, we described live-range splitting strategies that improve the efficiency and the precision of our analyses.
- Secondly, we showed how to deal with integer overflows, a phenomenon that can compromise the correctness of static algorithms validating memory accesses.

We validated our claims by incorporating our findings into AddressSanitizer (see https://code.google.com/p/ address-sanitizer/). We generated SPEC CINT 2006 code that is 17% faster and 9% more energy efficient than the code originally produced by this tool. Furthermore, our approach is 50% more effective than Pentagons, a state-of-the-art analysis to sanitize memory accesses. This work was published at the OOPSLA 2014 conference [9].

6.4. Scaling Termination Proofs

Participants: Laure Gonnord, Gabriel Radanne [ENS Rennes], David Monniaux [CNRS/VERIMAG], Fernando Pereira [Univ. Mineas Gerais, Brasil], Raphael Rodrigues [Univ. Mineas Gerais, Brasil].

In [15], we presented a new algorithm adapted from scheduling techniques to synthesize (multi-dimensional) affine functions from general flowcharts programs. But, as for other methods, our algorithm tried to solve linear constraints on each control point and each transition, which can lead to quasi-intractable linear programming instances. In contrast to these approaches, we proposed a new algorithm based on the following observations:

- Searching for ranking functions for loop headers is sufficient to prove termination.
- Furthermore, there exist loops such that there is a linear lexicographic ranking function that decreases along each path inside the loop, from one loop iteration to the next, but such that there is no lexicographic linear ranking function that decreases at each step along these paths. For these reasons, it is tempting to treat each path inside a loop as a single transition.

Unfortunately the number of paths may be exponential in the size of the program, thus the constraint system may become very large, even though it features fewer variables. To face this theoretical complexity, even though the number of paths may be large, we argue that, in practice, few of them actually matter in the constraint system (we formalize this concept by giving a characterization as geometric extremal points). Our algorithm therefore builds the constraint system lazily, taking paths into account *on demand*.

In 2014, we consolidated this approach with a work on complexity issues (inspired by [19]) and a new implementation: Termite (see Section 5.13). A corresponding paper is currently under submission for PLDI.

With Fernando Pereira's group in Brazil, we also studied the relevance of fast and simple solutions to compute approximations of the number of iterations of loops (*loop trip count*) of imperative real-world programs. The context of this work is the use of these approximations in compiler optimizations: most of the time, the optimizations yield greater benefits for large trip counts, and are either innocuous or detrimental for small ones. In our paper published at WST'14 [10], we have shown that, most of the time, there is no need to use computationally-expensive state-of-the-art methods to compute (an approximation of) it. We support our position with an actual case study. We show that a fast predictor can be used to speedup the JavaScript JIT compiler of Firefox - one of the most well-engineered runtime environments in use today.

6.5. Equivalence-Checking of Programs with Reductions

Participants: Guillaume Iooss, Christophe Alias, Sanjay Rajopadhye [Colorado State University, USA].

Program equivalence is a well-known problem with a wide range of applications, such as algorithm recognition, program verification, and program optimization. This problem is also known to be undecidable if the class of programs is rich enough, in which case semi-algorithms are commonly used.

We focused on programs represented as systems of affine recurrence equations (SARE), defined over parametric polyhedral domains, a well-known formalism for the *polyhedral model*. SAREs include, as a proper subset, the class of affine control loop programs. Several semi-algorithms for program equivalence were already proposed for this class. Some take into account algebraic properties such as associativity and commutativity. To the best of our knowledge, none of them manage reductions, i.e., accumulations of a *parametric* number of sub-expressions using an associative and commutative operator. Our main contribution has been a new semi-algorithm to manage reductions. In particular, we outlined the ties between this problem and the perfect matching problem in a parametric bipartite graph.

This work was published at the SAS 2014 conference [7].

6.6. Analysis and Transformation of Parallel Programs

Participants: Albert Cohen [Inria/PARKAS], Alain Darte, Paul Feautrier, Abdoulaye Gamatie [CNRS/LIRMM], Laure Gonnord, Alain Ketterlin [Inria/CAMUS], Sanjay Rajopadhye [Colorado State University], Vijay Saraswat [IBM Research], Eric Violard [Inria/CAMUS], Tomofumi Yuki.

While, historically, Compsys has applied polyhedral analysis to sequential programs, it was recently realized that it also applies to parallel programs, with the aim of checking their correctness or improving their performance. The prospect of having to program exascale architectures, with their millions of cores, has led to the development of new programming languages, whose objective is to increase the programmer productivity. Compsys has applied polyhedral techniques to synchronous languages (see [25], [26] and previous activity reports), to IBM's high-productivity language X10, and, in the context of the ManycoreLabs project, to a streaming language, OpenStream, developed by Albert Cohen's group.

X10 is based on the creation of independent *activities* (light-weight threads), which can synchronize either by a generalization of the fork/join scheme, or with *clocks*, an improved version of the familiar barriers. X10 is deadlock-free by construction but it is the programmer responsibility to insure determinism by a proper use of synchronizations. Non-determinism bugs may have a very low occurrence probability thus be very difficult to detect by testing, hence the interest for detecting races at compile time. In collaboration with CSU (S. Rajopadhye, T. Yuki) and IBM (V. Saraswat), we extended array dataflow analysis to polyhedral clock-free X10 programs [29]. We have been working on clocked programs too: race detection becomes undecidable [30], but realistic problems may still be solved by heuristics.

As a side-effect of this work, we have shown in cooperation with Eric Violard and Alain Ketterlin (Inria Team Camus, Strasbourg) that clocks can be removed and replaced by `async/finish` constructs without modifying the program semantics [6]. While this transformation incurs a large overhead for general programs, in the polyhedral case the overhead is negligible, thus improving the program performance.

In contrast to X10, OpenStream is deterministic by construction, but may have deadlocks. A usual way of disproving deadlocks is by exhibiting a schedule for the program operations, a well-known problem for polyhedral programs, where dependences can be described by affine constraints. In the case of OpenStream, communications use one-dimensional channels and, in a form of linearization, give rise to polynomial dependences for polyhedral OpenStream codes. In a ManycoreLabs project deliverable (see Section 7.1), we have formalized the problem and proved that deadlock detection is undecidable in general.

6.7. Handling Polynomials for Program Analysis and Transformation

Participant: Paul Feautrier.

As shown in the previous section, many problems in parallel programs analysis and verification can be reduced to proving or disproving properties of polynomials in the variables of the program. For instance, so-called "linearizations" (replacing a multi-dimensional object by a one-dimensional one) generate polynomial access functions. These polynomials then reappear in dependence testing, scheduling, and invariant construction. This is also the case in OpenStream where nested loops act on one-dimensional streams. What is needed here is a replacement for the familiar emptiness tests and for Farkas lemma (deciding whether an affine form is positive inside a polyhedron).

Recent mathematical results by Handelman and Schweighofer on the *Positivstellensatz* allow one to devise algorithms that are able to solve these problems. The difference is that one gets only sufficient conditions, and that complexity is much higher than in the affine cases. A paper presenting applications of these ideas to three use cases – dependence testing, scheduling, and transitive closure approximation – will be presented at the 5th International Workshop on Polyhedral Compilation Techniques (IMPACT'15) [5] in Amsterdam in January 2015.

6.8. Parametric Loop Tiling with Constant Aspect Ratio

Participants: Guillaume Iooss, Christophe Alias, Sanjay Rajopadhye [Colorado State University, USA].

Parametric tiling is a well-known transformation which is widely used to improve locality, parallelism, and granularity (see also the next section for more details). However, parametric tiling is also a non-linear transformation and this prevents polyhedral analysis or further polyhedral transformation after parametric tiling. It is therefore generally applied during the code generation phase.

To address this issue, we presented a method to remain in a polyhedral representation, in a special case of parametric tiling where all the dimensions are tiled and all tile sizes are constant multiples of a single tile size parameter. We call this *Constant Aspect Ratio Tiling*. We showed how to mathematically transform a polyhedron and an affine function into their tiled counterpart, which are the two main operations needed in such a transformation.

The approach is now implemented, and has been tested successfully on several kernels commonly used in the community (matrix multiply, jacobi 1D, jacobi 2D). A corresponding paper was published at the IMPACT 2014 workshop [8].

6.9. Exact and Approximated Data-Reuse Optimizations for Tiling with Parametric Sizes

Participants: Alain Darte, Alexandre Isoard.

Loop tiling is a loop transformation widely used to improve spatial and temporal data locality, to increase computation granularity, and to enable blocking algorithms, which are particularly useful when offloading kernels on computing units with smaller memories. When caches are not available or used, data transfers and local storage must be software-managed, and some useless remote communications can be avoided by exploiting data reuse between tiles. An important parameter of tiling is the sizes of the tiles, which impact the size of the required local memory. However, for most analyses involving several tiles, which is the case for inter-tile data reuse, the tile sizes induce non-linear constraints, unless they are numerical constants. This complicates or prevents a parametric analysis with polyhedral optimization techniques.

We showed that, when tiles are executed in sequence along tile axes, the parametric (with respect to tile sizes) analysis for inter-tile data reuse is nevertheless possible, i.e., one can determine, at compile-time and in a parametric fashion, the copy-in and copy-out data sets for all tiles, with inter-tile reuse, as well as sizes for the induced local memories. When approximations of transfers are performed, the situation is much more complex, and involves a careful analysis to guarantee correctness when data are both read and written. We provide the mathematical foundations to make such approximations possible, thanks to the introduction of the concept of *pointwise functions*. Combined with hierarchical tiling, this result opens perspectives for the automatic generation of blocking algorithms, guided by parametric cost models, where blocks can be pipelined

and/or can contain parallelism. Previous work on FPGAs and GPUs already showed the interest and feasibility of such automation with tiling, but in a non-parametric fashion. Our method is currently implemented with the `iscc` calculator of ISL, a library for the manipulation of integer sets defined with Presburger arithmetic, a complete implementation within the PPCG compiler is in progress.

We believe that our approximation technique can be used for other applications linked to the extension of the polyhedral model as it turns out to be fairly powerful. Our future work will be to derive efficient approximation techniques, either because the program cannot be fully analyzable, or because approximations can speed-up or simplify the results of the analysis without losing much in terms of memory transfers and/or memory sizes.

A preliminary version of this work has been presented at the IMPACT'14 workshop [3]. A revised version has been accepted for publication at the International Conference on Compiler Construction (CC'15) [4].

6.10. Studying Optimal Spilling in the Light of SSA

Participants: Florian Brandner [ENSTA ParisTech], Quentin Colombet [Apple, Cupertino], Alain Darte.

Recent developments in register allocation, mostly linked to static single assignment (SSA) form, have shown the benefits of decoupling the problem in two phases: a first spilling phase places load and store instructions so that the register pressure at all program points is small enough, a second assignment and coalescing phase maps the variables to physical registers and reduces the number of move instructions among registers. At the end of Quentin Colombet's PhD thesis, we focused on the first phase, for which many open questions remained: in particular, we studied the notion of optimal spilling (what can be expressed?) and the impact of SSA form (does it help?). To identify the important features for optimal spilling on load-store architectures, we developed a new integer linear programming formulation, more accurate and expressive than previous approaches. Among other features, we can express SSA ϕ-functions, memory-to-memory copies, and the fact that a value can be stored simultaneously in a register and in memory. Based on this formulation, we presented a thorough analysis of the results obtained for the SPECINT 2000 and EEMBC 1.1 benchmarks, from which we drawed the following conclusions: a) rematerialization is extremely important, b) SSA complicates the formulation of optimal spilling, especially because of memory coalescing when the code is not in conventional SSA (CSSA), c) micro-architectural features are significant and thus have to be accounted for, which is not the case with standard cost functions, d) significant savings can be obtained in terms of static spill costs, cache miss rates, and dynamic instruction counts, e) however, cost models based only on static spill costs are not always relevant, in particular when spilling is "at the limit": in this situation, bad interactions with register coalescing and post-pass scheduling can be exacerbated and it may be better to spill a bit more. This important observation indicates that more research is needed to explore alternative cost models that reliably guide spilling.

Parts of this work were already published at CASES 2011. The publication at ACM Transactions on Architecture and Code Optimization [1] contains more detailed discussions, more examples illustrating new concepts and existing approaches, and additional experiments covering the observed worst-case behavior, a new post-latency heuristic, and empiric evidence showing why static spill costs are a poor metric. Three configurations were added: Appel and George under SSA, Koes and Goldstein, and the heuristic of Braun and Hack. This work was partly supported by the Mediacom contract with STMicroelectronics (ended in 2013).

CONVECS Project-Team

6. New Results

6.1. New Formal Languages and their Implementations

LNT is a next generation formal description language for asynchronous concurrent systems, which attempts to combine the best features of imperative programming languages and value-passing process algebras. LNT is increasingly used by CONVECS for industrial case studies and applications (see § 6.5) and serves also in university courses on concurrency, in particular at ENSIMAG (Grenoble) and at Saarland University.

6.1.1. Translation from LNT to LOTOS

Participants: Hubert Garavel, Frédéric Lang, Wendelin Serwe.

In 2014, the translator from LNT to LOTOS was further improved. In addition to bug fixes and removal of incorrect warnings emitted by the translator itself or by the C compiler on the generated code, the following enhancements have been brought: the LNT language was extended with a "`!representedby`" pragma for processes, and a "`only if`" statement to concisely express guarded commands; the translator now performs static analysis and warns about unused variables, unused "`in`" or "`in out`" parameters, useless (deterministic or nondeterministic) assignments to variables, "`in out`" parameters that are never assigned, and dubious synchronizations between processes; checks for underflow/overflow on natural and integer numbers are now activated by default. The translator also generates better LOTOS code, and the LNT reference manual was shortened and updated in many places.

6.1.2. Translation from LOTOS to Petri nets and C

Participants: Hubert Garavel, Wendelin Serwe.

The LOTOS compilers CAESAR and CAESAR.ADT, which were once the flagship of CADP, now play a more discrete role since LNT (rather than LOTOS) has become the recommended specification language of CADP. Thus, CAESAR and CAESAR.ADT are mostly used as back-end translators for LOTOS programs automatically generated from LNT or other formalisms such as Fiacre, and are only modified when this appears to be strictly necessary.

In 2014, the CAESAR compiler has been modified to tolerate LOTOS specifications that would be normally rejected under the ISO/IEC 8807 standard definition of LOTOS. The first change extends the visibility scope of local definitions when the global definitions are empty. The second change uses the type information of process definitions to better resolve overloading ambiguities in expressions passed as actual parameters to process calls.

Conversely, CAESAR was made stricter by rejecting at compile-time LOTOS specifications containing out-of-bound constants, even if such constants are never used.

Performance has been increased by adding or strengthening a number of optimizations concerning, e.g., internal data structures, Boolean guards that can be statically evaluated, values belonging to singleton sorts, disconnected or otherwise unreachable Petri net places and transitions, etc.

The CAESAR.BDD tool of CADP, which analyzes hierarchical Petri nets generated from higher-level specifications (e.g., LOTOS or LNT) has been significantly enhanced. The semantic model accepted by CAESAR.BDD has been made more general and given the new name of NUPN (*Nested-Units Petri Nets*). The definition and theoretical properties of NUPN have been formalized.

The textual syntax for NUPN has been extended with pragmas intended to retain useful properties of non-ordinary and/or non-safe Petri nets translated to NUPN. An XML syntax for NUPN (compatible with the ISO standard PNML for the representation of Petri nets) has been defined and adopted by the Model Checking Contest [0]. A translator from PNML to NUPN has been developed at LIP6 (Paris, France).

---

[0]http://mcc.lip6.fr/nupn.php

The CAESAR.BDD tool has been updated accordingly, and extended to perform stricter checks and compute more structural and behavioral properties of NUPN models. CAESAR.BDD has been intensively used to correct the descriptions of the Model Checking Contest benchmarks: a first campaign (January-February 2014) detected 9 errors in structural properties and 8 errors in behavioral properties, and a second campaign (April 2014) revealed 23 more errors. CAESAR.BDD has also been used to automatically generate new benchmarks, together with their descriptions.

6.1.3. *Translation from GRL to LNT*

Participants: Fatma Jebali, Frédéric Lang, Eric Léo, Radu Mateescu.

In the context of the Bluesky project (see § 8.1.2.1), we study the formal modeling of GALS (*Globally Asynchronous, Locally Synchronous*) systems, which are composed of several synchronous subsystems evolving cyclically, each at its own pace, and communicating with each other asynchronously. Designing GALS systems is challenging due to both the high level of (synchronous and asynchronous) concurrency and the heterogeneity of computations (deterministic and nondeterministic). To bring our formal verification techniques and tools closer to the GALS paradigm, we designed a new formal language named GRL (*GALS Representation Language*), as an intermediate format between GALS models and purely asynchronous concurrent models. GRL combines the main features of synchronous dataflow programming and asynchronous process calculi into one unified language, while keeping the syntax homogeneous for better acceptance by industrial GALS designers. GRL allows a modular composition of synchronous systems (blocks), environmental constraints (environments), and asynchronous communication mechanisms (mediums), to be described at a level of abstraction that is appropriate to verification. GRL also supports external C and LNT code.

In 2014, we have continued to enhance the syntax and the formal semantics of GRL. We have written a detailed research report (82 pages) [25] containing the complete definition of the syntax, static semantics, and dynamic semantics (in the form of structural operational semantics rules), and also illustrating the checking of dynamic semantics rules on several examples of GRL programs. A paper describing GRL has been published in an international conference [14].

To equip GRL with verification features, we formally defined a translation from GRL to LNT. GRL blocks are translated into LNT functions, possibly encapsulated within LNT wrapper processes to enable asynchronous communication, whereas GRL environments and mediums are directly translated into LNT processes. The asynchronous composition of blocks, environments, and mediums is translated to an LNT parallel composition of the corresponding processes.

Using the SYNTAX and LOTOS NT compiler construction technology [44], we have developed a translator named GRL2LNT (about 25,000 lines of code), allowing an LNT program to be generated from a GRL specification automatically. GRL2LNT performs the lexical and syntactic analysis of GRL programs, together with almost all static semantic checks specified in its formal definition [25]. A stable version of GRL2LNT has been released in 2014. Additionally, we have developed an OPEN/CAESAR-compliant compiler GRL.OPEN (based on GRL2LNT and LNT.OPEN), which makes possible the on-the-fly exploration of the LTS underlying a GRL specification using CADP. We have also built a test base containing about 250 (correct and incorrect) GRL programs, and used it for non-regression testing of GRL2LNT. The correct GRL programs represent about 7,000 lines of code and produce about 18,000 lines of LNT code after translation using GRL2LNT.

A paper describing the formal verification of GALS systems using GRL and CADP, with a focus on the translation from GRL to LNT, has been submitted to an international conference [28].

6.1.4. *Coverage Analysis for LNT*

Participants: Gwen Salaün, Lina Ye.

In the classic verification setting, the designer has a specification of a system in a value-passing process algebra, a set of temporal properties to be verified on the corresponding LTS model, and a data set of examples (test cases) for validation purposes. At this stage, building the set of validation examples and debugging the specification is a complicated task, in particular for non-experts.

We propose a new framework for debugging value-passing process algebra through coverage analysis and we illustrate our approach with LNT. We define several coverage notions before showing how to instrument the specification without affecting original behaviors. Our approach helps one to improve the quality of a data set of examples used for validation purposes, but also to find ill-formed decisions, dead code, and other errors in the specification. We have implemented a tool for automating our debugging approach, and applied it to several real-world case studies in different application areas.

In 2014, a paper has been accepted in an international conference [19].

6.1.5. *Other Language Developments*

Participants: Hugues Evrard, Hubert Garavel, Frédéric Lang, Eric Léo, Wendelin Serwe.

The ability to compile and verify formal specifications with complex, user-defined operations and data structures is a key feature of the CADP toolbox since its very origins. A long-run effort has been recently undertaken to ensure a uniform treatment of types, values, and functions across all the various CADP tools.

In 2014, convergence between the LOTOS, LNT, BCG, and XTL data-type libraries has been increased by defining common libraries for eight predefined types: Boolean, Natural, Integer, Real, Character, String, Raw, and Gate. These libraries gather in the same place definitions of types, constants, and functions that were previously disseminated across different tools. Additionally, systematic checks for underflows and overflows have been set for the Natural and Integer types. Also, unprintable characters and C-like escape sequences are now uniformly handled by the Character, String, and Raw types.

To support the LNT language in the Emacs/XEmacs, jEdit, and Vim editors, configuration files have been added or updated, which provide for syntax highlighting/coloring, and enable autocompletion in Emacs using YASnippet.

6.2. Parallel and Distributed Verification

6.2.1. *Distributed Code Generation for LNT*

Participants: Hugues Evrard, Frédéric Lang.

Rigorous development and prototyping of a distributed verification algorithm in LNT involves the automatic generation of a distributed implementation. For the latter, a protocol realizing process synchronization is required. As far as possible, this protocol must itself be distributed, so as to avoid the bottleneck that would inevitably arise if a unique process would have to manage all synchronizations in the system. A particularity of such a protocol is its ability to support *branching synchronizations*, corresponding to situations where a process may offer a choice of synchronizing actions (which themselves may nondeterministically involve several sets of synchronizing processes) instead of a single one. Therefore, a classical barrier protocol is not sufficient and a more elaborate synchronization protocol is needed.

Using a synchronization protocol that we verified formally in 2013, we developed a prototype distributed code generator, named DLC (*Distributed LNT Compiler*), which takes as input the model of a distributed system described as a parallel composition of LNT processes.

In 2014, we continued the development of DLC. We improved the performances of DLC generated code by reducing the number of protocol messages when one or several processes are ready on a single gate. We experimented this optimization on a set of processes running on different computers and synchronizing all together on a single barrier interaction (i.e., all processes are ready on a single gate). In this situation, DLC now generates code that is faster than Java or Erlang.

The distributed program generated by DLC would be of little interest if it could not interact with its environment (e.g., users through human-computer interfaces, or other systems, such as databases, Web services, etc.). Therefore, we designed a mechanism to embed user-defined C functions, called *hook functions*, into the code generated by DLC. Hook functions are triggered on events related to actions in the system. This allows system actions to be, e.g., monitored by the user or controlled by external conditions. Using hook functions, the code generated by DLC can thus both take an account of and have an effect on its environment.

In order to demonstrate DLC on a real-world example, we applied it to the recent Raft [0] consensus algorithm [60]. We wrote an LNT specification of a simple key-value store made fault tolerant by replication of commands using the Raft consensus algorithm. During the modeling phase, we found a missing transition in the TLA+ specification of the protocol. We signaled it to the authors [0], who corrected the TLA+ specification. We used hook functions to implement interaction with the replicated store from external clients. The distributed implementation generated by DLC was successfully tested on clusters of the Grid5000 platform. We presented an overview of DLC, the hook functions and the Raft experiment in an article that has been accepted for publication in an international conference [12].

6.3. Timed, Probabilistic, and Stochastic Extensions

6.3.1. *Model Checking for Extended PCTL*

Participants: Hubert Garavel, Radu Mateescu, Jose Ignacio Requeno.

In the context of the SENSATION project (see § 8.2.1.1), we study the specification and verification of quantitative properties of concurrent systems.

In 2014, we defined an extension of PCTL (*Probabilistic Computation Tree Logic*) [49] with the manipulation of data values and actions. This logic is interpreted on extended DTMCs (*Discrete-Time Markov Chains*) containing visible transitions, labeled with channel names and data values, in addition to probabilistic transitions. Extended PCTL makes possible the specification of temporal properties involving discrete time, probabilities, and data values.

We devised a prototype model checker for extended PCTL in the form of an XTL library describing the denotational semantics of all PCTL operators (both primitive and derived ones), accompanied by external C code implementing the algorithms for LTS exploration and numerical computation of probabilities. The high-level programming language constructs of XTL (iterators, sets in comprehension, parameterized macro-definitions) allowed us to easily implement the advanced features (filters on arithmetic and logical operators, computation of probabilities, experiments over data series, etc.) of established probabilistic model checkers, such as PRISM [54]. Also, the manipulation of data values in XTL allows one to specify properties in which probabilities and discrete time deadlines depend on the values of state variables, a feature currently not provided by PRISM.

To experiment and cross-check our extended PCTL library w.r.t. PRISM, we developed an automated translator from the (state-based) DTMCs used by PRISM into the (action-based) DTMCs in BCG format used by CADP. State information is represented by means of special self-looping transitions containing the values of state variables, which are properly handled during the evaluation of probabilistic temporal operators.

The experiments we performed with our extended PCTL library on various examples of DTMCs (produced from communication protocols, chemical reactions, hazard games, etc.) showed a performance comparable to (explicit-state) PRISM for pure PCTL formulas.

Furthermore, in addition to many bug fixes, the XTL compiler and its XTL_EXPAND preprocessor have been strengthened to better detect and report potential mistakes in source XTL specifications. In particular, vacuity checks have been introduced, which warn the user when no label in a BCG graph has the right number of fields or the appropriate field types to satisfy an XTL label match expression (previously, this expression would silently evaluate to false).

The type checking system of XTL and its list of predefined functions have been extended to support the new Natural and Raw types of the BCG format, and to properly distinguish between Natural and Integer values, and Raw and String values, while achieving a high degree of backward compatibility. In particular, XTL now uses type information from the BCG labels to better solve overloading in label offers, so that certain XTL programs that were formerly invalid are now accepted. Finally, it is now possible to use the predefined types and functions of XTL when defining temporal operators directly using external C code.

---

[0]http://raftconsensus.github.io
[0]https://groups.google.com/forum/#!topic/raft-dev/yu-wOUx-gnA

6.4. Component-Based Architectures for On-the-Fly Verification

6.4.1. *Property-Dependent Reductions for the Modal Mu-Calculus*
Participant: Radu Mateescu.

In collaboration with Anton Wijs (Technical University of Eindhoven), we proposed a new method for enhancing the performance of model checking a temporal formula on an LTS by reducing the LTS as much as possible depending on the formula prior to (or simultaneously with) the verification. Given an LTS and a formula, the method consists of two steps:

- The maximal set of actions that one can hide (i.e., rename into the internal action τ) in the LTS without disturbing the interpretation of the formula is computed, and those actions are hidden in the LTS. This works for any formula of the full modal μ-calculus (i.e., of arbitrary alternation depth) and provides the highest potential for reducing the LTS, and hence for improving verification performance, w.r.t. that formula.

- The LTS is reduced modulo an equivalence relation preserving the formula. The reduction can be done before verification, either by constructing the LTS explicitly and using the direct minimization features provided by the BCG_MIN tool, or by constructing the minimized LTS incrementally using the compositional verification features provided by EXP.OPEN and SVL. The reduction can be also done simultaneously during verification, by detecting τ-confluent transitions and prioritizing them against their neighbors.

We defined a μ-calculus fragment, named Lμ-dsbr, and shown its adequacy w.r.t. divergence-sensitive branching bisimulation (divbranching for short). We also shown that Lμ-dsbr is equally expressive to the μ-ACTL$\diagdown$X logic, an extension of ACTL$\diagdown$X (Action-based CTL without the next time operator) with fixed point operators [39], [40]. This result also implies the adequacy w.r.t. divbranching of μ-ACTL$\diagdown$X, which was previously shown to be adequate w.r.t. strong bisimulation.

We experimented our method using the EVALUATOR model checker on various examples of protocols and distributed systems, by specifying the temporal properties in MCL and reducing the LTSs modulo strong and divbranching bisimulation. The experiments showed performance enhancements both in execution time (reduction by a factor 4 for strong bisimulation and 20 for divbranching) and memory consumption (reduction by a factor 2 for strong bisimulation and 5 for divbranching).

We also built a prototype MCL library regrouping the temporal operators of ACTL$\diagdown$X (which were already present in CADP) and the modal and temporal operators of Lμ-dsbr (which were newly added). Used in conjunction with the Boolean and fixed point operators of MCL, the operators of this library can be used to specify temporal formulas adequate w.r.t. divbranching, which allows one to reduce the LTS modulo this equivalence (after applying maximal hiding) and to increase the performance of verification accordingly. An article has been published in an international journal [8].

6.4.2. *Compositional Verification*
Participants: Hubert Garavel, Frédéric Lang.

The CADP toolbox contains various tools dedicated to compositional verification, among which EXP.OPEN, BCG_MIN, BCG_CMP, and SVL play a central role. EXP.OPEN explores on the fly the graph corresponding to a network of communicating automata (represented as a set of BCG files). BCG_MIN and BCG_CMP respectively minimize and compare behavior graphs modulo strong or branching bisimulation and their stochastic extensions. SVL (*Script Verification Language*) is both a high-level language for expressing complex verification scenarios and a compiler dedicated to this language.

In 2014, we corrected 2 bugs in EXP.OPEN, 6 bugs in BCG_MIN and BCG_CMP, and 5 bugs in SVL. We also enhanced these tools as follows:

- We corrected the diagnostic generation algorithm of BCG_CMP, which sometimes generated irrelevant diagnostics.

- We improved the messages displayed by SVL and EXP.OPEN when generating an LTS from a composition expression using the *smart reduction* strategy [38], so that the user can follow more easily the selected composition order.

- Following the recent progress made on the development of the language LNT (see § 6.1), the syntax of the SVL and EXP languages for comments, gate typing, and the "`par`", "`hide`", "`rename`", "`cut`", and "`prio`" operators was extended to be compatible with the syntax of LNT. This enables composition expressions (including comments, channel typing, etc.) copied from LNT programs to be pasted in SVL scripts while requiring as few syntactic changes as possible.

- The "`verify`" operator has been generalized to give access to all three model checkers of CADP (EVALUATOR 3, EVALUATOR 4, and XTL). A new statement "`|=`" has been added to SVL, which enables MCL and XTL formulas to be directly written in an SVL script, rather than being stored in external files.

- To provide for requirements expression and traceability in SVL, we introduced two new statements, "`property`" and "`check`", which increase the readability and good structure of SVL scripts by allowing to define and verify properties, each of which is given a name, instantiable parameters, an informal textual description, and (optionally) an expected truth value.

- We updated several demo examples of CADP in order to illustrate the above extensions.

6.4.3. *On-the-Fly Test Generation*

Participants: Hubert Garavel, Radu Mateescu, Wendelin Serwe.

In the context of the collaboration with STMicroelectronics, we study techniques for testing if a (hardware) implementation is conform to a formal model described in LNT. Our approach is inspired by the theory of conformance testing [62], as implemented for instance in TGV [53] and JTorX [33]. We have developed two prototype tools to support this approach. The first tool implements a dedicated OPEN/CAESAR-compliant compiler for the particular asymmetric synchronous product between the model and the test purpose. The second tool, based on slightly extended generic components for graph manipulation (τ-compression, τ-confluence reduction, determinization) and resolution of Boolean equation systems, generates the complete test graph (CTG), which can be used to extract concrete test cases or to drive the test of the implementation. The principal advantage of our approach compared to existing tools is the use of LNT for describing test purposes, which facilitates the manipulation of data values.

In 2014, we developed a third prototype tool that takes as input a CTG and extracts either a single test case (randomly chosen or the first encountered one), or the set of *all* test cases. This prototype tool was used in the case study with STMicroelectronics (see § 6.5.1).

The test-generation tool TGV has been streamlined by removing some obsolete options and replacing a large part of its code by calls to the standard CADP libraries. TGV has been made faster, it now supports the latest version of the AUT format, and ensures that test purposes provided in the BCG format are deterministic. The manual page has been updated and completed.

6.4.4. *Other Component Developments*

Participants: Soraya Arias, Hubert Garavel, Frédéric Lang, Radu Mateescu.

The AUT textual format for CADP for storing LTSs was extended to support recent languages (such as LNT and the PseuCo language developed at Saarland University) that manipulate character-string values. The AUT format, which was defined in the late 80s, did not support such values. A new version 2014 of the AUT format has been defined, which solves this problem and maintains backward compatibility. All the CADP tools that read or write AUT files have been updated accordingly.

The BCG format of CADP for storing LTSs has been upgraded with the advent of a new version 1.2, which replaces version 1.1 released in 2009. New predefined types have been added to BCG to express the difference between unsigned and signed integers, and between character strings and untyped raw-data values. The new version of the BCG format is also more compact and now uses variable-length encoding for strings. The rules for label parsing of the BCG_WRITE interface have been extended, and BCG_IO now supports version 2014 of the AUT format. The intrinsic difficulty of these changes was to preserve the backward compatibility with the BCG files generated over the last twenty years.

To simplify the installation of CADP on Windows systems, we studied an alternative execution environment based on Gnuwin32 and MinGW/Msys rather than Cygwin. Preliminary changes have been brought to CADP scripts to undertake such a migration.

6.5. Real-Life Applications and Case Studies

6.5.1. *ACE Cache Coherency Protocol*

Participants: Abderahman Kriouile, Radu Mateescu, Wendelin Serwe.

In the context of a CIFRE convention with STMicroelectronics, we study system-level cache coherency, a major challenge faced in the current System-on-Chip architectures. Because of their increasing complexity (mainly due to the significant number of computing units), the validation effort using current simulation-based techniques grows exponentially. As an alternative, we study formal verification.

We focused on the ACE (AXI Coherency Extensions) cache coherency protocol, a system-level coherency protocol proposed by ARM [29]. In previous years, we developed a formal LNT model (about $3,400$ lines of LNT) of a system consisting of an ACE-based cache coherent interconnect, processors, and a main memory. The model is parametric and can be instantiated with different configurations (number of processors, number of cache lines, number of memory lines) and different sets of supported elementary ACE operations (currently, a representative subset of 15 operations), including an abstract operation that represents any other ACE operation. We handled the global requirements of the ACE specification using a constraint oriented programming style, i.e., by representing each global requirement as a dedicated process observing the global behavior and inhibiting incorrect executions. We also specified temporal properties expressing cache coherence, data integrity, and successful completion of each transaction. Note that the former property required to transform state-based properties into action-based ones, by adding information about the cache state to the actions executed by the cache.

In 2014, we exploited the formal model to improve the validation of the architecture under design at STMicroelectronics. In a first step, we studied the sanity (soundness and completeness) of an industrial interface verification unit, consisting of a list of so-called *formal checks*. After modeling each check in LNT, we used the BISIMULATOR tool to verify that each check is an overapproximation of the corresponding projection of the formal model. When we tried to establish that the parallel composition of all checks is an overapproximation of the projection of the formal model, we discovered a missing check (a particular channel did not occur in any of the checks).

In a second step, we studied the derivation of system level test cases, using a two-phase approach:

- In the first phase, abstract test cases were extracted automatically from the formal model using a prototype tool (see § 6.4). To circumvent the complexity of extracting test cases from the complete model, we proposed an iterative approach based on the automatic selection of a comprehensive set of interesting scenarios leading to LTSs of tractable size. The selection of the interesting scenarios relies on the counterexamples provided by the EVALUATOR model checker for the properties of coherence and data integrity.

- In the second phase, the abstract test cases were translated into the input format of an industrial test bench in charge of refining them into concrete test cases to be executed on the RTL (*Register Transfer Level*) description of the architecture under study. Experiments with manually translated abstract test cases led to the early discovery of bugs in commercial verification blocks, which could therefore be corrected before their use became critical in the development process.

The tests derived from the formal model increased the coverage of problematic features of some blocks used in the architecture. In particular, our approach was able to detect a limitation concerning data integrity 20 months before it was confirmed by classical methods, and our methodology provides all the scenarios triggering the limitation.

This work led to a publication accepted in an international conference [15]. Also, a large Petri net derived from our LNT model was provided as benchmark example for the Model Checking Contest.

6.5.2. *Formal Verification of BPMN Processes*
Participants: Radu Mateescu, Gwen Salaün, Lina Ye.

A business process is a set of structured, related activities that aims at fulfilling a specific organizational goal for a customer or market. An important metric when developing a business process is its degree of parallelism, i.e., the maximum number of tasks that are executable in parallel in that process. The degree of parallelism determines the peak demand on tasks, providing a valuable guide for the problem of resource allocation in business processes.

In 2014, we investigated how to automatically measure the degree of parallelism for business processes, described using the BPMN standard notation. To this aim, we defined a formal model for BPMN processes in terms of LTSs, which are obtained through an encoding in LNT. We then proposed an approach for automatically computing the degree of parallelism by using model checking of parameterized MCL formulas and dichotomic search. We developed a prototype tool for automating this check and we applied it successfully to more than one hundred BPMN processes.

This work led to a publication in an international conference [16].

6.5.3. *Stability of Asynchronously Communicating Systems*
Participants: Gwen Salaün, Lina Ye.

Analyzing communicating systems that interact asynchronously via reliable FIFO buffers is an undecidable problem. A typical approach is to check whether the system is bounded, and if not, the corresponding state space can be made finite by limiting the presence of communication cycles in behavioral models or by fixing buffer sizes.

We followed a different approach, which aims at analyzing communicating systems without restricting them by imposing any arbitrary bounds. These systems are likely to be unbounded and therefore result in infinite state spaces. We introduce a notion of stability and prove that once the system is stable for a specific buffer bound (called stability bound), it remains stable whatever larger bounds are chosen for the buffers. This enables us to check certain properties on the (finite-state) system obtained for the stability bound and to ensure that the system will preserve them whatever larger bounds are used for buffers.

We have also proven that computing the stability bound is in general undecidable, and we proposed a semi-algorithm that successfully computes the stability bounds for many typical examples of communicating systems using heuristics and equivalence checking. This work is described in a research report [27].

6.5.4. *Deployment and Reconfiguration Protocols for Cloud Applications*
Participants: Rim Abid, Gwen Salaün.

In the context of the OpenCloudware project (see § 8.1.1.1), we collaborate with Noël de Palma and Fabienne Boyer (University Joseph Fourier), Xavier Etchevers and Thierry Coupaye (Orange Labs) in the field of cloud computing applications, which are complex distributed applications composed of interconnected software components running on distinct virtual machines (VMs). Setting up, (re)configuring, and monitoring these applications involve intricate management protocols, which fully automate these tasks while preserving application consistency as well as some key architectural invariants.

In 2014, we extended the specification of the self-deployment protocol to support VM failures. This led to a publication in an international conference [11], of which an extended version is under preparation for submission to an international journal.

We also worked on the dynamic reconfiguration of cloud applications. As a first attempt, we proposed to design this protocol using a publish-subscribe communication model [32]. In 2014, we improved the protocol to also support VM failures, and drastically validated the corresponding LNT specification using model checking. A paper presenting these results was submitted to an international journal. In parallel, we studied a version of this protocol where the different participants interact asynchronously via FIFO buffers. This led to a publication in an international conference [10].

As a new line of work, we undertook the study of controller synthesis techniques for the coordination of autonomic managers in asynchronous environments. Our approach relies on an encoding into LNT and on the application of several operations on automata (synchronous products, hiding, reduction) for synthesizing the corresponding controller using CADP tools. We also proposed automated techniques for generating Java code from an abstract model of the controller. For validation purposes, we applied our approach to real-world three-tier Web applications and showed that the introduction of a controller allows one to avoid erroneous situations due to the absence of coordination between autonomic managers.

6.5.5. *Networks of Programmable Logic Controllers*
Participants: Hubert Garavel, Fatma Jebali, Jingyan Jourdan-Lu, Frédéric Lang, Eric Léo, Radu Mateescu.

In the context of the Bluesky project (see § 8.1.2.1), we study the software applications embedded on the PLCs (*Programmable Logic Controllers*) manufactured by Crouzet Automatismes. One of the objectives of Bluesky is to enable the rigorous design of complex control applications running on several PLCs connected by a network. Such applications are instances of GALS (*Globally Asynchronous, Locally Synchronous*) systems composed of several synchronous automata embedded on individual PLCs, which interact asynchronously by exchanging messages. A formal analysis of these systems can be naturally achieved by using the formal languages and verification techniques developed in the field of asynchronous concurrency.

For describing the applications embedded on individual PLCs, Crouzet provides a dataflow language with graphical syntax and synchronous semantics, equipped with an ergonomic user-interface that facilitates the learning and use of the language by non-experts. To equip the PLC language of Crouzet with functionalities for automated verification, the solution adopted in Bluesky was to translate it into GRL (see § 6.1.3), which enables the connection to testing and verification tools covering the synchronous and asynchronous aspects.

In 2014, we have developed a set of GRL libraries implementing about 40 of the function blocks present in the PLC programming tool of Crouzet, to facilitate the integration of GRL in the PLC software design process. These function blocks include (among others) logic and comparison functions, timers, triggers, and counters. These GRL libraries have been used to model large applications provided by Crouzet. The GRL2LNT and GRL.OPEN tools (see § 6.1.3) provide a direct connection to all verification functionalities of CADP, in particular model checking and equivalence checking.

Regarding model checking, we have studied existing work in the verification of synchronous systems and GALS systems. We have identified a set of typical patterns of temporal properties (e.g., deadlocks, safety, liveness) relevant for GALS systems. These property patterns have been specified using MCL and checked on a set of feature-rich GRL examples using GRL.OPEN and EVALUATOR.

Regarding equivalence checking, the purpose is to compare the behavior of a GALS system with its *service*, which represents its desired observable behavior, modulo a suitable equivalence relation. We have studied existing work in equivalence checking for GALS systems and we have investigated how to formally define the expected service of a GALS system at the appropriate level of expressiveness and abstraction, which requires a careful identification of the observable actions corresponding to the interactions between the GALS system and its physical environment. We have modeled several examples of GALS systems in GRL, and experimented the definition of appropriate services and their usage for equivalence checking by means of GRL.OPEN and BISIMULATOR.

The validation approach we promote, together with our colleagues from the LCIS laboratory (Valence) in the Bluesky project, led to a common publication in a national conference [21].

6.5.6. *EnergyBus Standard for Connecting Electric Components*

Participants: Hubert Garavel, Wendelin Serwe.

The EnergyBus [0] is an upcoming industrial standard for electric power transmission and management, based on the CANopen field bus. It is developed by a consortium assembling all major industrial players (such as Bosch, Panasonic, and Emtas) in the area of light electric vehicles (LEV); their intention is to ensure interoperability between all electric LEV components. At the core of this initiative is a universal plug integrating a CAN-Bus [0] with switchable power lines. The central and innovative role of the EnergyBus is to manage the safe electricity access and distribution inside an EnergyBus network.

In the framework of the European FP7 project SENSATION (see § 8.2.1.1) a formal specification in LNT of the main EnergyBus protocols is being developed by Alexander Graf-Brill and Holger Hermanns at Saarland University [48], with the active collaboration of CONVECS.

In 2014, our joint work with Saarland University on the modeling, verification, and test case generation for the EnergyBus standard led to a common publication [13].

6.5.7. *Graphical User-Interfaces and Plasticity*

Participants: Hubert Garavel, Frédéric Lang, Raquel Oliveira.

In the context of the Connexion project (see § 8.1.1.2) and in close collaboration with Gaëlle Calvary, Eric Ceret, and Sophie Dupuy-Chessa (IIHM team of the LIG laboratory), we study the formal description and validation of graphical user-interfaces using the most recent features of the CADP toolbox. The case study assigned to LIG in this project is a prototype graphical user-interface [36] designed to provide human operators with an overview of a running nuclear plant. The main goal of the system is to inform the operators about alarms resulting from faults, disturbances, or unexpected events in the plant. Contrary to conventional control rooms, which employ large desks and dedicated hardware panels for supervision, this new-generation interface uses standard computer hardware (i.e., smaller screen(s), keyboard, and mouse), thus raising challenging questions on how to best provide synthetic views of the plant status. Another challenge is to introduce plasticity in such interface, so as to enable several supervision operators, including mobile ones outside of the control room, to get accurate information in real time.

We formally specified this new-generation interface in LNT, encompassing not only the usual components traditionally found in graphical user-interfaces, but also a model of the physical world (namely, a nuclear reactor with various fault scenarios) and a cognitive model of a human operator in charge of supervising the plant. Also, several desirable properties of the interface have been expressed in MCL and verified on the LNT model using CADP. This led to a publication in an international conference [17].

In 2014, we continued our activity along several directions. The LNT specification was matured in various respects. As a result of several interactions with EDF, the specification was enhanced with a more realistic representation of the plant (currently $5,358$ lines of LNT code). Besides, new desirable properties of the user-interfaces emerged with the evolution of the formal model, making a total of seven complex properties formally specified in MCL.

We initiated an integration of our formal model with an industrial control room prototype, provided by a partner in the project. To this aim, several improvements were done in the formal specification, and the integration is currently in progress.

We started to address the introduction of plasticity in the formal specification, a challenge that was identified in 2013. Plasticity is the capacity of a user-interface to withstand variations of its context of use (i.e., platform, user, environment) while preserving usability. We proposed two approaches to introduce plasticity in the analysis. The first one introduces in the formal model a representation of a plasticity engine (responsible for user-interfaces adaptation) and applies model checking to verify its properties. The second approach consists in formally specifying several versions of the user-interfaces, derived from adaptation, and applying equivalence checking to verify similarity relations on the user-interface models.

---

[0]http://www.energybus.org
[0]http://www.can-cia.org

6.5.8. *Fault-Tolerant Routing for Network-on-Chip Architectures*
Participant: Wendelin Serwe.

Fault-tolerant architectures provide adaptivity for on-chip communications, but also increase the complexity of the design, so that formal verification techniques are needed to check their correctness. In collaboration with Chris Myers and Zhen Zhang (University of Utah, USA), we studied an extension of the link-fault tolerant Network-on-Chip (NoC) architecture introduced by Wu *et al* [67] that supports multiflit wormhole routing.

To keep the state space manageable, the formal LNT model of the routing algorithm was constructed in several steps, applying different abstractions (structural and related to data). This modeling process led to several insights. First, it led to the discovery of a package leakage path that could lead to the complete loss of a packet and a deadlock. This error in the design of an arbiter was corrected in the subsequent models. Second, a buffering capacity in an arbiter was found to be crucial; this insight also led to a redesign of the arbiters. The resultant changes on the router and arbiter models uncovered interesting symmetries. Finally, we studied how deadlock freedom and tolerance of a single-link fault can be verified for a NoC architecture.

This work led to a publication in an international conference [20].

6.5.8.1. Other Case Studies

The demo examples of CADP, which have been progressively accumulated since the origins of the toolbox, are a showcase for the multiple capabilities of CADP, as well as a test bed to assess the new features of the toolbox. In 2014, the effort to maintain and enhance these demos has been pursued. The progressive migration to LNT has continued, by translating certain demos from LOTOS to LNT. Many demos have been enriched with value-passing temporal formulas that illustrate the conciseness and expressiveness of MCL and the capabilities of the EVALUATOR 4 model checker. Finally, many demos have been shortened and made more readable by using the new features of SVL, especially the "`property`" and "`|=`" statements that allow formulas to be gathered in a single SVL file rather than disseminated in a collection of MCL or XTL files.

<div align="center">

DICE Team

</div>

5. New Results

5.1. The economy of intermediation

We have presented in [6] an introductory panorama on the disruption of the intermediation revolution. Our efforts to measure data flows in the world, have been pursued [2] to estimate the concentration of the data industry. It is well known that the main platforms of the Web are concentrated in a few countries, mostly in the USA. Some countries, mostly in Asia, such as China, Russia, Korea or Japan have successfully developed their own Web 2.0 industry, while others, such as European countries, have failed to do so. We have explored in [7] the strategy of China, which has the largest Web industry behind the US and has made a priority of keeping its data at home, with systems in all activity sectors developed in general only one or two years after their main American counterparts. The innovation strategy of China aims in all fields to achieve technological independence, with at most 30% of foreign IP.

The rise of the economy of data disrupts values, such as privacy, and the way we think about our visibility. In [9], we investigate the digital world from an ethical perspective and a computer science viewpoint. We assess the structure and the dynamic of digital visibility and propose a model-driven approach to handle visibility in service compositions.

5.2. Architecture design for intermediation platforms

During our joint work with Worldline we built a JavaScript compiler for generating dataflow program from plain standard JavaScript sources. In an ACM Middleware conference poster session we raised the question of extracting a dataflow design from JavaScript callback hell. The compiler https://github.com/etnbrd/due-compiler is used to help JavaScript standard developers generate their equivalent dataflow scheme without the need of external libraries such as Promises, Async or Q. With this tool, developer may migrate their javascript legacy code towards a new flow based design. Our due npm module https://github.com/etnbrd/due is a first step towards a dynamic flow based architecture studied in Etienne's project.

The C3PO project provides a browser based application for interacting with other nearby participants in chat mode. The client architecture runs exclusively in the browser over a DTN layer and listens to posts send through a dedicated spontaneous and ephemeral social network (SESN) [5]. The client is organized around a display canvas hosting plugins. Each plugin registers for some tags it wishes to handle. The local DTN manager receives posts and propagates them to the plugins.

We have used intermediation technologies for voting systems. A brief presentation of our motivations has been made in [4]. A patent on the BitBallot protocol is on its way.

6. New Results

6.1. Highlights of the Year

Graduate Research Award of the OSU department in 2015 for Venmugil Elango (co-advised by Fabrice Rastello)

6.2. An interval constrained memory allocator for a GAS runtime

Participants: François Gindraud, Fabrice Rastello, Albert Cohen [ENS Ulm].

This work presents a memory allocator for global address space (GAS) runtime targeting distributed memory embedded architectures (MPSoC). MPSoC we are interested in are relatively new architectures, composed of several nodes with multiple general purpose cores and a local memory, linked by a network, all on one chip (NoC). They have promising energy and computing performances, but are hard to program due to the multilevel parallelism and the hardware constraints (limited memory, network structure). Existing programming framework are either thin but let the programmer do the hard choices (OpenMP + MPI) or heavy and automatic but target specific kind of applications on big systems (Global Arrays).

Givy 5.1 is a runtime currently developed to execute dynamic task graphs with data-flow dependencies on MPSoC. It has a focus on supporting irregular applications, using the dependencies to perform data-aware dynamic task scheduling and data transfer. Data blocks live in a GAS, and thus requires a GAS-aware memory allocator to avoid address collisions when they are dynamically allocated. The allocator implementation proposed in this paper does this with zero synchronization between nodes, while being memory efficient in the small distributed memories, and fast on each multithreaded node.

This work will be submitted at ACM ISMM Symposium.

6.3. A Framework for Enhancing Data Reuse via Associative Reordering

Participants: Kevin Stock [OSU], Martin Kong [OSU], Tobias Grosser [ENS Ulm], Louis-Noël Pouchet [UCLA], Fabrice Rastello, J. Ramanujam [LSU], P. Sadayappan [OSU].

The freedom to reorder computations involving associative operators has been widely recognized and exploited in designing parallel algorithms and to a more limited extent in optimizing compilers.

In this work, we develop a novel framework utilizing the associativity and commutativity of operations in regular loop computations to enhance register reuse. Stencils represent a particular class of important computations where the optimization framework can be applied to enhance performance. We show how stencil operations can be implemented to better exploit register reuse and reduce load/stores. We develop a multi-dimensional retiming formalism to characterize the space of valid implementations in conjunction with other program transformations. Experimental results demonstrate the effectiveness of the framework on a collection of high-order stencils.

This work is the fruit of the collaboration 8.1 with OSU and has been presented at the conference ACM PLDI'14.

6.4. Beyond Reuse Distance Analysis: Dynamic Analysis for Characterization of Data Locality Potential

Participants: Naznin Fauzia [OSU], Venmugil Elango [OSU], Mahesh Ravishankar [OSU], J. Ramanujam [LSU], Fabrice Rastello, Atanas Routnev [OSU], Louis-Noël Pouchet [UCLA], P. Sadayappan [OSU].

Emerging computer architectures will feature drastically decreased flops/byte (ratio of peak processing rate to memory bandwidth) as highlighted by recent studies on Exascale architectural trends. Further, flops are getting cheaper while the energy cost of data movement is increasingly dominant. The understanding and characterization of data locality properties of computations is critical in order to guide efforts to enhance data locality.

Reuse distance analysis of memory address traces is a valuable tool to perform data locality characterization of programs. A single reuse distance analysis can be used to estimate the number of cache misses in a fully associative LRU cache of any size, thereby providing estimates on the minimum bandwidth requirements at different levels of the memory hierarchy to avoid being bandwidth bound. However, such an analysis only holds for the particular execution order that produced the trace. It cannot estimate potential improvement in data locality through dependence preserving transformations that change the execution schedule of the operations in the computation.

In this work, we develop a novel dynamic analysis approach to characterize the inherent locality properties of a computation and thereby assess the potential for data locality enhancement via dependence preserving transformations.

This work is the fruit of the collaboration 8.1 with OSU and has been published at ACM TACO'14.

6.5. On Using the Roofline Model with Lower Bounds on Data Movement

Participants: Venmugil Elango [OSU], Naser Sedaghati [OSU], Fabrice Rastello, Louis-Noël Pouchet [UCLA], J. Ramanujam [LSU], Radu Teodorescu [OSU], P. Sadayappan [OSU].

The roofline model is a popular approach to "bounds and bottleneck" performance analysis. It focuses on the limits to performance of processors because of limited bandwidth to off-chip memory. It models upper bounds on performance as a function of operational intensity, the ratio of computational operations per byte of data moved from/to memory. While operational intensity can be directly measured for a specific implementation of an algorithm on a particular target platform, it is of interest to obtain broader insights on bottlenecks, where various semantically equivalent implementations of an algorithm are considered, along with analysis for variations in architectural parameters. This is currently very cumbersome and requires performance modeling and analysis of many variants.

In this work, we alleviate this problem by using the roofline model in conjunction with upper bounds on the operational intensity of computations as a function of cache capacity, derived using lower bounds on data movement. This enables bottleneck analysis that holds across all dependence-preserving semantically equivalent implementations of an algorithm. We demonstrate the utility of the approach in in assessing fundamental limits to performance and energy efficiency for several benchmark algorithms across a design space of architectural variations.

This work is the fruit of the collaboration 8.1 with OSU and is to be published at ACM TACO'15.

6.6. On Characterizing the Data Access Complexity of Programs

Participants: Venmugil Elango [OSU], Fabrice Rastello, Louis-Noël Pouchet [UCLA], J. Ramanujam [LSU], P. Sadayappan [OSU].

Technology trends will cause data movement to account for the majority of energy expenditure and execution time on emerging computers. Therefore, computational complexity will no longer be a sufficient metric for comparing algorithms, and a fundamental characterization of data access complexity will be increasingly important. The problem of developing lower bounds for data access complexity has been modeled using the formalism of Hong & Kung's red/blue pebble game for computational directed acyclic graphs (CDAGs). However, previously developed approaches to lower bounds analysis for the red/blue pebble game are very limited in effectiveness when applied to CDAGs of real programs, with computations comprised of multiple sub-computations with differing DAG structure. We address this problem by developing an approach for effectively composing lower bounds based on graph decomposition. We also develop a static analysis algorithm to derive the asymptotic data-access lower bounds of programs, as a function of the problem size and cache size.

This work is the fruit of the collaboration 8.1 with OSU and is to be presented at ACM POPL'15.

6.7. PolyCheck: Dynamic Verification of Iteration Space Transformations on Affine Programs

Participants: Sriram Krishnamoorthy [PNNL], Bao Wenlei [OSU], Louis-Noël Pouchet [UCLA], P. Sadayappan [OSU], Fabrice Rastello.

High-level compiler transformations, especially loop transformations, are widely recognized as critical optimizations to restructure programs to improve data locality and expose parallelism.

Guaranteeing the correctness of program transformations is essential, and to date three main approaches have been developed: proof of equivalence of affine programs, matching the execution traces of programs, and checking bit-by-bit equivalence of the outputs of the programs. Each technique suffers from limitations in either the kind of transformations supported, space complexity, or the sensitivity to the testing dataset. In this paper, we take a novel approach addressing all three limitations to provide an automatic bug checker to verify any iteration reordering transformations on affine programs, including non-affine transformations, with space consumption proportional to the original program data, and robust to arbitrary datasets of a given size. We achieve this by exploiting the structure of affine program control- and data-flow to generate at compile-time a lightweight checker code to be executed within the transformed program. Experimental results assess the correctness and effectiveness of our method, and its increased coverage over previous approaches.

This work is the result of the collaboration 8.1 with OSU.

6.8. On Using Lower Bounds for Discrimination of Utility/Futility of Loop Fusion

Participants: Samyam Rajbhandari [OSU], Martin Konk [OSU], P. Sadayappan [OSU], Robert J. Harrison [Stonybrook], Fabrice Rastello.

Fusion is an important loop transformation for data locality enhancement. However, it is very challenging to determine which of a set of possible fusion choices is best. In this paper, we pursue a novel approach to addressing this problem. Instead of the conventional approach of explicitly modeling different possible fused loop configurations and modeling the expected performance with each, we instead use lower bounds modeling to characterize conditions where fusion might have utility and where it will be futile because the maximal possible improvement from fusion is much lower than the minimal data movement overheads for each of the unfused components. We successfully demonstrate the use of such a methodology with two practically important codes from the quantum chemistry domain, i) with the affine 4-index transform code, and ii) unstructured tree operations with the MADNESS framework.

This work is the result of the collaboration 8.1 with OSU.

6.9. A Tiling Perspective for Register Optimization

Participants: Duco Van Amstel, Lukasz Domagala, P. Sadayappan [OSU], Fabrice Rastello.

Register allocation is a much studied problem. A particularly important context for optimizing register allocation is within loops, since a significant fraction of the execution time of programs is often inside loop code. A variety of algorithms have been proposed in the past for register allocation, but the complexity of the problem has resulted in a decoupling of several important aspects, including loop unrolling, register promotion, and instruction reordering.

In this work, we develop an approach to register allocation and promotion in a unified optimization framework that simultaneously considers the impact of loop unrolling and instruction scheduling. This is done via a novel instruction tiling approach where instructions within a loop are represented along one dimension and innermost loop iterations along the other dimension. By exploiting the regularity along the loop dimension, and imposing essential dependence based constraints on intra-tile execution order, the problem of optimizing register pressure is cast in a constraint programming formalism. Experimental results are provided from thousands of innermost loops extracted from the SPEC benchmarks, demonstrating improvements over the current state-of-the-art.

This work is the fruit of both the collaboration 8.1 with OSU and with Kalray 7.1 7.2 .

6.10. Hybrid Pointer Disambiguation

Participants: Fernando Pereira, Alexandros Labrineas, Péricles Alves, Fabian Gruber, Fabrice Rastello.

In order to provide effective optimizations, compilers must deal with memory dependences. However, the state-of-the-art heuristics available in the literature to track memory dependencies are inherently imprecise and computationally expensive. Consequently, the most advanced code transformations that compilers have today are ineffective when applied on real-world programs. The goal of this paper is to solve this conundrum - a goal that we accomplish through the hybrid disambiguation of pointers. We provide a static analysis that generates dynamic tests to determine when two memory locations can overlap. We then produce two versions of a loop: one that is aliasing-free - hence, easy to optimize - and another that is not. Our checks lets us safely branch to the optimizable region. We have applied these ideas on Polly-LLVM, a loop optimizer built on top of the LLVM compilation infrastructure. Our experiments indicate that our method is precise, effective and useful: we can disambiguate the vast majority of checks in benchmarks that go from the loop intensive Polybench suite to the more general SPEC CPU 2006 benchmark collection. The result of this precision is code quality: the binaries that we generate are 9.5% faster than those that Polly-LLVM produces without our optimization. Given the current technology to statically solve alias analysis, we believe that our ideas are a necessary step to make modern compiler optimizations useful in practice.

This work is the fruit of the collaboration with UFMG 8.1 and Kalray 7.1 7.2 .

6.11. Parameterized Construction of Program Representations for Sparse Dataflow Analyses

Participants: André Tavares [ENS Lyon], Benoit Boissinot [ENS Lyon], Fernando Pereira, Fabrice Rastello.

Data-flow analyses usually associate information with control flow regions. Informally, if these regions are too small, like a point between two consecutive statements, we call the analysis dense. On the other hand, if these regions include many such points, then we call it sparse. This paper presents a systematic method to build program representations that support sparse analyses. To pave the way to this framework we clarify the bibliography about well-known intermediate program representations. We show that our approach, up to parameter choice, subsumes many of these representations, such as the SSA, SSI and e-SSA forms. In particular, our algorithms are faster, simpler and more frugal than the previous techniques used to construct SSI - Static Single Information - form programs. We produce intermediate representations isomorphic to Choi *et al.*'s Sparse Evaluation Graphs (SEG) for the family of data-flow problems that can be partitioned per variables. However, contrary to SEGs, we can handle - sparsely - problems that are not in this family. We have tested our ideas in the LLVM compiler, comparing different program representations in terms of size and construction time.

This work is the fruit of the collaboration with UFMG 8.1 and has been presented at Springer CC'14.

6.12. Time-critical Computing on a Single-chip Massively Parallel Processor

Participants: Benoit Dupont-de-Dinechin [Kalray], Duco Van Amstel, Marc Poulhiès [Kalray], Guillaume Lager [Kalray].

In this work we demonstrate the capabilities of the MPPA(TM)-256 chip in the field of time-critical computations. This manycore chip features amongst others a Network-on-Chip (NoC) linking the seperate computational clusters each disposing of its own local memory and processing elements (PEs). The PEs architectural features induce a locally deterministic behaviour and the memory access arbitration that is used allows for a Worst-Case Execution Time (WCET) that is achieved for the combination of all local worst-cases. As such, in order to achieve a WCET analysis for a full MPPA(TM)-256 chip, we provide a Worst-Case Traversal Time (WCTT) analysis for the NoC to link the WCETs provided by each computational cluster. This part of the work is based on the (sigma, rho) model used for general network flow analysis and Quality-of-Service (QoS) parametrization.

This work has been presented at DATE'14.

6.13. Guaranteed Services of the NoC of a Manycore Processor

Participants: Benoit Dupont-de-Dinechin [Kalray], Yves Durand [CEA], Duco Van Amstel [Kalray], Alexandre Ghiti [Kalray].

In the case of the MPPA(TM)-256 chip the study of the integrated Network-on-Chip (NoC) is a fundamental subject for anyone using this architecture for time-critical purposes or real-time use-cases that need guarantees on the Worst-Case Traversal Time (WCTT) of the NoC. Previous work has already shown that the MPPA(TM)-256 NoC can be modelled using the (sigma, rho)-model. In the current work we will elaborate on this point by providing an indepth analysis of the NoC as well as the method to guarantee Quality-of-Service properties.

This work has been presented at the International Workshop on Network on Chip Architectures 2014.

<p style="text-align:center;">**PRIVATICS Project-Team**</p>

5. New Results

5.1. Highlights of the Year

Vincent Roca was awarded the 3rd Applied Research price of the Fédération des Industries Electriques, Electroniques et Communications (FIEEC), for his transfer activities to the Expway French SME, Lyon, October 8th, 2014.

The team got two major contributions:

- *A Case Study: Privacy Preserving Release of Spatio-temporal Density in Paris* was published by Gergely Acs and Castelluccia at KDD 2014.
- *Censorship in the Wild: Analyzing Internet Filtering in Syria* was published by Chaabane Abdelberi, Mathieu Cunche,and Mohamed Ali Kaafar at IMC 2014.

5.2. Filtering and blocking the Internet

Participants: Mohamed Ali Kaafar, Abdelberi Chaabane, Mathieu Cunche, Cédric Lauradoux, Amrit Kumar.

- **Censorship**

  Based on 600GB leaked logs from appliances used to filter Internet traffic in Syria, we performed an analysis of the Syrian censorship apparatus. This study have been published in ACM Internet Measurement Conference [7].

  We found that the Internet traffic in Syria was filtered in several ways using IP addresses, domain names and keywords. Content sharing, instant messaging and proxy technologies were heavily censored. Some social media such as badoo.com were fully censored, but others such as Facebook are only censored for specific political and religious pages. We also found evidences of successful usage of censorship-circumvention techniques such as Tor and VPN. We also found that P2P file-sharing and Google cache were used to escape censorship blockage.

  While our work might help organizations on both sides of the censorship line, we believe the presented results can help understand the underlying technologies, policies and can inform the design of tools designed to evade the censorship.

- **Attacking filters** Many major Internet companies use probabilistic techniques to filter the users requests or to prevent malicious attacks. In our work [35], [34], we show how they can be polluted/saturated using pre-image attacks and how it increases the false-positive probability. Then, we show how to forge false-positives to mount attacks. In the adversarial settings, we have the liberty to assume that the inputs to the filter are non uniformly distributed. This observation leads to our second contribution: we compute the worst case false-positive probability and obtain new equations for Bloom filter parameters. To support our contributions, we provide four attacks on software applications based on Bloom filter: Bloom-enabled SCRAPY web spider, BITLYDABLOOMS spam filter, SQUID web cache and GOOGLE Safe Browsing. Our attacks retain some form of DoS. They are all based on the forgery of Uniform Resource Locators (URLs) matching certain pre-image or second pre-image property. The impact of our attack ranges from denial-of-service to massively distributed denial-of-service with reflection.

5.3. Selling Off Privacy at Auction

Participants: Claude Castelluccia, Lukasz Olejnik, Minh-Dung Tran.

The first one is a privacy analysis of Real-Time Bidding (RTB) and Cookie Matching (CM). RTB is a technology that allows ad buyers (advertisers) and ad sellers (publishers) to buy and sell ad spaces at real-time auctions through ad exchanges. In RTB, when user visits a publisher page, the ad impression (i.e. one ad display in an ad space) and the user information are immediately broadcast by the ad exchange to a number of bidders (i.e. advertisers or their representatives) for them to bid for the chance to serve ads to this user. CM protocol allows the ad exchange and the bidder to synchronize their cookies of the same user, thus facilitating their exchange of user data.

In [13], we characterize and quantify the potential user web history leakage from ad exchanges to bidders in RTB as a result of exchanging user data. We also discuss and quantify the extent to which companies can potentially collude to increase their tracked user profiles using CM. In addition, we leverage a design characteristic of RTB to observe the winning price of each RTB auction. By analyzing these prices, we show how advertisers evaluate the value of user privacy. This work (titled Selling Off Privacy at Auction) will be presented in NDSS 2014, San Diego, USA in February, 2014.

5.4. Data anonymization

Participants: Claude Castelluccia, Gergely Acs.

With billions of handsets in use worldwide, the quantity of mobility data is gigantic. When aggregated they can help understand complex processes, such as the spread viruses, and built better transportation systems, prevent traffic congestion. While the benefits provided by these datasets are indisputable, they unfortunately pose a considerable threat to location privacy. At KDD 2014 [9], we present a new anonymization scheme to release the spatio-temporal density of Paris, in France, i.e., the number of individuals in 989 different areas of the city released every hour over a whole week. The density is computed from a call-data-record (CDR) dataset, provided by the French Telecom operator Orange, containing the CDR of roughly 2 million users over one week. Our scheme is differential private, and hence, provides provable privacy guarantee to each individual in the dataset. Our main goal with this case study is to show that, even with large dimensional sensitive data, differential privacy can provide practical utility with meaningful privacy guarantee, if the anonymization scheme is carefully designed. This work is part of the national project XData (http://xdata.fr) that aims at combining large (anonymized) datasets provided by different service providers (telecom, electricity, water management, postal service, etc.).

5.5. Wi-Fi and privacy

Participants: Jagdish Achara, Mathieu Cunche, Vincent Roca.

In Android, installing an application implies accepting the permissions it requests, and these permissions are then enforced at runtime. In our WISEC 2014 paper [29], we focus on the privacy implications of the `ACCESS_WIFI_STATE` permission. For this purpose, we analyzed permissions of the 2700 most popular applications on Google Play and found that the `ACCESS_WIFI_STATE` permission is used by 41% of them. We then performed a static analysis of 998 applications requesting this permission and based on the results, chose 88 applications for dynamic analysis. Our analyses reveal that this permission is already used by some companies to collect user Personally Identifiable Information (PII). We also conducted an online survey to study users' perception of the privacy risks associated with this permission. This survey shows that users largely underestimate the privacy implications of this permission. As this permission is very common, most users are therefore potentially at risk.

5.6. Sensor security and privacy

Participant: Marine Minier.

Wireless sensor networks (WSNs) are composed of a large number of low-cost, low-power, and multi-functional sensor nodes that communicate at short distance through wireless links. They are usually deployed in an open and uncontrolled environment where attackers may be present. Due to the use of low-cost materials, hardware components are not tamper-resistant and an adversary could access to a sensor's internal state. With Ochirkhand Erdene-Ochir and Pierre Brunisholz, we continue to work on the notion of resiliency in WSNs [17], [31].

5.7. Buidling blocks

Participant: Marine Minier.

In the context of the BLOC project funded by the ANR, we continue to work on Extended Generalized Feistel Network and on new ligthweight block cipher design (see [30]). We hope to obtain results in this area at the beginning of 2015. With Christine Solnon and Julia Reboul, we work on the formalism of related-key and chosen-key attacks against symmetric key primitives using constraint programming (CP). This preliminary work was presented at the CP 2014 workshop ModRef 2014 in [42].

5.8. Formal and legal issues of privacy

Participants: Thibaud Antignac, Denis Butin, Daniel Le Métayer.

- **Privacy Architectures: Reasoning About Data Minimization and Integrity** Privacy by design will become a legal obligation in the European Community if the Data Protection Regulation eventually gets adopted. However, taking into account privacy requirements in the design of a system is a challenging task. We present an approach based on the specification of privacy architectures at FM 2014 [12] and focus on a key aspect of privacy, data minimisation, and its tension with integrity requirements. We illustrate our formal framework through a smart metering case study.

- **Log Analysis for Data Protection Accountability**

  Accountability is increasingly recognized as a cornerstone of data protection, notably in European regulation, but the term is frequently used in a vague sense. For accountability to bring tangible benefits, the expected properties of personal data handling logs and the assumptions regarding the logging process must be defined with accuracy. At STM 2014 [10], we provide a formal framework for accountability and show the correctness of the log analysis with respect to abstract traces used to specify privacy policies. We also show that compliance with respect to data protection policies can be checked based on logs free of personal data, and describe the integration of our formal framework in a global accountability process.

SPADES Team

6. New Results

6.1. Components and Contracts

Participant: Jean-Bernard Stefani.

6.1.1. Location graph model

The design of configurable systems can be streamlined and made more systematic by adopting a component-based structure, as demonstrated with the Fractal component model [2]. However, the formal foundations for configurable component-based systems, featuring higher-order capabilities where components can be dynamically instantiated and passivated, and non-hierarchical structures where components can be contained in different composites at the same time, are still an open topic. We have developed recently the location graph model [15], where components are understood as graphs of locations hosting higher-order processes, and where component structures can be arbitrary graphs. We have developed a compositional operational semantics for the location graph model, which is parametric with respect to the family of processes. We have shown that the location graph model constitutes a conservative extension of a previous model, called CAB, that captures the key features of the BIP component model [5]. We have further worked on the behavioral theory of the location graph model, characterizing contextual equivalence in the model by means of a higher-order bisimularity relation, and begun the study of the encoding of different models, including the Synchronized Hyperedge Replacement model [45].

6.2. Real-Time multicore programming

Participants: Vagelis Bebelis, Adnan Bouakaz, Pascal Fradet, Alain Girault, Gregor Goessler, Jean-Bernard Stefani, Sophie Quinton, Partha Roop, Eugene Yip.

6.2.1. Analysis and scheduling of parametric dataflow models

Recent data-flow programming environments support applications whose behavior is characterized by dynamic variations in resource requirements. The high expressive power of the underlying models (*e.g.*, Kahn Process Networks or the CAL actor language) makes it challenging to ensure predictable behavior. In particular, checking *liveness* (*i.e.*, no part of the system will deadlock) and *boundedness* (*i.e.*, the system can be executed in finite memory) is known to be hard or even undecidable for such models. This situation is troublesome for the design of high-quality embedded systems.

Recently, we have introduced the *schedulable parametric data-flow* (SPDF) MoC for dynamic streaming applications [47]. SPDF extends the standard dataflow model by allowing rates to be parametric. Last year, we have proposed the *Boolean Parametric Data Flow* (BPDF) MoC which combines integer parameters (to express dynamic rates) and boolean parameters (to express the activation and deactivation of communication channels). High dynamicity is provided by integer parameters which can change at each basic iteration and boolean parameters which can change even within the iteration. We have presented static analyses which ensure the liveness and the boundedness of BDPF graphs.

Recently, we have proposed a generic and flexible framework to generate parallel schedules for BPDF applications [16]. The parametric dataflow graph is associated with user-defined specific constraints aimed at minimizing, timing, buffer sizes, power consumption, or other criteria. The scheduling algorithm executes with minimal overhead and can be adapted to different scheduling policies just by changing some constraints. The safety of both the dataflow graph and constraints can be checked statically and all schedules are guaranteed to be bounded and deadlock free. Our case studies are video decoders for high definition video streaming such as VC-1. One of the target architectures is the STHORM many-core platform designed by STMicroelectronics.

This research is the central topic of Vagelis Bebelis' PhD thesis. It is conducted in collaboration with STMicroelectronics.

6.2.2. *Typical Worst-Case Analysis of real-time systems*

Weakly hard time constraints have been proposed for applications where occasional deadline misses are permitted. We have recently developed Typical Worst Case Analysis (TWCA) to exploit similar constraints and bound response times of systems with sporadic overload. This year, we have applied this approach to a real-life automotive network [14]. Additionally, we have extended the approach for static priority preemptive (SPP) and static priority non-preemptive (SPNP) scheduling to determine the maximum number of deadline misses of a given task [21]. The approach is based on an optimization problem which trades off higher priority interference versus miss count. We formally derived a lattice structure for these combinations that lays the ground for an integer linear programming (ILP) formulation. The ILP solution was evaluated and provided far better results than previous TWCA.

In parallel, we have contributed to a systematic co-engineering approach that integrates TWCA into functional analysis [19]. We combine physical, control and timing models by representing them as a network of hybrid automata. Closed-loop properties can then be verified on this hybrid automata network by using standard model checkers for hybrid systems. The use of the Logical Execution Time (LET) semantics where data is written back deterministically at the typical worst-case response time (rather than the usual worst-case bound) is a new and particularly powerful approach for addressing the computational complexity of the model checking problem.

6.2.3. *Time predictable programming*

In the context of the RIPPES associated team with UC Berkeley and U Auckland, we have finalized ongoing work on our synchronous programming language for time predictability PRET-C [10]. PRET-C extends C with synchronous constructs inspired by ESTEREL, to allow an easy programming of concurrent reactive programs. These constructs allow the programmer to express concurrency, interaction with the environment, looping, and a synchronization barrier (like the `pause` statement in ESTEREL). PRET-C's semantics is deterministic, and it can be efficiently compiled towards sequential code, either executed on a dedicated processor for the best predictability of the program's Worst-Case Reaction Time (WCRT), or executed on a generic processor.

We have also continued our work on FOREC, a time predictable synchronous programming language for multi-core chips. Like PRET-C, it extends C with a small set of ESTEREL-like synchronous primitives. FOREC threads communicate with each other via shared variables, the values of which are combined at the end of each tick to maintain deterministic execution. FOREC is compiled into threads that are then statically scheduled for a target multi-core chip. This is the main difference with PRET-C. We have finalized the semantics of FOREC, which led us to propose several ways to combine shared variables at the tick boundaries, such that the semantics remains deterministic. This part was inspired by the so-called concurrent revisions [38].

Finally, with colleagues from the former ARTISTDESIGN European Network of Excellence, we have also participated in a survey on predictable embedded systems [11].

6.2.4. *Tradeoff exploration between energy consumption and execution time*

We have continued our work on multi-criteria scheduling, in the particular context of dynamic applications that are launched and terminated on an embedded multi-core chip, under execution time and energy consumption constraints. We have proposed a two layer adaptive scheduling method. In the first layer, each application (represented as a DAG of tasks) is scheduled statically on sets of cores: 2 cores, 3 cores, 4 cores, and so on. For each size of these sets (2, 3, 4, ...), there may be only one topology or several topologies. For instance, for 2 or 3 cores there is only one topology (a "line"), while for 4 cores there are three distinct topologies ("line", "square", and "T shape"). Moreover, for each topology, we generate statically several schedules, each one subject to a different total energy consumption constraint, and consequently with a different Worst-Case Reaction Time (WCRT). Coping with the energy consumption constraints is achieved thanks to Dynamic Frequency and Voltage Scaling (DVFS). In the second layer, we use these pre-generated static schedules to reconfigure dynamically the applications running on the multi-core each time a new application is launched or

an existing one is stopped. The goal of the second layer is to perform a global optimization of the configuration, such that each running application meets a pre-defined quality-of-service constraint (translated into an upper bound on its WCRT) and such that the total energy consumption is minimized. For this, we (1) allocate a sufficient number of cores to each active application, (2) allocate the unassigned cores to the applications yielding the largest gain in energy, and (3) choose for each application the best topology for its subset of cores (i.e., better than the by default "line" topology).

This is a joint work with Ismail Assayad (U. Casablanca, Morocco) who visits the team regularly.

6.3. Language Based Fault-Tolerance

Participants: Dmitry Burlyaev, Pascal Fradet, Alain Girault, Yoann Geoffroy, Gregor Goessler, Jean-Bernard Stefani.

6.3.1. *Automatic transformations for fault tolerant circuits*

In the past years, we have studied the implementation of specific fault tolerance techniques in real-time embedded systems using program transformation [1]. We are now investigating the use of automatic transformations to ensure fault-tolerance properties in digital circuits. To this aim, we consider program transformations for hardware description languages (HDL). We consider both single-event upsets (SEU) and single-event transients (SET) and fault models of the form "at most 1 SEU or SET within n clock signals".

We have expressed several variants of triple modular redundancy (TMR) as program transformations. We have proposed a verification-based approach to minimize the number of voters in TMR [17]. Our technique guarantees that the resulting circuit *(i)* is fault tolerant to the soft-errors defined by the fault model and *(ii)* is functionally equivalent to the initial one. Our approach operates at the logic level and takes into account the input and output interface specifications of the circuit. Its implementation makes use of graph traversal algorithms, fixed-point iterations, and BDDs. Experimental results on the ITC'99 benchmark suite indicate that our method significantly decreases the number of inserted voters, which entails a hardware reduction of up to 55% and a clock frequency increase of up to 35% compared to full TMR. We address scalability issues arising from formal verification with approximations and assess their efficiency and precision.

We have proposed novel fault-tolerance transformations based on time-redundancy. In particular, we have presented a transformation using double-time redundancy (DTR) coupled with micro-checkpointing, rollback and a speedup mode [18]. The approach is capable to mask any SET every 10 cycles and keeps the same input/output behavior regardless error occurrences. Experimental results on the ITC'99 benchmark suite indicate that the hardware overhead is 2.7 to 6.1 times smaller than full TMR with double loss in throughput. It is an interesting alternative to TMR for logic intensive designs.

We have also designed a transformation that allows the circuit to change its level of time-redundancy. This feature permits to dynamically and temporarily give up (resp. increase) fault-tolerance and speed up (resp. slow down) the circuit. The motivations for such changes can be based on the observed change in radiation environment or the processing of (non)critical data. These different time redundancy transformations have been patented [23]

We have started the formal certification of such transformations using the Coq proof assistant [40]. The transformations are described on a simple gate-level hardware description language inspired from μFP [68]. The fault-model is described in the operational semantics of the language. The main theorem states that, for any circuit, for any input stream and for any SET allowed by the fault-model, its transformed version produces a correct output. A TMR and triple time redundancy transformations have already been proved correct. The proof of the DTR transformation is in progress.

6.3.2. *Concurrent flexible reversibility*

In the recent years, we have been investigating reversible concurrent computation, and investigated various reversible concurrent programming models, with the hope that reversibility can shed some light on the common semantic features underlying various forms of fault recovery techniques (including, exceptions, transactions, and checkpoint/rollback schemes).

We have revisited our encoding of our reversible higher-order π-calculus in (a variant of) the higher-order π-calculus, in order to obtain a much tighter result than our original encoding. In essence, we now have a form of strong bisimilarity (modulo administrative reductions) between a reversible higher-order π-calculus process and its translation in higher-order π. We have also studied the relation between the causality information used in our reversible higher-order π and a causal higher-order π-calculus, inspired by the causal π-calculus [35]. This work has been submitted for publication [24]. This work was done in collaboration with Inria teams FOCUS in Bologna, as part of the ANR REVER project.

6.3.3. Blaming in component-based systems

The failure of one component may entail a cascade of failures in other components; several components may also fail independently. In such cases, elucidating the exact scenario that led to the failure is a complex and tedious task that requires significant expertise.

The notion of causality *(did an event e cause an event e'?)* has been studied in many disciplines, including philosophy, logic, statistics, and law. The definitions of causality studied in these disciplines usually amount to variants of the counterfactual test "e is a cause of e' if both e and e' have occurred, and in a world that is as close as possible to the actual world but where e does not occur, e' does not occur either". Surprisingly, the study of logical causality has so far received little attention in computer science, with the notable exception of [51] and its instantiations. However, this approach relies on a causal model that may not be known, for instance in presence of black-box components. For such systems, we have been developing a framework for blaming that helps us establish the causal relationship between component failures and system failures, given an observed system execution trace. The analysis is based on a formalization of counterfactual reasoning. We have shown in [12] how our approach can be used for log analysis to help establishing liability in the context of legal contracts.

We have proposed in [6] an approach for blaming in component-based real-time systems whose component specifications are given as timed automata. The analysis is based on a single execution trace violating a safety property P. We have formalized blaming using counterfactual reasoning to distinguish component failures that actually contributed to the outcome from failures that had no impact on the violation of P. We have shown how to effectively implement blaming by reducing it to a model-checking problem for timed automata. The approach has been implemented in LoCA (Section 5.1.1). We have further demonstrated the feasibility of our approach on the model of a dual-chamber implantable pacemaker.

6.3.4. Synthesis and implementation of fault-tolerant embedded systems

We have integrated a complete workflow to synthesize and implement correct-by-construction fault tolerant distributed embedded systems consisting of real-time periodic tasks. Correct-by-construction is provided by the use of discrete controller synthesis [63] (DCS), a formal method thanks to which we are able to guarantee that the synthe-sized controlled system satisfies the functionality of its tasks even in the presence of processor failures. For this step, our workflow uses the Heptagon domain specific language [43] and the Sigali DCS tool [59]. The correct implementation of the resulting distributed system is a challenge, all the more since the controller itself must be tolerant to the processor failures. We achieve this step thanks to the libDGALS real-time library [22] (1) to generate the glue code that will migrate the tasks upon processor failures, maintaining their internal state through migration, and (2) to make the synthesized controller itself fault-tolerant. We have demonstrated the feasibility of our work-flow on a multi-tasks multi-processor fault-tolerant distributed system.

<p style="text-align:center">**BIPOP Project-Team**</p>

6. New Results

6.1. Highlights of the Year

- Bernard Brogliato: keynote speaker at ICDVC-2014, 4th International Conference on Dynamics, Vibration and Control, August 23-25, 2014 in Shanghai, China. http://www.icdvc2014.org/

6.2. Multiple impacts modelling

Participant: Bernard Brogliato.

The work consists of studying two systems: the rocking block and tapered chains of balls, using the Darboux-Keller model of multiple impacts previously developed. The objectives are threefold: 1) show that the model predicts well the motion by careful comparisons with experimental data found in the literature, 2) study the system's dynamics and extract critical kinetic angles that allow the engineer to predict the system's gross motion, 3) develop numerical code inside the SICONOS platform that incorporates the model of multiple impact. The influence of the kinetic angles in the rocking block motion with friction is analysed as well, numerically. Extensive experimental works have been conducted by our colleague C. Liu at PKU on a disc-ball system. Results are in [32] [67], and in the monograph [16]. Multiple impacts have also been tackled through generalized kinematic models using the kinetic metric [20].

6.3. The contact complementarity problem

Participants: Bernard Brogliato, Florence Bertails-Descoubes, Alejandro Blumentals.

The contact linear complementarity problem is an set of equalities and complementarity conditions whose unknowns are the acceleration and the contact forces. It has been studied in a frictionless context with possibly singular mass matrix and redundant constraints in [21], using results on well-posedness of variational inequalities obtained earlier by the authors. This is also the topic of the first part of the Ph.D. thesis of Alejandro Blumentals where the frictional case is treated as a perturbation of the frictionless case. The contact LCP is directly related to the so-called Painlevé's paradox of contact mechanics. In collaboration with C. Liu (Beijing university PKU) some results habe been obtained from the analysis of a compliant model in the limit. It shows on the classical sliding rod system that the inconsistent mode yield to instantaneous transition to a sticking mode. This is quite coherent with previous results obtained by Le xuan Anh in 1991 on the Painlevé-Klein system (bilateral constraints with Coulomb friction). The results will appear in Multibody System Dynamics in 2015.

6.4. Discrete-time sliding mode control

Participants: Vincent Acary, Bernard Brogliato, Olivier Huber.

This topic concerns the study of time-discretized sliding-mode controllers. Inspired by the discretization of nonsmooth mechanical systems, we propose implicit discretizations of discontinuous, set-valued controllers. This is shown to result in preservation of essential properties like simplicity of the parameters tuning, suppression of numerical chattering, reachability of the sliding surface after a finite number of steps, and disturbance attenuation by a factor h or h^2 [41], [42], [43], [45], [61]. This work is part of the ANR project CHASLIM. Within the framework of CHASLIM we have performed many experimental validations on the electropneumatic setup of IRCCyN (Nantes), which nicely confirm our theoretical and numerical predictions: the implicit implementation of sliding mode control, drastically improves the input and output chattering behaviours. In particular the high frequency bang-bang controllers which are observed with explicit discretizations, are completely suppressed. The implicit discretization has been applied to the classical equivament-based-control SMC, and also to the twisting sliding-mode controller [43].

6.5. Lur'e set-valued dynamical systems

Participants: Bernard Brogliato, Christophe Prieur.

Lur'e systems are quite popular in Automatic Control since the fifties. Set-valued Lur'e systems possess a static feedback nonlinearity that is a multivalued function. This study consists in the mathematical analysis (existence and uniqueness of solutions) and the stability analysis (Lyapunov stability, invariance principle) of classes of set-valued Lur'e systems, with applications in complementarity dynamical systems, relay systems, mechanical systems with dry friction, electrical circuits, etc. Our works in this field started in [62]. The results in [64] extend those in [63] with an accurate characterization of the maximal monotonicity of the central operator of these systems, which consists of a projection-like operator. Concrete and verifiable criteria are provided for the above classes (complementarity, relay systems). Results on state observers and output feedback control for classes of Lur'e systems (namely: Moreau's sweeping process of first and second order, and with prox-regular sets) are proposed in [29], [44], [34]. Therein the convexity is replaced by the far more general notion of prox-regularity, which destroys the monotonicity. The input to state stability of measure driven differential equations has been tackled in [22], where some results from [29] are adapted.

6.6. Simulation and stability of piecewise linear gene networks

Participants: Vincent Acary, Arnaud Tonnelier, Bernard Brogliato.

This work has been done in collaboration with the IBIS project team, it is reported in [19]. Gene regulatory networks control the response of living cells to changes in their environment. A class of piecewise-linear (PWL) models, which capture the switch-like interactions between genes by means of step functions, has been found useful for describing the dynamics of gene regulatory networks. The step functions lead to discontinuities in the right-hand side of the differential equations. This has motivated extensions of the PWL models based on differential inclusions and Filippov solutions, whose analysis requires sophisticated numerical tools. We present a method for the numerical analysis of one proposed extension, called Aizerman-Pyatnitskii (AP)-extension, by reformulating the PWL models as a mixed complementarity system (MCS). This allows the application of powerful methods developed for this class of nonsmooth dynamical systems, in particular those implemented in the Siconos platform. We also show that under a set of reasonable biological assumptions, putting constraints on the right-hand side of the PWL models, AP-extensions and classical Filippov (F)-extensions are equivalent. This means that the proposed numerical method is valid for a range of different solution concepts. We illustrate the practical interest of our approach through the numerical analysis of three well-known networks developed in the field of synthetic biology.

In addition, we have investigated oscillatory regimes in repressilator-type models with piecewise linear dynamics [30]. We derived exact analytical conditions for oscillations and showed that the relative location between the dissociation constants of the Hill functions and the ratio of kinetic parameters determines the possibility of oscillatory activities. We also computed analytically the probability of oscillations. Results suggest that a switch-like coupling behaviour, a time-scale separation and a repressilator-type architecture with an even number of elements facilitate the emergence of sustained oscillations in biological systems.

6.7. Numerical analysis and simulation of mechanical systems with constraints

6.7.1. *Event-capturing schemes for nonsmooth mechanical systems*
Participant: Vincent Acary.

To perform the numerical time integration of nonsmooth mechanical systems, the family of event-capturing time–stepping schemes are the most robust and efficient tools. Nevertheless, they suffer from several draw-backs : a) a low–order accuracy (at best at order one), b) a drift phenomena when the unilateral constraints are treated at the velocity level and c) a poor "energetic" behavior in terms of stabilizing the high–frequency dynamics. We proposed self-adapting schemes by applying time–discontinuous Galerkin methods to the measure differential equation in [28]. In order to satisfy in discrete time, the impact law and the constraints at the position and the velocity level, an adaptation of the well-known Gear–Gupta–Leimkuhler approach has been

developed. In [58], the approach is algorithmically specified, improved and applied to nonlinear multi-contact examples with friction. It is compared to other numerical schemes and it is shown that the newly proposed integration scheme yields a unified behavior for the description of contact mechanical problems. Especially, we provide time-integration of the nonimpulsive dynamics with semi–explicit Runge–Kutta method previously developed for differential algebraic equations.

6.7.2. *Numerical time–integration methods for event-detecting schemes.*

Participants: Vincent Acary, Bernard Brogliato, Mounia Haddouni.

The CIFRE thesis of M. Haddouni concerns the numerical simulation of mechanical systems subject to holonomic bilateral constraints, unilateral constraints and impacts. This work is performed in collaboration with ANSYS and the main goal is to improve the numerical time–integration in the framework of event-detecting schemes. Between nonsmooth events, time integration amounts to numerically solving a differential algebraic equations (DAE) of index 3. We have compared dedicated solvers (Explicit RK schemes, Half-explicit schemes, generalizes α-schemes) that solve reduced index formulations of these systems. Since the drift of the constraints is crucial for the robustness of the simulation through the evaluation of the index sets of active contacts, we have proposed some recommendations on the use of the solvers of dedicated to index-2 DAE. A manuscript has been submitted to Multibody System Dynamics.

6.7.3. *Multibody systems with clearances (dynamic backlash)*

Participants: Vincent Acary, Bernard Brogliato, Narendra Akadkhar.

The PhD thesis of N. Akadkhar under contract with Schneider Electric concerns the numerical simulation of mechanical systems with unilateral constraints and friction, where the presence of clearances in imperfect joints plays a crucial role. A first work deals with four-bar planar mechanisms with clearances at the joints, which induce unilateral constraints and impacts, rendering the dynamics nonsmooth. The objective is to determine sets of parameters (clearance value, restitution coefficients, friction coefficients) such that the system's trajectories stay in a neighborhood of the ideal mechanism (*i.e.* without clearance) trajectories. The analysis is based on numerical simulations obtained with the projected Moreau-Jean time-stepping scheme. These results have been reported in [37]. It is planned to extend these simulations to frictional cases and to mechanisms of circuit breakers.

6.8. Inverse modeling with contact and friction

6.8.1. *Inverse statics of plates and shells*

Participants: Florence Bertails-Descoubes, Romain Casati, Gilles Daviet.

We have started to investigate the problem of interpreting an arbitrary 3D mesh as an equilibrium configuration of an elastic plate/shell, subject to gravity and frictional contact forces. We have first considered a simple nodal shell model accounting for stretch, shear and bending. For such a model, inverse statics formulates as an ill-posed minimization problem with a nonlinear objective and nonsmooth constraints. Our objective is to examine this problem in the case where the rest pose of the system is left as unknown, while material parameters (mass, stiffness) are assumed to be known (inverse design problem). In some specific cases (cloth modeling), we use a priori information such as locally low Gaussian curvature so as to help the retrieval of most natural solutions. We plan to submit our results to Siggraph 2015. Targeted applications include virtual garment modeling and parameter retrieval from 3D image-based capture.

6.9. Modeling of fibrous medium

6.9.1. *Continuous modeling of fiber assemblies*

Participants: Florence Bertails-Descoubes, Gilles Daviet.

Following the exploratory project funded by Persyval (2013-2014), we have started to model an assembly of long elastic fibers (such as hair) using a continuous approach (continuum mechanics equations coupled with a nonsmooth stress-strain law). Interactions between air and fibers can then be naturally accounted for, increasing the realism of some macroscopic features compared to our previous discrete elements model. This is still work in progress and we will make some of our results publicly available in 2015.

6.10. Threshold in spiking neural models

Participant: Arnaud Tonnelier.

We studied the threshold for spike initiation in two-dimensionnal spiking neural models. A threshold criterion that depends on both the membrane voltage and the recovery variable is proposed. This approach provides a simple and unified framework that accounts for numerous voltage threshold properties. Implications for neural modeling are also discussed [31].

6.11. Nonsmooth modes in chains of impact oscillators

Participants: Vincent Acary, Guillaume James, Franck Pérignon.

Chains of impact oscillators arise for example as finite-element models of thin oscillating mechanical structures (a string under tension or a clamped beam) contacting rigid obstacles. Nonlinear periodic waves are observed in experiments on such systems, but relatively little is known from a theoretical point of view on their existence and stability. In 2008, Gendelman and Manevitch have analyzed the existence and stability of nonlinear localized modes (breathers) for discrete linear chains with a single node undergoing rigid impacts. In this work, we introduce a numerical method allowing to compute branches of time-periodic solutions when an arbitrary number of nodes undergo rigid impacts without energy dissipation. For this purpose, we reformulate the search of periodic solutions as a boundary value problem incorporating unilateral constraints. We illustrate this numerical approach by computing different families of breathers and nonlinear normal modes. Our method is much more effective than a numerical continuation of periodic solutions based on compliant models, which requires to integrate stiff differential equations and lead to costly numerical continuation. These results have been communicated in two international conferences, ENOC 2014 [35] and 11th World Congress on Computational Mechanics [36].

6.12. Traveling waves in spatially discrete excitable media

Participants: José Eduardo Morales, Arnaud Tonnelier, Guillaume James.

The propagation of traveling waves in excitable media is a widespread phenomenon, with applications ranging from forest fires to electrical signals propagating along nerve fibers. The case of spatially discrete excitable models is notoriously difficult to analyze. In particular, for the discrete FitzHugh-Nagumo reaction-diffusion system, the existence of pulses for a general class of bistable nonlinearities has been proved only recently (Hupkes and Sandstede, 2010). The existence of pulses under more general types of interactions (e.g. elastic instead of diffusive) remains an open question, as well as traveling wave propagation in higher-dimensional systems. These problems will be tackled in the PhD thesis of J.-E. Morales (advisors A. Tonnelier and G. James), which started on November 2013. J.-E. Morales has started to analyze pulse propagation in the excitable Burridge-Knopoff (BK) model, which finds applications in the context of nonlinear friction. This model includes elastic interactions between particles, and an additional difficulty linked with nonsmoothness of the (multivalued) Coulomb friction law. Using an idealized piecewise linear friction force, we have studied the propagation of a pulse wave in the discrete BK model. Using asymptotic methods, we proved the existence of a pulse wave and derived quantitative results for travelling wave properties.

6.13. Nonlinear waves in granular chains

Participants: Guillaume James, Bernard Brogliato.

Granular chains made of aligned beads interacting by contact (e.g. Newton's cradle) are widely studied in the context of impact dynamics and acoustic metamaterials. When a large number of beads are present, their dynamics can be described by infinite-dimensional differential equations, which possess a limited smoothness when unilateral Hertzian contact interactions are considered. In this context, we have developed and analyzed new reduced-order models describing nonlinear wave propagation in such systems. In the work [25] (collaboration with D.Pelinovsky, McMaster Univ.), we analyze small amplitude slowly modulated compression waves in the limit when the exponent of the Hertz force is close to unity. From a multiple scale analysis, we derive a Korteweg-de Vries equation with logarithmic nonlinearity allowing to approximate wave profiles, in particular solitary wave solutions. In the work [50] (collaboration with Y. Starosvetsky, Technion IIT), we prove existence of spatially localized nonlinear modes (breathers) in the DpS equation, an amplitude equation describing small oscillations in Newton's cradle over long time scales. For Hertz force exponents close to unity, we show that breather envelopes are well approximated by a Gaussian solution of the logarithmic nonlinear Schrödinger equation. This result is generalized to traveling localized oscillations (traveling breathers) generated by an impact in Newton's cradle (G. James, article in preparation). The existence of breathers is also analyzed in granular metamaterials consisting of hollow beads with internal masses (G. James) in collaboration with L. Liu, A. Vainchtein (Pittsburgh Univ.) and P. Kevrekidis (UMass Amherst) - article in preparation. In addition the LZB model introduced in [15] has been extensively used to numerically investigate wave phenomena in chains of aligned balls (tapered, monodisperse, anti-tapered, stepped chains). Thorough comparisons with experimental results reported in the Granular Matter literature have been made. The results are reported in the monograph [16].

6.14. Robotics

6.14.1. *Lexicographic Least-Squares solver*

Participants: Pierre-Brice Wieber, Dimitar Dimitrov.

We have been working on Multi-Objective Least-Squares problems with inequality constraints for the last few years, focusing especially on the Lexicographic case. The focus this year has been on nonlinear problems, in collaboration with Adrien Escande from JRL, Tsukuba, Japan. Questions concerning the second-order approximation, using a Gauss-Newton approach or considering more precise second-order information, and questions concerning the globalization scheme, trust-region and/or filter methods have been approached, but results are still preliminary.

6.14.2. *Mobile manipulation by humanoid robots*

Participants: Pierre-Brice Wieber, Dimitar Dimitrov, Alexander Sherikov, Jory Lafaye.

The realization of mobile manipulation by humanoid robots requires the handling of two simultaneous problems: taking care of the dynamic balance of the robot, what is usually done with Model Predictive Control (MPC) schemes, and redundant motion and force control of the whole body of the robot, what is usually done with a Quadratic Program, or a more advanced Lexicographic Least-Squares problem (see above). These two problems are usually solved in sequence: an MPC scheme first computes the necessary motion of the feet and Center of Mass (CoM) of the robot, then motion and force redundancy of the whole body of the robot is resolved. We have observed that this sequence corresponds to a lexicographic order between two objectives, feet and CoM motion first, the rest of the body after, which limits the possibility to tackle scenarios where we would like the motion of the CoM of the robot to be driven by the motion of the rest of the body of the robot, for example to catch an object with the hand. We have proposed therefore to reorganize the order between these different objectives, building on the LexLS solver presented above. The focus this year has been on non-coplanar multi-contact situations.

6.14.3. *Reactive trajectory generation*

Participants: Pierre-Brice Wieber, Dimitar Dimitrov, Saed Al Homsy, Matthieu Guilbert.

The goal of the ongoing collaboration with Adept Technologies is to generate near time optimal trajectories in the presence of moving obstacles in real time. Results are not public yet due to industrial constraints.

6.15. Optimization

6.15.1. Semidefinite programming and combinatorial optimization

Participant: Jérôme Malick.

We have worked with Frederic Roupin (Prof. at Paris XIII) and Nathan Krislock (Assistant Prof. at North Illinois University, USA) on the use of semidefinite programming to solve combinatorial optimization problems to optimality.

We proposed a new family of semidefinite bounds for 0-1 quadratic problems with linear or quadratic constraints [65]. We have embedded the new bounds within branch-and-bound algorithms to solve 2 standard combinatorial optimization problems to optimality.

- *Max-cut.* We developed [26] an improved bounding procedure obtained by reducing two key parameters (the target level of accuracy and the stopping tolerance of the inner Quasi-Newton engine) to zero, and iteratively adding triangle inequality cuts. We also precisely analyzed its theoretical convergence properties. We show that our method outperform the state-of-the-art solver ([66]) on the large test-problems.

- *Heaviest k-subgraph problems*. Adapting the techniques we developed for the max-cut problem, we have proposed in [60] an algorithm able to solve exactly k-cluster instances of size 160. In practice, our method works particularly fine on the most difficult instances (with a large number of vertices, small density and small k).

We have also been working on a generic online semidefinite-based solver for binary quadratic problems using the generality of [65]. Finally, a first web interface for our solvers and our data sets are available online at http://lipn.univ-paris13.fr/BiqCrunch/.

6.15.2. Quadratic stabilization of Benders decomposition

Participants: Jérôme Malick, Sofia Zaourar.

The Benders decomposition, a fundamental method in operation research, is known to have the inherent instability of cutting plane-based methods. The PhD thesis of Sofia Zaourar proposes a algorithmic improvement of the method inspired from the level-bundle methods of nonsmooth optimisation. We illustrate the interest of the stabilization on two classical network problems: network design problems and hub location problems. We also prove that the stabilized Benders method have the same theoretical convergence properties as the usual Benders method. An article about this research was submitted this summer.

<div align="center">

MISTIS Project-Team

</div>

6. New Results

6.1. Highlights of the Year

6.1.1. P-Locus software and Pixyl start-up project

The work on the P-Locus software has been exploited in order to create a start-up in January 2015. The project called Pixyl have been accepted by the GATE1 incubator and has been awarded a BPI emergence prize. It is leaded by Senan Doyle (future CEO). The other co-founders are Michel Dojat (INSERM, GIN), Florence Forbes (Inria, Mistis) and IT-Translation.

6.2. Mixture models

6.2.1. Parameter estimation in the heterogeneity linear mixed model

Participant: Marie-José Martinez.

Joint work with: Emma Holian (National University of Ireland, Galway)

In studies where subjects contribute more than one observation, such as in longitudinal studies, linear mixed models have become one of the most used techniques to take into account the correlation between these observations. By introducing random effects, mixed models allow the within-subject correlation and the variability of the response among the different subjects to be taken into account. However, such models are based on a normality assumption for the random effects and reflect the prior belief of homogeneity among all the subjects. To relax this strong assumption, Verbeke and Lesaffre (1996) proposed the extension of the classical linear mixed model by allowing the random effects to be sampled from a finite mixture of normal distributions with common covariance matrix. This extension naturally arises from the prior belief of the presence of unobserved heterogeneity in the random effects population. The model is therefore called the heterogeneity linear mixed model. Note that this model does not only extend the assumption about the random effects distribution, indeed, each component of the mixture can be considered as a cluster containing a proportion of the total population. Thus, this model is also suitable for classification purposes.

Concerning parameter estimation in the heterogeneity model, the use of the EM-algorithm, which takes into account the incomplete structure of the data, has been considered in the literature. Unfortunately, the M-step in the estimation process is not available in analytic form and a numerical maximisation procedure such as Newton-Raphson is needed. Because deriving such a procedure is a non-trivial task, Komarek et al. (2002) proposed an approximate optimization. But this procedure proved to be very slow and limited to small samples due to requiring manipulation of very large matrices and prohibitive computation.

To overcome this problem, we have proposed in an alternative approach which consists of fitting directly an equivalent mixture of linear mixed models. Contrary to the heterogeneity model, the M-step of the EM-algorithm is tractable analytically in this case. Then, from the obtained parameter estimates, we can easily obtain the parameter estimates in the heterogeneity model.

6.2.2. Taking into account the curse of dimensionality

Participants: Stéphane Girard, Alessandro Chiancone, Seydou-Nourou Sylla.

Joint work with: C. Bouveyron (Univ. Paris 5), M. Fauvel (ENSAT Toulouse) and J. Chanussot (Gipsa-lab and Grenoble-INP)

In the PhD work of Charles Bouveyron (co-advised by Cordelia Schmid from the Inria LEAR team) [67], we propose new Gaussian models of high dimensional data for classification purposes. We assume that the data live in several groups located in subspaces of lower dimensions. Two different strategies arise:

- the introduction in the model of a dimension reduction constraint for each group
- the use of parsimonious models obtained by imposing to different groups to share the same values of some parameters

This modelling yields a new supervised classification method called High Dimensional Discriminant Analysis (HDDA) [4]. Some versions of this method have been tested on the supervised classification of objects in images. This approach has been adapted to the unsupervised classification framework, and the related method is named High Dimensional Data Clustering (HDDC) [3]. Our recent work consists in adding a kernel in the previous methods to deal with nonlinear data classification and heterogeneous data [12]. We also investigate the use of kernels derived from similary measures on binary data. The targeted application is the analysis of verbal autopsy data (PhD thesis of N. Sylla): Indeed, health monitoring and evaluation make more and more use of data on causes of death from verbal autopsies in countries which do not keep records of civil status or with incomplete records. The application of verbal autopsy method allows to discover probable cause of death. Verbal autopsy has become the main source of information on causes of death in these populations.

6.2.3. *Location and scale mixtures of Gaussians with flexible tail behaviour: properties, inference and application to multivariate clustering*

Participant: Florence Forbes.

Joint work with: Darren Wraith from QUT, Brisbane Australia.

Clustering concerns the assignment of each of N, possibly multidimensional, observations $y_1, ..., y_N$ to one of K groups. A popular way to approach this task is via a parametric finite mixture model. While the vast majority of the work on such mixtures has been based on Gaussian mixture models in many applications the tails of normal distributions are shorter than appropriate or parameter estimations are affected by atypical observations (outliers). The family of location and scale mixtures of Gaussians has the ability to generate a number of flexible distributional forms. It nests as particular cases several important asymmetric distributions like the Generalised Hyperbolic (GH) distribution. The Generalised Hyperbolic distribution in turn nests many other well known distributions such as the Normal Inverse Gaussian (NIG) whose practical relevance has been widely documented in the literature. In a multivariate setting, we propose to extend the standard location and scale mixture concept into a so called multiple scaled framework which has the advantage of allowing different tail and skewness behaviours in each dimension of the variable space with arbitrary correlation between dimensions. The approach builds upon, and develops further, previous work on scale mixtures of Gaussians [21]. Estimation of the parameters is provided via an EM algorithm with a particular focus on NIG distributions. Inference is then extended to cover the case of mixtures of such multiple scaled distributions for application to clustering. Assessments on simulated and real data confirm the gain in degrees of freedom and flexibility in modelling data of varying tail behaviour and directional shape. In addition, comparison with other similar models of GH distributions shows that the later are not as flexible as claimed.

6.2.4. *Bayesian mixtures of multiple scaled distributions*

Participants: Florence Forbes, Alexis Arnaud.

Joint work with: Emmanuel Barbier and Benjamin Lemasson from Grenoble Institute of Neuroscience.

In previous work [21], inference for mixtures of multiple scaled distributions has been carried out based on maximum likelihood principle and using the EM algorithm. In this work we consider a Bayesian treatment of these models for the many advantages that the Bayesian framework offers in the mixture model context. Mainly it avoids the ill-posed nature of maximum likelihood due to the presence of singularities in the likelihood function. A mixture component may collapse by becoming centered at a single data vector sending its covariance to 0 and the model likelihood to infinity. A Bayesian treatment protects the algorithm from this problem occurring in ordinary EM. Also, Bayesian model comparison embodies the principle that states

that simple models should be preferred. Typically, maximum likelihood does not provide any guidance on the choice of the model order as more complex models can always fit the data better. For standard scale mixture of Gaussians, the usual Normal-Wishart prior can be used for the Gaussian parameters. For multiple scaled distributions, the specific decomposition of the covariance requires appropriate separated priors on the eigenvectors and eigenvalues of the scale matrix. Such a decomposition has been already examined in various works on priors for covariance matrix. In this work we consider several possibilities. We derive an inference scheme based on variational approximation and show how to apply this to model selection. In particular, we consider the issue of selecting automatically an appropriate number of classes in the mixtures. We show how to select this number from a single run avoiding the repetitive inference and comparison of all possible models.

6.2.5. *EM for Weighted-Data Clustering*

Participant: Florence Forbes.

Joint work with: Israel Gebru, Xavier Alameda-Pined and Radu Horaud from the Inria Perception team.

Data clustering has received a lot of attention and many methods, algorithms and software packages are currently available. Among these techniques, parametric finite-mixture models play a central role due to their interesting mathematical properties and to the existence of maximum-likelihood estimators based on expectation-maximization (EM). In this work we propose a new mixture model that associates a weight with each observed data point. We introduce a Gaussian mixture with weighted data and we derive two EM algorithms: the first one assigns a fixed weight to each observed datum, while the second one treats the weights as hidden variables drawn from gamma distributions. We provide a general-purpose scheme for weight initialization and we thoroughly validate the proposed algorithms by comparing them with several parametric and non-parametric clustering techniques. We demonstrate the utility of our method for clustering heterogeneous data, namely data gathered with different sensorial modalities, e.g., audio and vision. See also an application in [40].

6.3. Statistical models for Neuroscience

6.3.1. *Physiologically informed Bayesian analysis of ASL fMRI data*

Participants: Florence Forbes, Aina Frau Pascual, Thomas Vincent.

Joint work with: Philippe Ciuciu from Team Parietal and Neurospin, CEA in Saclay.

ASL fMRI data provides a quantitative measure of blood perfusion, that can be correlated to neuronal activation. In contrast to BOLD measure, it is a direct measure of cerebral blood flow. However, ASL data has a lower SNR and resolution so that the recovery of the perfusion response of interest suffers from the contamination by a stronger BOLD component in the ASL signal. In this work [38], [39] we consider a model of both BOLD and perfusion components within the ASL signal. A physiological link between these two components is analyzed and used for a more accurate estimation of the perfusion response function in particular in the usual ASL low SNR conditions.

6.3.2. *Physiological models comparison for the analysis of ASL fMRI data*

Participants: Florence Forbes, Aina Frau Pascual.

Joint work with: Philippe Ciuciu from Team Parietal and Neurospin, CEA in Saclay.

Physiological models have been proposed to describe the processes that underlie the link between neural and hemodynamic activity in the brain. Among these, the Balloon model describes the changes in blood flow, blood volume and oxygen concentration when an hemodynamic response is ensuing neural activation. Next, a *BOLD signal model* links these variables to the measured BOLD signal. Taken together, these equations allow the precise modeling of the coupling between the cerebral blood flow (CBF) and hemodynamic response (HRF). However, several competing versions of BOLD signal model have been described in the past. In this work, we compare different physiological models linking CBF to HRF and different BOLD signal models too in terms of least squares error and log-likelihood, and we assess the impact of this setting in the context of Arterial Spin Labelling (ASL) functional Magnetic Resonance Imaging (fMRI) data analysis.

6.3.3. *Variational EM for the analysis of ASL fMRI data*

Participants: Florence Forbes, Aina Frau Pascual.

Joint work with: Philippe Ciuciu from Team Parietal and Neurospin, CEA in Saclay.

In this work, the goal is to analyse ASL data by accounting jointly for both the BOLD and perfusion components in the signal. Using the model proposed in [77], we design a variational EM approach to estimate the model parameters as a faster alternative to the MCMC approach used in [77] and [39].

6.3.4. *Metaheuristics for the analysis of fMRI data*

Participants: Florence Forbes, Pablo Mesejo Santiago.

Joint work with: Jan Warnking from Grenoble Institute of Neuroscience.

The undergoing work is focused on the optimization of nonlinear models for fMRI data analysis, specially Blood-oxygen-level dependent (BOLD) MR modality. The current optimization procedure consists of a Bayesian inversion of the nonlinear model using a Gauss-Newton/Expectation-Maximization algorithm. Such an optimization procedure is time-consuming and achieves sub-optimal results. Therefore, the current research work is mainly focused on improving these results by experimenting with global search optimization methods, like metaheuristics (MHs). Secondly, MHs can also be of great help in the development of minimization algorithms for solving problems with orthogonality constraints (like in polynomial optimization, combinatorial optimization, eigenvalue problems, sparse PCA, matrix rank minimization, etc.). Thus, another main research line is concerned with the application of MHs to this problem and, if necessary, the design and implementation of new evolutionary operators that preserve orthogonality. And, finally, we are also trying to create advanced statistical models for coupling Arterial Spin Labeling (ASL) and BOLD MR modalities to study brain function.

6.3.5. *Model selection for hemodynamic brain parcellation in fMRI*

Participant: Florence Forbes.

Joint work with: Lotfi Chaari, Mohanad Albughdadi, Jean-Yves Tourneret from IRIT-ENSEEIHT in Toulouse and Philippe Ciuciu from Neurospin, CEA in Saclay.

Brain parcellation into a number of hemodynamically homogeneous regions (parcels) is a challenging issue in fMRI analyses. This task has been recently integrated in the joint detection-estimation (JDE) resulting in the so-called joint detection-parcellation-estimation (JPDE) model. JPDE automatically estimates the parcels from the fMRI data but requires the desired number of parcels to be fixed. This is potentially critical in that the chosen number of parcels may influence detection-estimation performance. In this paper [30], we propose a model selection procedure to automatically fix the number of parcels from the data. The selection procedure relies on the calculation of the free energy corresponding to each concurrent model, within the variational expectation maximization framework. Experiments on synthetic and real fMRI data demonstrate the ability of the proposed procedure to select an adequate number of parcels. We also investigated the use of Latent Dirichlet Processes.

6.3.6. *Partial volume estimation in brain MRI revisited*

Participant: Florence Forbes.

Joint work with: Alexis Roche from Siemens Advanced Clinical Imaging Technology, Department of Radiology, CHUV, Signal Processing Laboratory (LTS5), EPFL, Lausanne, Switzerland.

Image-guided diagnosis of brain disease calls for accurate morphometry algorithms, e.g., in order to detect focal atrophy patterns relating to early-stage progression of particular forms of dementia. To date, widely used brain morphometry packages rest upon discrete Markov random field (MRF) image segmentation models that ignore, or do not fully account for partial voluming, leading to potentially inaccurate estimation of tissue volumes. Although several partial volume (PV) estimation methods have been proposed in the literature from the early 90's, none of them seems to be in common use. In [43], we propose a fast algorithm to estimate brain tissue concentrations from conventional T1-weighted images based on a Bayesian maximum a posteriori

formulation that extends the "mixel" model developed in the 90's. A key observation is the necessity to incorporate additional prior constraints to the "mixel" model for the estimation of plausible concentration maps. Experiments on the ADNI standardized dataset show that global and local brain atrophy measures from the proposed algorithm yield enhanced diagnosis testing value than with several widely used soft tissue labeling methods.

6.3.7. *Tumor classification and prediction using robust multivariate clustering of multiparametric MRI*

Participants: Florence Forbes, Alexis Arnaud.

Joint work with: Emmanuel Barbier and Benjamin Lemasson from Grenoble Institute of Neuroscience.

Advanced statistical clustering approaches are promising tools to better exploit the wealth of MRI information especially on large cohorts and multi-center studies. In neuro-oncology, the use of multiparametric MRI may better characterize brain tumor heterogeneity. To fully exploit multiparametric MRI (e.g. tumor classification), appropriate analysis methods are yet to be developed. They offer improved data quality control by allowing automatic outlier detection and improved analysis by identifying discriminative tumor signatures with measurable predictive power. In this work, we show on small animals data that advanced statistical learning approaches can help 1) in organizing existing data by detecting and excluding outliers and 2) in building a dictionary of tumor fingerprints from a clustering analysis of their microvascular features. Future work should include the integration in a joint statistical model of both automatic ROI delineation and clustering for whole brain data analysis, with a better use of anatomical information. This work has been submitted to the ISMRM 2015 conference and accepted in the SFMRMB 2015 conference [45].

6.4. Markov models

6.4.1. *Identifying Interactions between Tropical Plant Species: A Correlation Analysis of High-Throughput Environmental DNA Sequence Data based on Random Matrix Theory*
Participants: Florence Forbes, Angelika Studeny.

This is joint work with: Eric Coissac and Pierre Taberlet from LECA (Laboratoire d'Ecologie Alpine) and Alain Viari from Inria team Bamboo.

The study of species cooccurence pattern has always been central to community ecology. The rise of high-throughput molecular methods and their use in ecology nowadays allows for a facilitated access to new data of an unprecedented quantity. We address the question about the identification of genuine species interactions in the light of these novel data. The statistical analysis has to be tailored to the data specifics: the large amount of available data as well as biases inherent to the data extraction methods. The latter can cause spurious interactions while the former complicates any statistical modelling approach. In addition, the resolution of the data provided is rarely on the species level. In this work, we conduct a thorough correlation analysis between MOTUs (molecular operating taxonomic unit) on different spatial scales to investigate global as well as local spatial pattern. Although this type of analysis is per se exploratory, we suggest it here in order to separate true species interaction from random pattern and to identify species subgroups for further in detail modelling. A random-matrix approach allows us to derive objective cut-off values for genuine correlations. We compare the results with those derived by the application of a model-based, sparse regression approach. Our study shows that despite their seemingly less precise nature when it comes to species identification, these data enable us to reveal mechanisms that structure an ecological community. In the light of the nowadays facilitated access to molecular data, this points the way to a novel set of efficient methods for community analysis.

6.4.2. *Modelling multivariate counts with graphical Markov models.*
Participant: Jean-Baptiste Durand.

Joint work with: Pierre Fernique (Montpellier 2 University, CIRAD and Inria Virtual Plants) and Yann Guédon (CIRAD and Inria Virtual Plants)

Multivariate count data are defined as the number of items in different states issued from sampling within a population, which individuals own items in various numbers and states. The analysis of multivariate count data is a recurrent and crucial issue in numerous modelling problems, particularly in the fields of biology and ecology (where the data can represent, for example, children counts associated with multitype branching processes), sociology and econometrics. Denoting by K the number of states, multivariate count data analysis relies on modelling the joint distribution of the K-dimensional random vector $N = (N_0, ..., N_{K-1})$ with discrete components. Our work focused on I) Identifying states that appear simultaneously, or on the contrary that are mutually exclusive. This was achieved by identifying conditional independence relationships between the K variables; II)Building parsimonious parametric models consistent with these relationships; III) Characterizing and testing the effects of covariates on the distribution of N, particularly on the dependencies between its components.

Our context of application was characterised by zero-inflated, often right skewed marginal distributions. Thus, Gaussian and Poisson distributions were not *a priori* appropriate. Moreover, the multivariate histograms typically had many cells, most of which were empty. Consequently, nonparametric estimation was not efficient.

We developed an approach based on probabilistic graphical models (Koller & Friedman, 2009 [73]) 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 considered models are based on chain graphs. Model selection procedures are necessary to infer the graph and specify parsimonious distributions. 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. These families were chosen by selection model procedures among different parametric families [36]. To relax the strong constraints regarding dependencies induced by using parametric distributions, mixture of graphical models were also considered [49].

Further extensions will be considered, and particularly

- Hidden Markov tree models (see 6.4.3) where the hidden state process is a multitype branching process with graphical generation distributions.
- Gaussian chain graph models, where the chain components can be identified using lasso methods.

6.4.3. *Statistical characterization of tree structures based on Markov tree models and multitype branching processes, with applications to tree growth modelling.*

Participant: Jean-Baptiste Durand.

Joint work with: Pierre Fernique (Montpellier 2 University and CIRAD) and Yann Guédon (CIRAD), Inria Virtual Plants.

Algorithmic issues in hidden Markov tree models were considered by Durand *et al.* (2004) [68]. This family of models was used to represent local dependencies and heterogeneity within tree-structured data. It relied on a tree-structured hidden state process, where the children states were assumed independent given their parent state. The latter assumption has been relaxed in an extension of these models and new algorithmic solutions for model inference have been proposed in Pierre Fernique's PhD [70]. An application to the study of the cell lineage in biological tissues responsible for the plant growth has been considered. In this setting, the number of children is small (between 0 and 2) and a saturated model has been considered to model transitions between parent and configurations of children states. Extensions will be proposed, based on the parametric discrete multivariate distributions developed in Section 6.4.2 .

6.4.4. *Change-point models for tree-structured data*

Participant: Jean-Baptiste Durand.

Joint work with: Pierre Fernique (Montpellier 2 University and CIRAD) and Yann Guédon (CIRAD), Inria Virtual Plants.

As an alternative to the hidden Markov tree models discussed in Section 6.4.3 , subtrees with similar attributes can be identified using multiple change-point models. These approaches are well-developed in the context of sequence analysis, but their extensions to tree-structured data are not straightforward. Their advantage on hidden Markov models is to relax the strong constraints regarding dependencies induced by parametric distributions and local parent-children dependencies. Heuristic approaches for change-point detection in trees were proposed and applied to the analysis of patchiness patterns (consisting of canopies made of clumps of either vegetative or flowering botanical units) in mango trees [70].

6.4.5. *Hidden Markov models for the analysis of eye movements*

Participant: Jean-Baptiste Durand.

Joint work with: Anne Guérin-Dugué (GIPSA-lab) and Benoit Lemaire (Laboratoire de Psychologie et Neurocognition)

In the last years, GIPSA-lab has developed computational models of information search in web-like materials, using data from both eye-tracking and electroencephalograms (EEGs). These data were obtained from experiments, in which subjects had to make some kinds of press reviews. In such tasks, reading process and decision making are closely related. Statistical analysis of such data aims at deciphering underlying dependency structures in these processes. Hidden Markov models (HMMs) have been used on eye movement series to infer phases in the reading process that can be interpreted as steps in the cognitive processes leading to decision. In HMMs, each phase is associated with a state of the Markov chain. The states are observed indirectly through eye-movements. Our approach was inspired by Simola *et al.* (2008) [76], but we used hidden semi-Markov models for better characterization of phase length distributions. The estimated HMM highlighted contrasted reading strategies (i.e., state transitions), with both individual and document-related variability.

However, the characteristics of eye movements within each phase tended to be poorly discriminated. As a result, high uncertainty in the phase changes arose, and it could be difficult to relate phases to known patterns in EEGs.

As a perspective, we aim at developing an integrated model coupling EEG and eye movements within one single HMM for better identification of the phases. Here, the coupling should incorporate some delay between the transitions in both (EEG and eye-movement) chains, since EEG patterns associated to cognitive processes occur lately with respect to eye-movement phases. Moreover, EEGs and scanpaths were recorded with different time resolutions, so that some resampling scheme must be added into the model, for the sake of synchronizing both processes. Probabilistic graphical models (see Section 6.4.2) will be inferred from the channel correlations to represent interactions between brain zones. The variability of these graphs is partly explained by individual differences in text exploration, which will have to be quantified.

6.4.6. *Hyper-Spectral Image Analysis with Partially-Latent Regression and Spatial Markov Dependencies*

Participant: Florence Forbes.

Joint work with: Antoine Deleforge, Sileye Ba and Radu Horaud from the Inria Perception team.

Hyper-spectral data can be analyzed to recover physical properties at large planetary scales. This involves resolving inverse problems which can be addressed within machine learning, with the advantage that, once a relationship between physical parameters and spectra has been established in a data-driven fashion, the learned relationship can be used to estimate physical parameters for new hyper-spectral observations. Within this framework, we propose a spatially-constrained and partially-latent regression method which maps high-dimensional inputs (hyper-spectral images) onto low-dimensional responses (physical parameters). The proposed regression model comprises two key features. Firstly, it combines a Gaussian mixture of locally-linear mappings (GLLiM) with a partially-latent response model described in [17]. While the former makes high-dimensional regression tractable, the latter enables to deal with physical parameters that cannot be observed or, more generally, with data contaminated by experimental artifacts that cannot be explained with

noise models. Secondly, spatial constraints are introduced in the model through a Markov random field (MRF) prior which provides a spatial structure to the Gaussian-mixture hidden variables. Experiments conducted on a database composed of remotely sensed observations collected from the Mars planet by the Mars Express orbiter demonstrate the effectiveness of the proposed model. A preliminary version of the work can be found in [31].

6.5. Semi and non-parametric methods

6.5.1. *Conditional extremal events*
Participant: Stéphane Girard.

Joint work with: L. Gardes (Univ. Strasbourg), A. Daouia (Univ. Toulouse I and Univ. Catholique de Louvain), J. Elmethni (Univ. Paris 5) and S. Louhichi (Univ. Grenoble 1)

The goal of the PhD thesis of Alexandre Lekina was to contribute to the development of theoretical and algorithmic models to tackle conditional extreme value analysis, *ie* the situation where some covariate information X is recorded simultaneously with a quantity of interest Y. In such a case, the tail heaviness of Y depends on X, and thus the tail index as well as the extreme quantiles are also functions of the covariate. We combine nonparametric smoothing techniques [71] with extreme-value methods in order to obtain efficient estimators of the conditional tail index and conditional extreme quantiles. The strong consistency of such estimator is established in [53]. When the covariate is functional and random (random design) we focus on kernel methods [58].

Conditional extremes are studied in climatology where one is interested in how climate change over years might affect extreme temperatures or rainfalls. In this case, the covariate is univariate (time). Bivariate examples include the study of extreme rainfalls as a function of the geographical location. The application part of the study is joint work with the LTHE (Laboratoire d'étude des Transferts en Hydrologie et Environnement) located in Grenoble.

6.5.2. *Estimation of extreme risk measures*
Participant: Stéphane Girard.

Joint work with: E. Deme (Univ. Gaston-Berger, Sénégal, J. Elmethni (Univ. Paris 5), L. Gardes and A. Guillou (Univ. Strasbourg)

One of the most popular risk measures is the Value-at-Risk (VaR) introduced in the 1990's. In statistical terms, the VaR at level $\alpha \in (0, 1)$ corresponds to the upper α-quantile of the loss distribution. The Value-at-Risk however suffers from several weaknesses. First, it provides us only with a pointwise information: VaR(α) does not take into consideration what the loss will be beyond this quantile. Second, random loss variables with light-tailed distributions or heavy-tailed distributions may have the same Value-at-Risk . Finally, Value-at-Risk is not a coherent risk measure since it is not subadditive in general. A coherent alternative risk measure is the Conditional Tail Expectation (CTE), also known as Tail-Value-at-Risk, Tail Conditional Expectation or Expected Shortfall in case of a continuous loss distribution. The CTE is defined as the expected loss given that the loss lies above the upper α-quantile of the loss distribution. This risk measure thus takes into account the whole information contained in the upper tail of the distribution. It is frequently encountered in financial investment or in the insurance industry. In [52], we have established the asymptotic properties of the CTE estimator in case of extreme losses, *i.e.* when $\alpha \to 0$ as the sample size increases. We have exhibited the asymptotic bias of this estimator, and proposed a bias correction based on extreme-value techniques. In [20], we study the situation where some covariate information is available. We thus has to deal with conditional extremes (see paragraph 6.5.1). We also proposed a new risk measure (called the Conditional Tail Moment) which encompasses various risk measures, such as the CTE, as particular cases.

6.5.3. *Multivariate extremal events*
Participants: Stéphane Girard, Gildas Mazo, Florence Forbes.

Joint work with: C. Amblard (TimB in TIMC laboratory, Univ. Grenoble I), L. Gardes (Univ. Strasbourg) and L. Menneteau (Univ. Montpellier II)

Copulas are a useful tool to model multivariate distributions [75]. At first, we developed an extension of some particular copulas [1]. It followed a new class of bivariate copulas defined on matrices [55] and some analogies have been shown between matrix and copula properties.

However, while there exist various families of bivariate copulas, much fewer has been done when the dimension is higher. To this aim an interesting class of copulas based on products of transformed copulas has been proposed in the literature. The use of this class for practical high dimensional problems remains challenging. Constraints on the parameters and the product form render inference, and in particular the likelihood computation, difficult. We proposed a new class of high dimensional copulas based on a product of transformed bivariate copulas [64]. No constraints on the parameters refrain the applicability of the proposed class which is well suited for applications in high dimension. Furthermore the analytic forms of the copulas within this class allow to associate a natural graphical structure which helps to visualize the dependencies and to compute the likelihood efficiently even in high dimension. The extreme properties of the copulas are also derived and an R package has been developed.

As an alternative, we also proposed a new class of copulas constructed by introducing a latent factor. Conditional independence with respect to this factor and the use of a nonparametric class of bivariate copulas lead to interesting properties like explicitness, flexibility and parsimony. In particular, various tail behaviours are exhibited, making possible the modeling of various extreme situations [42]. A pairwise moment-based inference procedure has also been proposed and the asymptotic normality of the corresponding estimator has been established [66].

In collaboration with L. Gardes, we investigate the estimation of the tail copula which is widely used to describe the amount of extremal dependence of a multivariate distribution. In some situations such as risk management, the dependence structure can be linked with some covariate. The tail copula thus depends on this covariate and is referred to as the conditional tail copula. The aim of our work is to propose a nonparametric estimator of the conditional tail copula and to establish its asymptotic normality [57].

6.5.4. *Level sets estimation*
Participant: Stéphane Girard.

Joint work with: A. Guillou and L. Gardes (Univ. Strasbourg), A. Nazin (Univ. Moscou), G. Stupfler (Univ. Aix-Marseille) and A. Daouia (Univ. Toulouse I and Univ. Catholique de Louvain)

The boundary bounding the set of points is viewed as the larger level set of the points distribution. This is then an extreme quantile curve estimation problem. We proposed estimators based on projection as well as on kernel regression methods applied on the extreme values set, for particular set of points [10]. We also investigate the asymptotic properties of existing estimators when used in extreme situations. For instance, we have established in collaboration with G. Stupfler that the so-called geometric quantiles have very counter-intuitive properties in such situations [63], [62] and thus should not be used to detect outliers. These resuls are submitted for publication.
In collaboration with A. Daouia, we investigate the application of such methods in econometrics [16]: A new characterization of partial boundaries of a free disposal multivariate support is introduced by making use of large quantiles of a simple transformation of the underlying multivariate distribution. Pointwise empirical and smoothed estimators of the full and partial support curves are built as extreme sample and smoothed quantiles. The extreme-value theory holds then automatically for the empirical frontiers and we show that some fundamental properties of extreme order statistics carry over to Nadaraya's estimates of upper quantile-based frontiers.
In collaboration with A. Nazin, we define new estimators of the frontier function based on linear programming methods. The frontier is defined as the solution of a linear optimization problem under inequality constraints. The estimator is shown to be strongly consistent with respect to the L_1 norm and we establish that it reaches the optimal minimax rate of convergence [58].

In collaboration with G. Stupfler and A. Guillou, new estimators of the boundary are introduced. The regression is performed on the whole set of points, the selection of the "highest" points being automatically performed by the introduction of high order moments [22].

6.5.5. *Retrieval of Mars surface physical properties from OMEGA hyperspectral images.*

Participants: Stéphane Girard, Alessandro Chiancone.

Joint work with: S. Douté from Laboratoire de Planétologie de Grenoble, J. Chanussot (Gipsa-lab and Grenoble-INP) and J. Saracco (Univ. Bordeaux).

Visible and near infrared imaging spectroscopy is one of the key techniques to detect, to map and to characterize mineral and volatile (eg. water-ice) species existing at the surface of planets. Indeed the chemical composition, granularity, texture, physical state, etc. of the materials determine the existence and morphology of the absorption bands. The resulting spectra contain therefore very useful information. Current imaging spectrometers provide data organized as three dimensional hyperspectral images: two spatial dimensions and one spectral dimension. Our goal is to estimate the functional relationship F between some observed spectra and some physical parameters. To this end, a database of synthetic spectra is generated by a physical radiative transfer model and used to estimate F. The high dimension of spectra is reduced by Gaussian regularized sliced inverse regression (GRSIR) to overcome the curse of dimensionality and consequently the sensitivity of the inversion to noise (ill-conditioned problems) [15]. We have also defined an adaptive version of the method which is able to deal with block-wise evolving data streams [13].

In his PhD thesis work, Alessandro Chiancone studies the extension of the SIR method to different sub-populations. The idea is to assume that the dimension reduction subspace may not be the same for different clusters of the data [46]. He also published a paper on a previous work in the field of hierarchical segmentation of images [14].

<p align="center" style="color:red">NANO-D Project-Team</p>

5. New Results

5.1. Variance Analysis of ARPS-Langevin dynamics

Participants: Zofia Trstanova, Gabriel Stoltz, Stephane Redon.

In order to analyze statistical convergence speed-up that can be achieved by using Adaptively Restrained Particle Simulations (ARPS) dynamics, we proposed a formula that combines the variance of the sampled process and the algorithmic speed-up:

$$S_\sigma = S_A \frac{\sigma_0^2}{\sigma_\epsilon^2} \tag{1}$$

where S_σ is the convergence speed-up, S_A is the algorithmic speed-up, σ_0^2 is the variance of the original system and σ_ϵ^2 is the variance of the ARPS-Langevin system. This led to a need of a detailed analysis of the variance of ARPS-Langevin process. We performed many numerical simulations, from the simple one-dimensional case up to more real- istic dimer-solvent models, in order to observe the behavior of the variance and the quantitative dependence on the ARPS coefficients. For the one-dimensional case we managed to compute by using Galerkin approximations the numerical approximation of the variance. We are also studying analytically by use of standard techniques the properties of the ARPS-Langevin dynamics such as the existence of an invariant measure. We are also interested in the relationship between the variance of the Langevin dynamics and the ARPS-Langevin dynamics. We showed that for small ARPS coefficients the ARPS-Langevin process can be seen as a perturbation of a standard Langevin process by a perturbation operator that depends on the ARPS coefficient ϵ.

5.2. Parallel adaptively restrained particle simulations

Participants: Krishna Kant Singh, Stephane Redon.

We have continued our work on the development of *parallel* adaptively restrained particle simulations. We have integrated the ARPS algorithm in LAMMPS (Large-scale Atomic/ Molecular Massively Parallel Simulator). LAMMPS is a computationally efficient simulator, which contains a wide range of potentials and force fields for simulating systems like solid-state materials (metals, semiconductors), soft matter (biomolecules, polymers) and coarse-grained or mesoscopic systems.

In order to verify our implementation of ARPS in LAMMPS, we have generated a trajectory of 1 ns by simulating 108 Argon particles using the ARPS algorithm and the NVE ensemble (constant Number of particles, Volume and Energy). All the particles were placed in an orthogonal box with a side length of 17.158 angstrom. We used periodic boundary conditions with 8.5 angstrom cut-off for the Lennard-Jones potential. We used a threshold $\epsilon_r = 0.0000001$ for applying restraints and a threshold $\epsilon_f = 0.005$ for releasing restraints. The system was simulated at different step sizes: using $0.5, 1, 2, 3, 4, 5, 10, 50, 60, 70, 80$ and 90 femtoseconds.

Our results show that ARPS in LAMMPS preserves the total energy during simulation (Figure 4) as well as the radial distribution function (Figure 5). We are now in the process of modifying the parallel force calculation algorithms in LAMMPS to make them incremental, i.e. make their cost proportional to the number of active particles in the simulation at a given time.

5.3. Molecular Modeling

5.3.1. *The CARBON method*

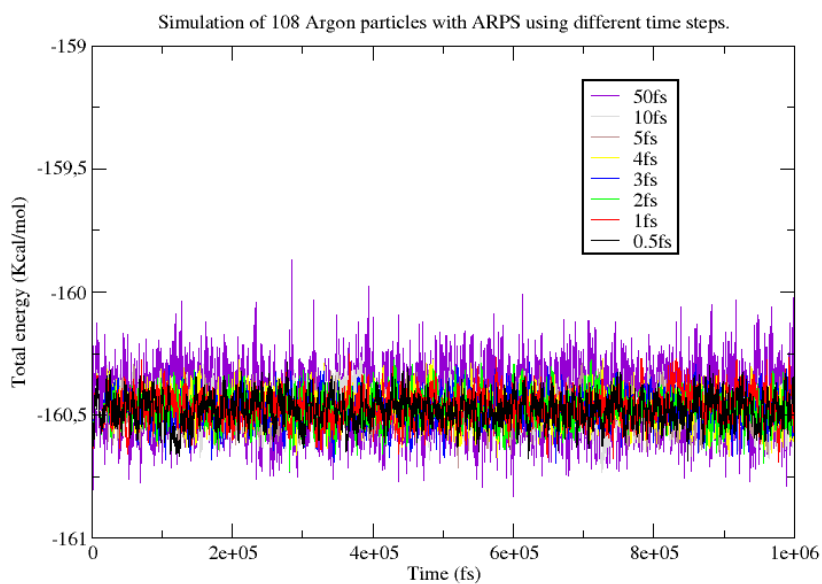**Participants:** Sergei Grudinin, Stephane Redon, Petr Popov.

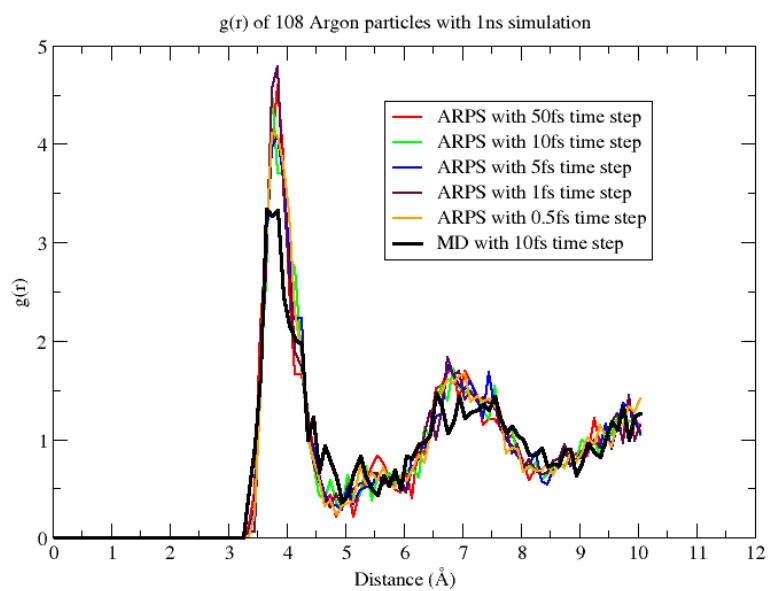*Figure 4. Energy conservation in LAMMPS using ARPS.*

Figure 5. Preservation of the radial distribution function in LAMMPS using ARPS.

In molecular docking, various refinement algorithms are implied either to take into account flexibility of molecular complexes or to get rid of the docking artefacts, e.g. steric clashes. To address the latter problem, one possibility is to continuously minimize the energy of the complex with respect to the affine transformations, i.e. rigid transformations. Petr Popov developed a fast and efficient method called CARBON, where one considers the rigid-body optimization problem as the calculation of quasi-static trajectories of rigid bodies influenced by the inverse-inertia-weighted energy gradient. In order to determine the appropriate step-size in the direction of the net generalized force, we introduce the concept of advancement region, which is the interval of step-sizes that provide movements of the rigid body within a certain range of root mean square deviation from the initial conformation. We tested and validated CARBON on several benchmarks using both a classical force-field and a knowledge-based scoring function and demonstrated that CARBON significantly improves the quality of docking pre- dictions an also remains stable when monomers of a molecular complex significantly overlap. CARBON will be made available as a SAMSON Element for the SAMSON software platform at http://www.samson-connect.net.

5.3.2. *The KSENIA method*

Participants: Petr Popov, Sergei Grudinin.

Molecular docking as an integral part of the drug discovery involves the scoring stage, where one selects the best binding candidates from the set of ligand poses. The scoring stage incorporates sophisticated scoring functions based on the empirical force-fields or the information derived from known structures of protein complexes. The latter type of scoring functions belongs to the family of the knowledge-based or statistical scoring functions. Typically, for the training of a knowledge-based scoring function, modern methods require an ensemble of generated non-native decoy structures and a computation of the reference state, which is challenging. Petr Popov developed a method that does not require neither the computation of the reference state nor the ensemble of non-native complexes. Furthermore, the developed approach fully relies on the structures of protein complexes in their native configurations. More precisely, Petr trained the knowledge-based scoring function based on sets of near-native conformations. These are composed using the fluctuations along the direction of low-frequency normal modes computed at the native configurations. The obtained scoring function is capable to distinguish the native and near-native protein-protein interactions from the non-native ones. The robustness of the method was verified on several protein-protein docking benchmarks. Our methodology can be easily adapted to the recognition of other types of molecular interactions, such as protein-ligand, protein-RNA, etc. KSENIA will be made publicly available as a part of the SAMSON software platform at http://www.samson-connect.net.

5.3.3. *Optimization solvers*

Participants: Petr Popov, Anatoli Juditsky, Sergei Grudinin.

To derive a knowledge-based scoring function, we map non-native and near-native molecular complexes to the vectors of descriptors in a high- dimensional space. In this space, we formulate an optimization problem to construct the scoring function in such a way, that the projection of a descriptor vector onto the scoring vector corresponds to the score of a molecular complex. The formulated problem contains the regularization term and the penalty term and might vary depending on the method applied to solve the optimization problem. Different methods provide different convergence rates and cost per operation. We implemented several modern first- and second- order optimization techniques and explored which one works the best on the given data. Namely, we tested the standard gradient descent method, the conjugate gradients method, the Nesterov method, the Fista and Fista-descent methods, and the proximal gradient method.

5.3.4. *Novel Docking Criterion*

Participants: Petr Popov, Sergei Grudinin.

Generally, to assess the prediction capabilities of a scoring function for protein- protein interactions, one evaluates the success rate of the scoring function on widely used protein-protein benchmarks. The percentage of correctly predicted complexes is taken as the characteristic of the scoring function. However, all existing benchmarks nowadays consists on many non-native and only few near- native conformations. However, the ability of the scoring function to distinguish a particular near-native conformation from the non-native decoys does not guarantee that the scoring function is able to distinguish another near-native conformation. The same is applied if the scoring function fails on a particular molecular complex. Thus, the success rate is not a robust criterion, since it depends on the near-native and non-native conformations presented in the benchmark. We proposed the new robust method to evaluate the predictive capability of a scoring function, which does not suffer from such drawback. The method uses the probability density function of the score computed from the set of the near-native conformations and complementary empirical distribution function of the score computed from the set on non-native conformations. We tested the criterion on the previously derived scoring functions and showed that the criterion also provides an insight on some limits and restrictions of the atom-atom distant-dependent knowledge-based scoring functions.

5.4. Flexible molecular fitting

Participants: Alexandre Hoffmann, Sergei Grudinin.

We have started a PhD on flexible molecular fitting. The first part of the PhD aims at developing a new method for non-rigid molecular fitting. The problem is the following : We have two proteins $\mathcal{P}_1$ and $\mathcal{P}_2$ and we know $d_1 : \mathbb{R}^3 \to \mathbb{R}$, the electron density of $\mathcal{P}_1$ and $(Y_k)_{k=0\cdots N_{atoms}-1}$, the average positions of the atoms of $\mathcal{P}_2$. Assuming we can generate an artificial electron density $d_2 : \mathbb{R}^3 \to \mathbb{R}$ from $(Y_k)_{k=0\cdots N_{atoms}-1}$, our problem is to find a transformation of the atoms $T : \mathbb{R}^3 \to \mathbb{R}^3$ that minimizes the L^2 distance between d_1 and d_2.

In image processing this problem is usually solved using the optimal transport theory, but this method assumes that both densities have the same L^2 norm, which is not necessarily the case for the fitting problem. To solve this problem, one instead starts by splitting T into a rigid transformation T_{rigid} (which is a combination of translation and rotation) and a flexible transformation $T_{flexible}$. Two classes of methods have been developed to find T_{rigid} :

- the first one uses optimization techniques such as gradient descent, and
- the second one uses Fast Fourier Transform (FFT) to compute the Cross Correlation Function (CCF) of d_1 and d_2.

We have already developed several algorithms based on the FFT to find T_{rigid} and we now want to develop an efficient algorithm to find $T_{flexible}$.

The majority of algorithms first finds the best T_{rigid} and then use Normal Mode Analysis (NMA) to improve their fitting, the problem with such a method is that one can miss the optimal solution. We aim at developing a method that uses convex optimization to find the best $T_{flexible}$ for each T_{rigid} sampled on a grid, and therefore find the best T possible on a grid.

The rest of the PhD will be focused on the improvement of the modeling of the atom's motion, by using machine learning algorithms and methods that go beyond linear NMA. We hope that such an improvement can improve the quality of the fitting method.

5.5. PEPSI-Dock : Fast predictions of putative docking poses using accurate knowledge-based potentials functions to describe interaction between proteins

Participants: Emilie Neveu, Sergei Grudinin, David Ritchie, Petr Popov.

Many biological tasks involve finding proteins that can act as an inhibitor for a virus or a bacteria, fir example. Such task requires knowledge on the structure of the complex to be formed. Protein Data Bank can help but only a small fraction of its proteins are complexes [16]. Therefore, computational docking predictions, being low-cost and easy to perform, are very attractive if they describe accurately the interactions between proteins while being fast to find which conformation will be the most probable. We have been developing a fast and accurate algorithm that combines the FFT-accelerated docking methods with the precise knowledge-based potential functions describing interactions between the atoms in the proteins.

Docking methods can be described as a two ingredients recipe. First, a certain approximation for the binding free energy needed to describe the interactions between the proteins. Second, an efficient sampling algorithm is used to find the lowest-energy conformations. Commonly, as going through all the possibilities with a realistic energy function is very costly, it is approximated with a very simple energy function. Then, a much more precise energy function is typically used to re-score the most promising predictions.Considering the numerous local minima that can be found, it is important to use the most accurate free energy from the beginning not to miss some important docking solutions. In the *Hex* code, an exhaustive search combined with a spherical polar Fourier representation enables the fast exploration of all the conformations. By now it is still the most efficient and reliable search algorithm [21]. However, only a few types of energies have been accelerated using this technic (shape complementarity and electrostatics, for example). Knowledge-based potential functions are much more precise but have been used only at the re-scoring stage of the protein docking predictions pipeline. Thus, our aim is to take advantage of the fast exhaustive search by integrating the very-detailed knowledge-based potentials into the *Hex* exhaustive search method.

We have demonstrated that we can adapt the machine learning process so that the knowledge-based potentials describing atom interactions can be translated into the polynomial basis used in *Hex*. Then, the knowledge-based scores are calculated in *Hex* using the fast polynomial expansions accelerated by the fast Fourier transform. The current evaluations of the knowledge-based scores takes more time than a shape+electrostatic representation but is still fast. More precisely, docking predictions for a single complex takes on average 5-10 minutes on a regular laptop computer. The preliminary results on the data set used for training shows significant improvements in accuracy of the method. Indeed, considering the prediction is correct if its Root Mean Square distance from the true solution is smaller than 5 Å, we currently obtain more than 50% of correct predictions rank first.

5.6. Extended Universal Force Field

Participants: Svetlana Artemova, Leonard Jaillet, Stephane Redon.

In parallel with the implementation of a Universal Force Field module in SAMSON (see Section 5.10.3), we have developed an extension of this force field to allow soft transitions for both topologies and atoms' typizations. In classical UFF topologies and atoms' typizations are set in the initialization phase and remain fixed for the entire simulation. In the proposed extension, they can vary continuously to allow the transition from one given topology to another (see Figure 6). This extended UFF combined with the interaction modeling tools already present in SAMSON allows to interactively build and modify molecules while being driven by UFF forces to ensure the relevance of the corresponding structures. The validity of this extended version of UFF was also tested on the same type of benchmarks as those used to test UFF.

5.7. Incremental Algorithms for Orbital-Free Density Functional Theory

Participants: François Rousse, Stephane Redon.

We have started a new PhD to develop incremental algorithms for electronic structure calculation.
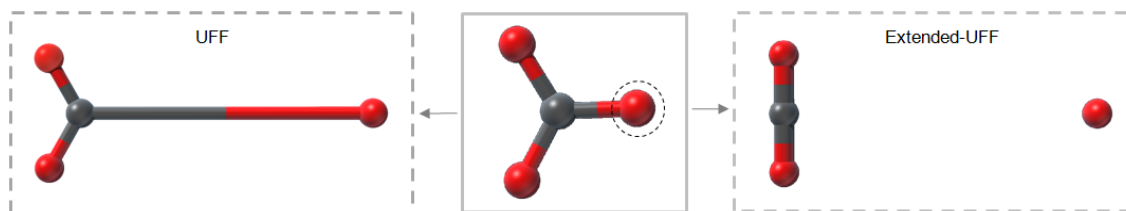
Figure 6. An oxygen atom (dashed circle) of the carbonate ion CO_3^{2-} is displaced using the interactive simulation framework in SAMSON (center). With standard UFF, the topology remains unchanged which leads to unrealistic geometries (left). With extended UFF, the covalent bond is broken forming a Carbon dioxide CO_2 and an isolated Oxygen (right).

Density Functional Theory (DFT) permits to simulate the electronic structure of a molecular system without solving the Schrödinger equation, but by finding incrementally the electronic density that minimizes the system's energy. The most used method is based on the determination of molecular orbitals. It has been shown to be an accurate method but the computation of the energy makes it too slow for the study of big systems ($> 10^3$ atoms) or dynamical ones. The Orbital-Free DFT, although less precise, is faster and can simulate the electronic density of systems up to 10^6 atoms. The aim of the PhD research is to develop new algorithms for Orbital-Free DFT that are *incremental*, i.e. whose complexity depends on the atoms that are adaptively simulated.

5.8. Robotics-inspired methods for large nanosystems

Participants: Minh Khoa Nguyen, Leonard Jaillet, Stephane Redon.

We have started a new PhD to develop robotics-inspired methods for modeling and simulating large nanosystems. Several motion planning methods issued from robotics have been successfully applied to solve problems in the field of biological molecular systems such as, including probabilistic roadmap and rapidly-exploring random trees [12]. However, large systems are still challenging due to the high number of degree of freedom. Our aim is to apply dimensionality reduction methods and/or smart conformational-space exploration techniques inspired from robotics to overcome this difficulty. The PhD topic has started since 1 Oct 2014. Reviews of the state of art and preliminary implementations have been done.

5.9. Incremental algorithms for long-range interactions

Participants: Semeho Edorh, Stephane Redon.

We have started a PhD to develop incremental algorithms for calculating long-range molecular interactions. Numerical simulation of molecular dynamics are very expensive in terms of CPU resources, especially because of the evaluation of the interaction potential. In large crystalline ionic systems, *Ewald summation* is the most popular method for computing Coulombic interactions. It rewrites the interaction potential ϕ as the sum of two terms: $\phi(r) = \phi_{dir}(r) + \phi_{rec}(r)$. The so-called "short-range" contribution ϕ_{dir} can be easily calculated in a direct space , where as the "long-range" contribution ϕ_{rec} is calculated using a Fourier transform.

Direct evaluation of the Ewald summation is an order N^2 computational problem. Over the past three decades, many techniques were developed and reduced the evaluation of the potential to an order $N \log(N)$ problem. We want to develop a new approach that can reduce the computational cost by using incremental algorithms. The key idea is to use, at each time step of the simulation, information that has been computed in previous steps.

5.10. Software development of SAMSON

5.10.1. Development of SAMSON Connect

Participants: Mohamed Yengui, Jocelyn Gate, Stephane Redon.

We have continued the development of SAMSON Connect, the web site that will contribute to the diffusion and promotion of SAMSON and SAMSON Elements (modules for SAMSON).

SAMSON Elements are adapted to different application domain and help users build new models, perform calculations, run interactive or offline simulations, visualize and interpret results, etc. The goal of SAMSON Connect is to bring together a set of users and developers of SAMSON Elements in all areas of nanoscience (physics, biology, chemistry, electronics, etc...). It offers a set of features available depending on the user role:

- Developers (who have obtained the SAMSON-SDK) can develop SAMSON Elements and upload them to SAMSON Connect through the tools provided.
- Users (who have obtained the SAMSON Core application) can add SAMSON Elements to their instance of SAMSON Core in one click. The download process is performed during startup of SAMSON and without outside intervention.

All users can give feedbacks, review and rate SAMSON Elements after adding them to their SAMSON Core (Figure 7).
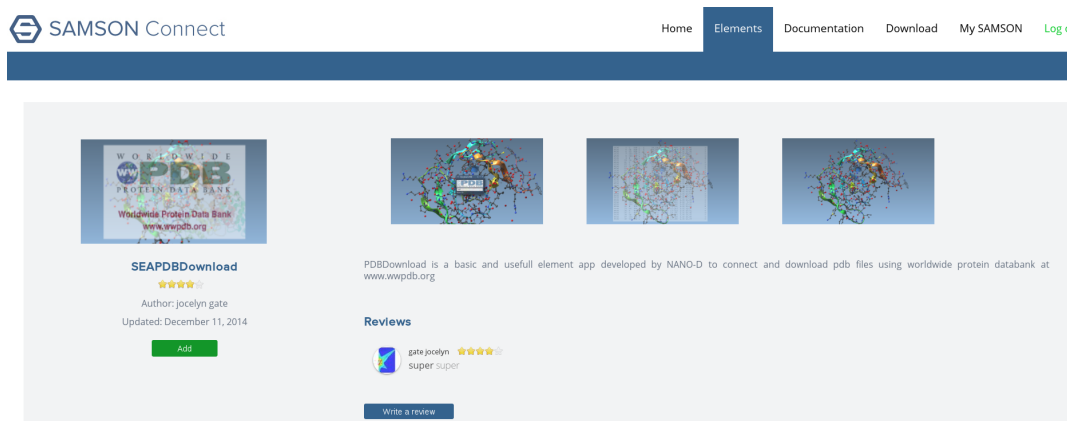
Figure 7. Screenshot of a SAMSON Element on SAMSON Connect.

SAMSON Connect also features some documentation to develop new elements for SAMSON (Figure 8).

SAMSON Connect will be available at http://samson-connect.net.

5.10.2. Deployment of SAMSON and the SAMSON SDK

Participants: Jocelyn Gate, Mohamed Yengui, Stephane Redon.

The SAMSON installer has been split in two parts: SAMSON-setup (installation of the SAMSON application, Figure 9) and SAMSON-Developer-setup (installation of the SAMSON SDK). internet. It is very useful to increase security.

Several helper tools related to SAMSON Elements management were developed to facilitate Element deployment. For example, the element packager is a tool useful for developers who want to distribute a new SAMSON Element on the SAMSON Connect platform. With this packager we can control many things: check whether the file is valid, if the SAMSON Element is readable with SAMSON, add a description file that contains useful information (name, author ID, checksum, element version, SDK version, operating system, etc.).

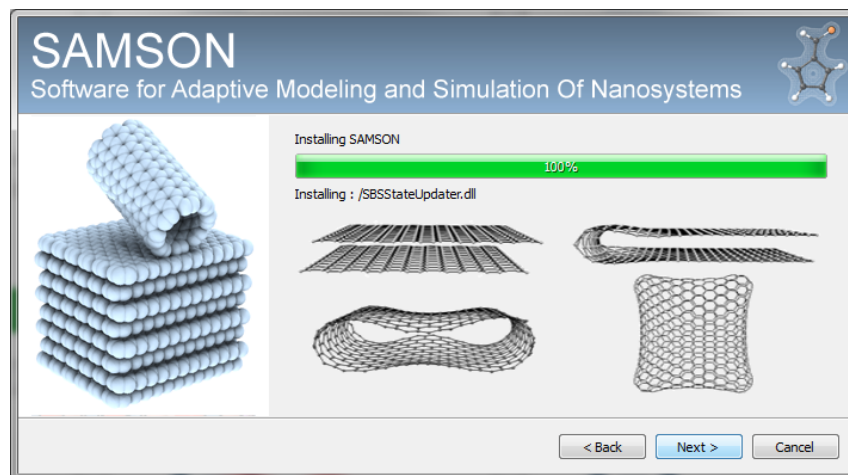*Figure 8. Screenshot of documentation on SAMSON-Connect.*

Figure 9. The SAMSON Installer

We added a service requester to SAMSON to communicate with SAMSON Connect and

- Check users/developers status
- Easily download new SAMSON Elements
- Be notified about updates

5.10.3. Universal Force Field

Participants: Svetlana Artemova, Leonard Jaillet, Stephane Redon.

We have implemented a version of the Universal Force Field (UFF) [19] in SAMSON, as a SAMSON Element embedding an interaction model. UFF is a classical force field, which can take as input almost every atom of the periodic table. Such flexibility allows to potentially use UFF on a large spectrum of systems and since its introduction, it has been applied to simulate problems involving main group compounds, organic molecules, metal complexes and has even been recently extended to MOF (Metal Organic Framework) [11]. The general energy expression for UFF as described in [19] is:

$$E_{UFF} = E_R + E_\theta + E_\phi + E_\omega + E_{vdw} + E_{el},$$

where E_R stands for bond stretching, E_θ describes angle bending, E_ϕ is dihedral angle torsion term, E_ω represents inversion, E_{vdw} stands for van der Waals interactions and E_{el} represents electrostatics (this last term is rarely considered for UFF, we do not study it neither). Forces involved in the atoms interactions can then be derived from the previous expression. Each energetic term in UFF can be computed based on simple rules deduced from a set of parameters. This set is based on the atoms' elements, their hybridization, and the overall connectivity of the molecular system.

In our implementation, we took into account several corrections and refinements that have been lately proposed in the literature for Universal Force Field. Our contribution also concerns the development of algorithms to automatically perceive the system's topology (covalent bonds and bond orders assignments). Moreover, we have introduced a method to automatically find the correct typization of the atoms. Precisely, atoms' hybridizations and oxidation states are computed, and resonance groups within or out of cycles are detected and treated. The implementation provided is computationally efficient enough to allow interactive simulation in SAMSON. The validity of the force field was tested on several groups of molecules proposed as benchmarks in the literature.

5.10.4. Integration of existing tools

Participants: Nadhir Abdellatif, Svetlana Artemova, Stephane Redon.

We have obtained funding from the Nanosciences Foundation in Grenoble to integrate in SAMSON some tools developed and used by the Grenoble community, in the form of SAMSON Elements, i.e. modules that integrate into SAMSON and may interact with SAMSON's main data graph. In particular, we have been meeting with some biologists and physicists to determine which tools and methods used (or developed) in Grenoble would be most appropriate for integration.

We integrated our first Element which is Babel, a chemical toolbox designed to "speak the many languages of chemical data", i.e. read, write and convert data files (over 110 chemical file formats) from molecular modeling, chemistry, solid-state materials, biochemistry, or related areas (see http://openbabel.org). The corresponding SAMSON element is an *app* that delegates all calculations to the Babel external executable. The app also makes it possible to import the data files to SAMSON to visualize the molecular data and proceed with other SAMSON elements.

We have also integrated Clustal, a tool for multiple sequence alignment. Thanks to Clustal's license, all source code is wrapped into the SAMSON Element (whose source code will be made available as well), and SAMSON users do not need to install Clustal separately.

5.10.5. Various

Participants: Stephane Redon, Svetlana Artemova, Marc Aubert.

- Units management was added to SAMSON. The mechanism relies on C++ template meta-programming techniques to perform dimensional analysis and automatic conversions at compile time, and has no runtime overhead. This was a significant undertaking, but one that will be very helpful to integrate in SAMSON different domains of nanoscience that have come to use different units for identical dimensions (e.g. kilocalories per mole in biology, electron volts in chemistry, etc.).

- SAMSON's reflection mechanism was improved to perform type registration and casting, and facilitate scripting and pipelining of SAMSON Elements.

- SAMSON now handles multiple documents.

- SAMSON has its own file format, which allows it to save the data graph information.

- More data graph nodes are now visible in SAMSON's data graph view.

- The split between classical and quantum interaction models was abandoned, for simplicity.

- SAMSON now handles multiple cameras.

- Selection methods have been improved, and selection is now undoable. Selections may be saved, retrieved, have boolean operations performed onto them, etc.

- The documentation of the SAMSON SDK has been improved.

- Controllers, a new type of data graph nodes, were added to SAMSON. Controllers are used to act on other data graph nodes (e.g. translate and rotate models).

- The object lifecycle of SAMSON was improved.

- SAMSON now has a mechanism for serialization.

- SAMSON now has preferences (e.g. for rendering).

- Existing parsers for input and output of molecular information in SAMSON have been improved and accelerated, and property windows for these parsers have been added.

- The Lennard-Jones potential has been added as an interaction model to SAMSON.

- A new editor for adding atoms corresponding to a chemical formula (in disorder) has been created.

- The work on a new editor containing functional groups and frequently-used molecular patterns has been started.

- Periodic Boundary Conditions (an important concept in molecular simulations) were implemented in SAMSON.

- General code debugging and improvement has been performed

- We decided to use the Qt5 framework for shaders management, for some maintenance reasons especially. This structure implied some other type changes to adapt to Qt5, such as the vertex buffers.

- We changed the way viewports display text. It is now possible to run SAMSON on every platform (Windows, Linux and Mac) and display text, and it provides Elements programmers a simple way to add text where they want in the 3D view.

<p style="text-align: center">**NECS Project-Team**</p>

6. New Results

6.1. Highlights of the Year

- C. Canudas de Wit serves as General Chair for the Europeen Control Conference (ECC'14), Strasbourg, France, Jul. 2014 (http://www.ecc14.eu/).
- Launch of the SPEEDD EU FP7 project in Feb. 2014.
- Launch of the COMFORT project, which supports the associate Team between Inria project-team NeCS and the Berkeley University project PATH (http://necs.inrialpes.fr/pages/projects/comfort. php).
- Launch of the LOCATE-ME Persyval project (Apr. 2014 to Aug. 2015) in collaboration with the Tyrex team.
- The team has organized the Hycon2 Show day in May 2014 (http://www.inria.fr/en/centre/grenoble/ calendar/hycon2-show-day-traffic-modeling-estimation-and-control).

6.2. Networked systems and graph analysis

6.2.1. *Distributed solution to the network reconstruction problem*

Participants: A. Kibangou [Contact person], F. Morbidi.

It has been recently shown in [45] that by collecting noise-contaminated time series generated by a coupled-oscillator system at each node of a network, it is possible to robustly reconstruct its topology, i.e. determine the graph Laplacian. Restricting ourselves to linear consensus dynamics over undirected communication networks. In [18], we have introduced a new dynamic average consensus least-squares algorithm to locally estimate these time series at each node, thus making the reconstruction process fully distributed and more easily applicable in the real world. We have also proposed a novel efficient method for separating the off-diagonal entries of the reconstructed Laplacian, and examined several concepts related to the trace of the dynamic correlation matrix of the coupled single integrators, which is a distinctive element of our network reconstruction method.

6.2.2. *Distributed estimation of Laplacian eigenvalues and network robustness assessment*

Participants: A. Kibangou [Contact person], T.-M. D. Tran, J. Hendrickx [Univ. Louvain-la-neuve].

As recently shown, Laplacian eigenvalues can be estimated by solving the factorization of the average consensus Matrix [46]. The problem was viewed as a constrained consensus optimization one. The main assumption was about the knowledge of the final consensus value. Indeed, estimation of the Laplacian eigenvalues can be carried out using measurements of the transient of the consensus protocol and the steady state (consensus value). In [34], we relaxed the assumptions by considering that the consensus value is only approximately known. We formulated a convex optimization, which allowed us to make use of recent well-known techniques and results dealing with convex optimization problem proposed in the literature (the Alternating Direction of Multipliers Method, ADMM), [40], [42]. Recently, we assumed that the consensus value is completely unknown and has to be found simultaneously with Laplacian eigenvalues. In such a case the problem becomes a convex combination problem where the cost function comprises two terms, one that is average consensus problem, and the rest is the consensus problem to estimate the Laplacian eigenvalues. The simulations indicate that the proposed algorithm is efficient enough to provide the nonzero distinct Laplacian eigenvalues with high accuracy. These eigenvalues are then used to assess the robustness of the graph by means of some spectral metrics, the number of spanning trees and the Kirchoff index precisely.

6.2.3. *Observability and privacy preserving features in consensus networks*

Participants: A. Kibangou [Contact person], C. Commault [Grenoble INP].

In [16], we have studied of observability in consensus networks modeled with strongly regular graphs or distance regular graphs. The first result consists in a Kalman-like simple algebraic criterion for observability in distance regular graphs. This criterion consists in evaluating the rank of a matrix built with the components of the Bose-Mesner algebra associated with the considered graph. Then, we have defined some bipartite graphs that capture the observability properties of the graph to be studied. In particular, we showed that necessary and sufficient observability conditions are given by the nullity of the so-called local bipartite observability graph (resp. local unfolded bipartite observability graph) for strongly regular graphs (resp. distance regular graphs). When the nullity cannot be derived directly from the structure of these bipartite graphs, the rank of the associated bi-adjacency matrix allows evaluating observability. Eventually, as a by-product of the main results we have shown that non-observability can be stated just by comparing the valency of the graph to be studied with a bound computed from the number of vertices of the graph and its diameter. Similarly non-observability can also be stated by evaluating the size of the maximum matching in the above mentioned bipartite graphs. Non-observability is strongly linked to privacy preserving feature of a given network. Indeed, when a node is neighborhood non-observable, it means that the data of the other nodes (excluding those of its neighborhood) cannot be retrieve from such a node. Therefore security efforts in order to preserve privacy of the entire network must be focused on nodes that are neighborhood-observable.

6.2.4. *Average and parametric consensus*
Participants: A. Kibangou [Contact person], F. Morbidi.

We have studied average consensus in wireless sensor networks with aim of providing a way to reach consensus in a finite number of steps [17]. In particular, we investigate the design of consensus protocols when, for security reasons for instance, the underlying graph is constrained to be strongly regular or distance regular. The proposed design method is based on parameters of the intersection array characterizing the underlying graph. With this protocol, at execution time, average consensus is achieved in a number of steps equal to the diameter of the graph, i.e. the smallest possible number of steps to achieve consensus. We have extended the parametric consensus protocol recently introduced by F. Morbidi, to more realistic agents modeled as double integrators and interacting over an undirected communication network. The stability properties of the new protocol in terms of the real parameter "s" are studied for some relevant graph topologies, and the connection with the notion of bipartite consensus is highlighted. The theory is illustrated with the help of two worked examples, dealing with the coordination of a team of quadrotor UAVs and with cooperative temperature measurement in an indoor environment [32].

6.3. Collaborative and distributed algorithms

6.3.1. *Distributed computation methods for large-scale multidimensional data*
Participants: A. Kibangou [Contact person], T.-M. D. Tran, A. de Almeida [UFC Brazil].

From Internet to large research infrastructures, the volume of data generated by our societies is continuously increasing. A deluge faced by the producers of these data as well as their users. The big data issue is a significant scientific challenge that requires deep investigations in both engineering and fundamental science. Low-rank matrix factorization has received a particular attention in recent years, since it is fundamental to a variety of mining tasks that are increasingly being applied to massive datasets. In large applications, matrix factorizations can involve matrices with billions of entries. At this massive scale, distributed algorithms for matrix factorization are essential to achieve reasonable performance [43]. However, in many disciplines, data inherently has more than two axes of variation and can be arranged as tensors (i.e. multi-way arrays). Computing tensor decompositions of multi-way datasets is particularly useful to extract hidden patterns and structure in data analytics problems. Specifically, CPD (Canonical Polyadic Decomposition) also known as PARAFAC (Parallel factor analysis) is an extension of a low rank matrix decomposition to tensors. In [26], we have introduced a fully distributed method to compute the CPD of a large-scale data tensor across a network of machines with limited computation resources. The proposed approach is based on collaboration between the machines in the network across the three modes of the data tensor. Such a multi-modal collaboration allows an essentially unique reconstruction of the factor matrices in an efficient way. We provide an analysis

of the computation and communication cost of the proposed scheme and address the problem of minimizing communication costs while maximizing the use of available computation resources.

6.3.2. *Collaborative source seeking*
Participants: C. Canudas de Wit [Contact person], R. Fabbiano, F. Garin, Y. Gaudfrin, J. Dumon.

The problem of source localization consists in finding, with one or several agents possibly cooperating with each other, the point or the spatial region from which a quantity of interest is being emitted. Source-seeking agents can be fixed sensors, that collect and exchange some information about the signal field and try to identify the position of the source (or the smallest region in which it is included), or moving devices equipped with one or more sensors, that physically reach the source in an individual or cooperative way. This research area is attracting a rapidly increasing interest, in particular in applications where the agents have limited or no position information and GPS navigation is not available, as in underwater navigation or in cave exploration: for instance, source localization is relevant to many applications of vapor emitting sources such as explosive detection, drug detection, sensing leakage or hazardous chemicals, pollution sensing and environmental studies. Other fields of interest are sound source localization, heat source localization and vent sources in underwater field. Techniques present in literature either are based on a specific knowledge of the solution of the diffusion process, or make use of an extremum-seeking approach, exciting the system with a periodic signal so as to explore the field and collect enough information to reconstruct the gradient of the quantity of interest. Our approach [13] lies in the computation of derivatives (potentially of any order) from Poisson integrals that, for isotropic diffusive source in steady-state, whose solution satisfies the Laplace equation, allows for a gradient search with a small computation load (derivatives are computed by integrals) and without requiring any knowledge of the closed-form solution, avoiding in the same time extremum-seeking oscillations; this has the additional advantage of an intrinsic high-frequency filtering, that makes the method robust to measurement noise. We also propose a distributed version of this algorithm [28], where agents communicate in order to reconstruct gradient information from local pointwise measurements, and a control law combines the two objectives of formation control (to have a circular formation, so that measurements are taken around circle) and gradient ascent (so as to move towards the source); differently from previous literature [41], the moving agents do not need to know their absolute position, but only relative bearing angle of their neighbours.

6.4. Sensor networks: estimation and data fusion

6.4.1. *Data fusion approaches for motion capture by inertial and magnetic sensors*
Participants: H. Fourati [Contact person], A. Makni, A. Kibangou.

The problem of rigid body attitude estimation under external acceleration from a small inertial/magnetic sensor module containing a triaxial gyroscope, accelerometer, and magnetometer is considered [15]. We are focused on two main challenges. The first one concerns the attitude estimation during dynamic conditions, in which external acceleration occurs [30]. Although external acceleration is one of the main source of loss of performance in attitude estimation methods, this problem has not been sufficiently addressed in the literature. A quaternion based adaptive Kalman filter (q-AKF) compensating external acceleration from the residual in the accelerometer is designed. At each step, the covariance matrix associated with the external acceleration is estimated to adaptively tune the filter gain. The second challenge deals with the energy consumption issue of gyroscope for a long-term battery life of Inertial Measurement Units (IMUs). We study the way to reduce the gyro measurement acquisition by switching on/off the sensor while maintaining acceptable attitude estimation. A smart detection approach is proposed to decide whether the body is in dynamic or static motion. The efficiency of the q-AKF is investigated through numerical simulations and experimental tests, under external acceleration and parsimonious use of gyroscope. This work is described in a submitted in IEEE/ASME Transactions on Mechatronics.

6.4.2. *Pedestrian dead-reckoning navigation*
Participant: H. Fourati [Contact person].

We propose a foot-mounted Zero Velocity Update (ZVU) aided Inertial Measurement Unit (IMU) filtering algorithm for pedestrian tracking in indoor environment. The algorithm outputs are the foot kinematic parameters, which include foot orientation, position, velocity, acceleration, and gait phase. The foot motion filtering algorithm incorporates methods for orientation estimation, gait detection, and position estimation. A novel Complementary Filter (CF) is introduced to better pre-process the sensor data from a foot-mounted IMU containing tri-axial angular rate sensors, accelerometers, and magnetometers and to estimate the foot orientation without resorting to GPS data. A gait detection is accomplished using a simple states detector that transitions between states based on acceleration measurements. Once foot orientation is computed, position estimates are obtained by using integrating acceleration and velocity data, which has been corrected at step stance phase for drift using an implemented ZVU algorithm, leading to a position accuracy improvement. We illustrate our findings experimentally by using of a commercial IMU during regular human walking trial in a typical public building. Experiment results show that the positioning approach achieves approximately a position accuracy less than 1 m and improves the performance regarding a previous work of literature [14].

6.4.3. *Sensor placement of unreliable sensors*

Participants: F. Garin [Contact person], P. Frasca [U. Twente], B. Gerencsér [U. Catholique de Louvain], J. Hendrickx [U. Louvain-la-neuve].

We consider problems in which sensors have to be deployed in a given environment in such a way to provide good coverage of it. It is clear that sensor failures may deteriorate the performance of the resulting sensor network. Then, it is also natural to ask if taking into account such uncertainties changes the coverage optimization problem and leads to a different optimal solution. For simplicity, we start considering a one-dimensional problem, where sensors are to be placed on a line in such a way to optimize the disk-coverage cost. The optimal solution for reliable sensors is simply an equally-spaced configuration of the sensors. If we allow that the sensors may fail to take or communicate their measurements, this solution may instead not be optimal. In our work, we assume that sensor can fail, independently and with a same failure probability, and we aim to minimize, in expectation, the largest distance between a point in the environment and an active sensor. Our first result states that the problem at hand is equivalent to a linear program, albeit with a number of variables growing exponentially with the number of sensors. This fact allows for a computational solution that is tractable if the number of sensors is not large. Secondly, we show that for large number of sensors n, the cost of the equispaced placement is asymptotically optimal, i.e., the ratio between its cost and the optimal cost tends to 1 when n grows. By contrast, we show in that a random sensor placement has an expected cost which is larger. This work has been presented at MTNS conference [35] and is described in a submitted journal paper (see http://arxiv.org/abs/1404.7711).

6.5. Control design and co-design

6.5.1. *Energy-aware networked control*

Participants: C. Canudas de Wit [Contact person], F. Garin, N. Cardoso de Castro, D. Quevedo [U. Newcastle].

We have considered an event-based approach to energy-efficient management of the radio chip in the sensor node of a wireless networked control system. Indeed the radio is the main energy consumer, and intermittent data transmission allows one to reduce the use of the radio. While the existing literature in the control community on event-based control only addresses policies using two radio modes (transmitting/sleep), our work follows some considerations on the radio chip modes well-known in the communication networks literature, and introduces various radio-modes: different 'idle' non-transmitting modes, where only part of the radio chip is switched off (thus consuming more energy than 'sleep', but allowing for faster transition to transmission), and various transmitting modes, with different power levels. We propose an event-based radio-mode switching policy, which allows to perform a trade-off between energy saving and performance of the control application; to this end, a switched model describes the system, taking into account control and communication. The optimal switching policy is computed using dynamic programming, considering a cost either over an infinite time-horizon (see [36]) or over a finite receding horizon (joint work with D. Quevedo, Univ. Newcastle, Australia, described in a paper in preparation).

6.5.2. *Adaptive control strategy based reference model for spacecraft motion trajectory*

Participants: H. Fourati [Contact person], Z. Samigulina [Kazakh National Technical University], O. Shiryayeva [Institute of Informatics and Control Problems].

In aerospace field, the economic realization of a spacecraft is one of the main objectives which should be accomplished by conceiving the optimal propulsion system and the best control algorithms. Our work focuses on the development of a viable Adaptive Control Approach (ACA) for Spacecraft Motion Trajectory (SMT), see [19]. The proposed strategy involves the nonlinear mathematical model of SMT expressed in the central field, which is linearized by the Taylor expansion, and the second Lyapunov method to offer a high rate and unfailing performance in the functioning. The adaptive control system is composed of the cascade of adaptation loop and feedback control loop. When the spacecraft deviates from its reference trajectory model, the ACA acts on the control system to correct this deviation and follow the optimal reference trajectory. Therefore, when the states of the adjustable model are different from its reference values, then the error signal is provided as an input to the adaptation law, which contains the adaptation algorithm. The output will be the state variable feedback control matrix which will be used to calculate the new control law vector. The efficiencies of the linearization procedure and the control approach are theoretically investigated through some realistic simulations and tests under MATLAB. The steady state errors of control between the reference model and the adjustable model of SMT converge to zero. This work is described in [20].

6.5.3. *Control design for hydro-electric power-plants*

Participants: C. Canudas de Wit [Contact person], S. Gerwig, F. Garin, B. Sari [Alstom].

We have initiated a collaboration with Alstom on collaborative and resilient control of hydro-electric power-plants, with the CIFRE PhD thesis of Simon Gerwig. The first goal of this research is to improve performance of a hydro-electric power-plant outside its design operation conditions, by adaptive cancellation of oscillations that occur in such an operation range. Indeed, current operation of power-plants often requires to operate on a variety of conditions, often different from the ones initially considered when designing the plant. At off-design operation pressure, the hydraulic turbine exhibits a vortex rope below the runner. This vortex generates pressure fluctuations after the turbine and can excite the hydraulic pipes. Indeed the water is compressible and the pipe walls elastic, so the system can oscillate. The goal is to damp these pressure oscillations as they create vibrations in the system and can lead to damages. Our first contribution has been to model the effect of the vortex rope on the hydraulic system as an external perturbation source acting on pipes. The pipes themselves are described with equations taking into account water compressibility and pipe-wall elasticity. The resulting model is nonlinear with hyperbolic functions in the equations (analogous to high-frequency transmission lines), from which we obtain a suitably linearized model.

6.5.4. *Controller for switched linear systems*

Participants: H. Fourati [Contact person], Djamel. E. C. Belkhiat [U. Setif], D. Jabri [U. Setif].

We designed a robust output feedback tracking controller for a class of Switched Linear Systems (SLS) subject to external disturbances [23]. The proposed synthesis approach, based on a descriptor redundancy formulation, allows to avoid of the crossing terms appearance between the switched Proportional-Derivative (PD) controller's and the switched system's matrices. Using the multiple Lyapunov functional methods, a robust output feedback tracking performance has been formulated in terms of set of Linear Matrix Inequality (LMI). The effectiveness of the proposed synthesis procedure has been illustrated by a numerical example [24].

6.6. Transportation networks and vehicular systems

6.6.1. *Traffic estimation and prediction*

Participants: C. Canudas de Wit [Contact person], A. Kibangou, L. Leon Ojeda, F. Morbidi.

In the PhD thesis of Leon Ojeda, we have been concerned with the design of a methodology for the real-time multi-step ahead travel time forecasting using flow and speed measurements from an instrumented freeway. Two main methodologies have been considered. The first one, a signal-based, uses only speed measurements collected from the freeway, where a mean speed is assumed between two consecutive collection points. The travel time is forecasted using a noise Adaptive Kalman Filter (AKF) approach. The process noise statistics are computed using an online unbiased estimator, while the observations and their noise statistics are computed using the clustered historical traffic data. Forecasting problems are reformulated as filtering ones through the use of pseudo-observations built from historical data. The second one, a model-based, uses mainly traffic flow measurements. Its main appealing is the use of a mathematical model in order to reconstruct the internal state (density) in small road portions, and consequently exploits the relation between density and speed to forecast the travel time. The methodology uses only boundary conditions as inputs to a switched Luenberger state observer, based on the "Cell Transmission Model" (CTM), to estimate the road initial states. The boundary conditions are then forecasted using the AKF developed above. Consequently, the CTM model is run using the initial conditions and the forecasted boundaries in order to obtain the future evolution of densities, speeds, and finally travel time. The added innovation in this approach is the space discretization achieved: indeed, portions of the road, called "cells", can be chosen as small as desired and thus allow obtaining a finer tracking of speed variations. The developed methodologies were assessed using the city-lab GTL [31]. Features and activities of this platform are described in [39].

6.6.2. *Traffic control*

Participants: C. Canudas de Wit [Contact person], D. Pisarski.

The work was mainly focused on the final design of a distributed controller and its implementation to the model of the south ring of Grenoble in the context of the project Hycon2. For the sake of the controller design, a distributed optimal control method for balancing of freeway traffic density was studied. The optimization was performed in a distributed manner by utilizing the controllability properties of the freeway network represented by the Cell Transmission Model. By using these properties, the subsystems to be controlled by local ramp meters were identified. The optimization problem was then formulated as a non-cooperative Nash game. The game was solved by decomposing it into a set of two-players hierarchical and competitive games. The process of optimization employed the communication channels matching the switching structure of system interconnectivity. By defining the internal model for the boundary flows, local optimal control problems were efficiently solved by utilizing the method of Linear Quadratic Regulator. The developed control strategy was tested via numerical simulations on the macroscopic model in two scenarios for uniformly congested and transient traffic. The controller was also validated through a microscopic simulations with the use of Aimsun software. The controller was implemented through Matlab under which a relevant program simulating distributed architecture was designed. The controller was then plugged to the Aimsun micro-simulator. The simulated scenario was based on real traffic data collected from the south ring of Grenoble. Were examined both, the balancing metric (optimized) and a set of standard traffic metrics (not optimized). The results showed that the balancing has a positive impact on the traffic flow, in particular, by smoothing the vehicle dynamics, it can potentially increase the average velocity (and thus, reduce the travelling time) and reduce the fuel consumption (and related emissions). The proposed modular architecture enabled to perform the optimization for long freeway sections in the real-time.

6.6.3. *Control of urban traffic networks*

Participants: C. Canudas de Wit [Contact person], F. Garin, P. Grandinetti.

This work deals with efficient operation of urban traffic networks, by controlling traffic lights. A first contribution has been to devise a model for urban networks, based on the Cell-Transmission-Model adapted to signalized intersections, and then simplified with an average-based approximation. Then, based on this model, a control law has been designed, where the duty cycle of each traffic light is optimized in real time, globally considering the whole network. We have chosen a simple one-step-ahead optimization, which can be formulated as a linear program, thus resulting in simple and fast optimization. This work is part of the PhD thesis of Pietro Grandinetti.

6.6.4. *Stability of Monotone Dynamical Flow Networks*
Participants: E. Lovisari [Contact person], G. Como [U. Lund], K. Savla [U. of Southern California].

The stability properties of monotone dynamical flow networks are studied [22]. Demand and supply functions relate states and flows of the network, and the dynamics at junctions are subject to fixed turning rates. Our main result consists in the characterization of a stability region such that: If the inflow vector in the network lies strictly inside the stability region and a certain graph theoretical condition is satisfied, then a globally asymptotically stable equilibrium exists. In contrast, if the inflow vector lies strictly outside the region, then every trajectory grows unbounded in time. As a special case, our framework allows for the stability analysis of the Cell Transmission Model on networks with arbitrary topologies. These results extend and unify previous work by Gomes et al. on stability of the Cell Transmission Model on a line topology as well as that by the authors on throughput optimality in monotone dynamical flow networks.

6.6.5. *Control of communicating vehicles in urban environment*
Participants: C. Canudas de Wit [Contact person], G. de Nunzio.

The stability properties analysis of the Variable Length Model (proposed by Prof. Canudas de Wit in 2011), adapted to the urban environment was studied. It has been found that the canonical definition of Lyapunov stability for the equilibrium points does not hold for the system under analysis. A different approach for the analysis of the stability properties of the system has been introduced. Furthermore, an energy map of the equilibrium points has been obtained. Namely, a cost was associated with each feasible equilibrium point of the system, thus obtaining an assessment of the efficiency of any operation point of the system. A Variable-Speed-Limits tracking controller of the desired operation point (i.e. equilibrium) has been also devised, in order to simulate the response of the driver to the energy-efficient speed advisory. This work was submitted and accepted at the IEEE Conference on Decision and Control 2014, with the title "Urban Traffic Eco-Driving: Speed Advisory Tracking". A previous work on the steady-state analysis of the Variable Length Model in urban environment was carried out in [27]. The effort has been put also on the validation of the macroscopic model (i.e. the Variable Length Model), used for traffic evolution prediction and control synthesis. The validation procedure was run with a microscopic traffic simulator, and aims at proving that the evolution of the state of the mathematical model replicates accurately the true evolution of the traffic conditions. In particular, an important variable modeled by the system is the length of the congested area of the road section under analysis, which may be thought of as the queue length. It has been shown that the macroscopic model is able to depict the evolution in time of the queue length, with only a small error with respect to the real congestion simulated by the highly-detailed microscopic simulator. Furthermore, the validation process aims, not only at confirming the reliability of the dynamical model, but also the accuracy of the energy consumption model and the other macroscopic traffic performance metrics that have been defined in order to formulate the optimization problem. Within the COMFORT project exchange program, the work on bandwidth maximization on signalized arterials by introducing VSL as an additional degree of freedom, and by considering the energetic aspects of the problem was expanded. The canonical bandwidth maximization problem is defined as the maximization of the time interval that the vehicles can use to drive though a sequence of signalized intersections without stopping; this is achieved solely by offset control. The extension of this framework aims at showing that the additional degree of freedom (i.e. variable speed limits) improves in every case the bandwidth. A further simulation campaign in a microscopic simulator shows the benefits of the theoretical bandwidth maximization on the standard traffic performance metrics. In particular, fluidity of traffic and lower number of stops result to be highly beneficial in terms of energy consumption, without losing much in terms of traveling time.

OPALE Project-Team

6. New Results

6.1. Highlights of the Year

Paola Goatin was awarded the *"Prix Inria - Académie des sciences du jeune chercheur"*.

6.2. Mathematical analysis and control of macroscopic traffic flow models

6.2.1. *Vehicular traffic*

Participants: Enrico Bertino, Guillaume Costeseque, Maria Laura Delle Monache, Paola Goatin, Sheila Scialanga, Alexandre Bayen [UC Berkeley, CA, USA], Sebastien Blandin [IBM Research Collaboratory, Singapore], Christophe Chalons [LJLL, UP7].

In collaboration with UC Berkeley, and as part of the Associated Team ORESTE activity (see http://www-sop.inria.fr/members/Paola.Goatin/ORESTE/index.html), we have considered the System Optimal Dynamic Traffic Assignment problem with Partial Control (SO-DTA-PC) for general road networks with horizontal queuing. The goal of which is to optimally control any subset of the networks agents to minimize the total congestion of all agents in the network. We adopt a flow dynamics model that is a Godunov discretization of the Lighthill-Williams-Richards (LWR) partial differential equation with a triangular flux function and a corresponding multi-commodity junction solver. Full Lagrangian paths are assumed to be known for the controllable agents, while we only assume knowledge of the aggregate turn ratios for the non-controllable (selfish) agents. We solve the resulting finite horizon non-linear optimal control problem using the discrete adjoint method, see [75].

As part of our TRAM3 activity and in collaboration with C. Chalons (UVSQ), we designed a new finite volume conservative algorithm to track the trajectory of a bus in the surrounding traffic using a locally non-uniform moving mesh, see [70].

In collaboration with S. Blandin (IBM), we proved the existence and stability of entropy weak solutions of a scalar conservation law with non-local flux arising in traffic flow modeling. The result is obtained providing accurate L^∞, BV and L^1 estimates for the sequence of approximate solutions constructed by an adapted Lax-Friedrichs scheme.

In collaboration with the University of Mannheim and in the framework of the PHC Procope project "Transport Networks Modeling and Analysis", we studied how to manage variable speed limits combined with coordinated ramp metering within the framework of the LWR network model. Following a "first discretize then optimize" approach, we derived the first order optimality system and explained the switch of speeds at certain fixed points in time and the boundary control for the ramp metering as well. Sequential quadratic programming methods are used to solve the control problem numerically. For application purposes, we present experimental setups where variable speed limits are used as a traffic guidance system to avoid traffic jams on highway interchanges and on-ramps, see [71].

Finally, E. Bertino internship was devoted to uncertainty quantification in macroscopic traffic flow models.

6.2.2. *Crowd motion*

Participants: Aekta Aggarwal, Régis Duvigneau, Paola Goatin, Matthias Mimault, Rinaldo M. Colombo [Brescia University, Italy].

A. Aggarwal postdoc is devoted to the analytical and numerical study of systems of conservation laws with non-local fluxes in several space dimensions. In collaboration with R.M. Colombo, we presented a Lax-Friedrichs type algorithm to numerically integrate this class of systems. The convergence of the approximate solutions was proved, also providing the existence of solution in a slightly more general setting than in other results in the current literature. An application to a crowd dynamics model is considered. This numerical algorithm is then used to test the conjecture that as the convolution kernels converge to a Dirac δ, the nonlocal problem converges to its non-nonlocal analogue.

M. Mimault is working on scalar conservation laws with non-local flow in two space dimensions. These equations are meant to model crowd motion, where the movement direction of each pedestrian depends on a weighted mean of the crowd density around him. In particular, M. Mimault is implementing a finite volume numerical scheme which will be used for flow optimization purposes.

The above researches were partially funded by the ERC Starting Grant "TRAM3 - Traffic management by macroscopic models".

6.3. Optimum design and control in fluid dynamics and its couplings

In computational sciences for physics and engineering, Computational Fluid Dynamics (CFD) are playing one of the major roles in the scientific community to foster innovative developments of numerical methodologies. Very naturally, our expertise in compressible CFD has led us to give our research on numerical strategies for optimum design a particular, but not exclusive focus on fluids.

The framework of our research aims to contribute to numerical strategies for PDE-constrained multi-objective optimization, with a particular emphasis on CPU-demanding computational applications in which the different criteria to be minimized (or reduced) originate from different physical disciplines that share the same set of design variables. These disciplines are often fluids, as a primary focus, coupled with some other disciplines, such as structural mechanics.

Our approach to *competitive optimization* is focused on the two-discipline problem. It is based on a particular construction of *Nash games*, relying on a *split of territory* in the assignment of individual strategies. A methodology has been proposed for the treatment of two-discipline optimization problems in which one discipline, the primary discipline, is preponderant, or fragile. Then, it is recommended to identify, in a first step, the optimum of this discipline alone using the whole set of design variables. Then, an orthogonal basis is constructed based on the evaluation at convergence of the Hessian matrix of the primary criterion and constraint gradients. This basis is used to split the working design space into two supplementary subspaces to be assigned, in a second step, to two virtual players in competition in an adapted Nash game, devised to reduce a secondary criterion while causing the least degradation to the first. The formulation has been proved to potentially provide a set of Nash equilibrium solutions originating from the original single-discipline optimum point by smooth continuation, thus introducing competition gradually [65] (see also subsectionsubsect:helico).

Our approach to *cooperative optimization*, in theory, is not limited in number of objective functions. It is based on a result of convex analysis established for a general unconstrained multiobjective problem in which all the gradients are assumed to be known. The theorem [66] states that in the convex hull of the gradients, there exists a unique vector of minimal norm, ω; if it is nonzero, the vector ω is a descent direction common to all criteria; otherwise, the current design point is Pareto-stationary. This result led us to generalize the classical steepest-descent algorithm by using the vector ω as search direction. We refer to the new algorithm as the multiple-gradient descent algorithm (MGDA). The MGDA yields to a Pareto-stationary point, and in practice actual Pareto-optimality is verified a posteriori.

The two approaches have been combined to explore the Pareto front segment-wise as illustrated on Figure 2 .

6.3.1. *Multiple-Gradient Descent Algorithm (MGDA)*
Participants: Jean-Antoine Désidéri, Régis Duvigneau, Camilla Fiorini, Matteo Giacomini, Abderrahmane Habbal, Adrien Zerbinati.
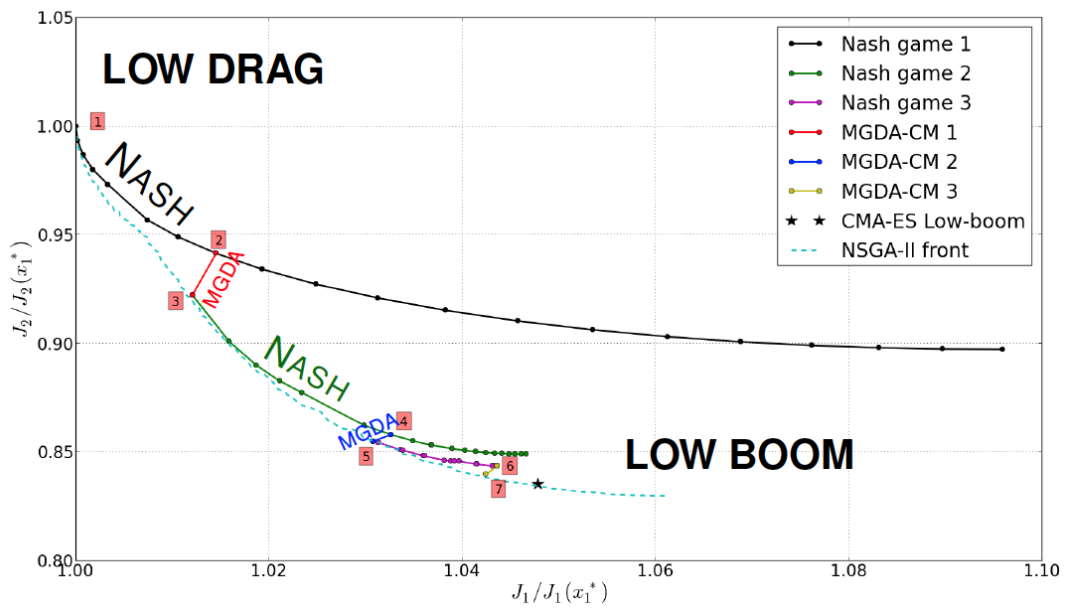
*Figure 2. **Two-discipline optimization of a generic geometry of a supersonic aircraft, for concurrent drag and sonic-boom reduction** (from A. Minelli's doctoral thesis, 2013). The wave drag is calculated by the ONERA elsA code in 3D finite-volume Eulerian flow mode over a 6M-node mesh and the sonic boom using a three-layer approach. The Nash-game paths have been devised by appropriate territory splitting in order to be tangent to the Pareto front, and they are interrupted whenever the Pareto-stationarity condition is judged excessively violated. The MGDA paths converge rapidly back to the front. The simulation demonstrates how the two algorithms complement each other and provide a potential for a piecewise description of the Pareto front, evaluated more economically than a stochastic algorithm operating on a large population.*

6.3.1.1. Theory and numerical experimentation of the MGDA construction

In multi-objective optimization, the knowledge of the Pareto set provides valuable information on the reachable optimal performance. A number of evolutionary strategies (PAES, NSGA-II, etc), have been proposed in the literature and proved to be successful to identify the Pareto set. However, these derivative-free algorithms are very demanding in terms of computational time. Today, in many areas of computational sciences, codes are developed that include the calculation of the gradient, cautiously validated and calibrated [66].

The notion of Pareto-stationarity, originally established to be a necessary condition of optimality in differentiable multi-objective optimization of unconstrained problems, has been extended to problems subject to equality constraints. On this basis, we were able to establish that by augmenting, in a classical manner, the objective-functions of a penalty term equal to the square of the constraint violation, and applying the standard MGDA to it, would result in converged solutions that are Pareto-stationary in the extended sense. Numerical experimentation on this is on-going.

6.3.1.2. Meta-model-assisted CFD optimization by MGDA

Using MGDA in a multi objective optimization problem requires the evaluation of a large number of points with regard to criteria, and their gradients. In the particular case of a CFD problems, each point evaluation is very costly since it involves a flow computation, possibly the solution of an adjoint-equation. To alleviate this difficulty, we have proposed to construct meta-models of the functionals of interest (lift, drag, etc) and to calculate approximate gradients by local finite differences. These meta-models are updated throughout the convergence process to the evaluation of the new design points by the high-fidelity model, here the 3D compressible Euler equations.

This variant of MGDA has been tested successfully over several aerodynamic shape optimization problems : lift concurrently with drag optimization for transonic aircraft; drag (under lift constraint) concurrently with sonic boom reduction for 3D supersonic configuration (at ONERA); drag (under lift constraint) concurrently with mass reduction for transport aircraft (at ONERA) [56].

6.3.1.3. Exact shape gradients

MGDA has successfully been tested over a two-objective optimization problem governed by two-dimensional elasticity. The deformation of a plate is calculated using an isogeometric approximation and compliance derived from it. The exact parametric shape gradient is calculated, yielding the gradient of the objective function in two antagonistic situations differing by the loading. Pareto-fronts are thus identified [68].

6.3.1.4. Optimization of an unsteady system using a multiobjective formulation

An approach has been developed to solve optimization problems in which the functional that has to be minimized is time dependent. In the literature, the most common approach when dealing with unsteady problems, is to consider a time-average criterion. However, this approach is limited since the dynamical nature of the state is neglected. Our alternative consists in building a set of cost functionals by evaluating a single criterion at different sampling times. In this way, the optimization of the unsteady system is formulated as a multi-objective optimization problem, solved using an appropriate descent algorithm (MGDA). Moreover, we also consider a hybrid approach in which the set of cost functionals is built by doing a time-average operation over multiple intervals. These strategies have been illustrated and applied to a non-linear unsteady system governed by a one-dimensional convection-diffusion-reaction partial differential equation [67].

6.3.1.5. Perspectives

MGDA offers the possibility to handle in a rational way several objective-functions for which gradients are known or approximated concurrently. This potential opens methodological paths to several themes of interest in high-fidelity simulation-based optimization: optimization of complex systems whose performance is evaluated w.r.t. several criteria originating from different, coupled disciplines; optimization under uncertainties, by introducing sensitivities as additional objectives; optimization of time-dependent systems, such as optimization of flow-control devices that generate a periodic flow (see next subsection), by converting the problem into a multipoint problem by time-discretization of the time and parameter-dependent functional (as above); etc.

6.3.2. *Flow control*

Participants: Régis Duvigneau, Jérémie Labroquère, Emmanuel Guilmineau [Ecole Centrale de Nantes].

Shape optimization methods are not efficient to improve the performance of fluid systems, when the flow is characterized by a strong unsteadiness related to a massive detachment. This is typically the case for the flow around an automotive body or a wing in stall condition. To overcome this difficulty, flow control strategies are developed, that aim at manipulating vortex dynamics by introducing some active actuators, such as periodic blowing/suction jets. In this context, the choice of the control parameters (location, amplitude, frequency) is critical and not straightforward. Therefore, we develop a methodology to determine optimal control parameters by coupling the simulation of unsteady actuated flows with optimization algorithms. Two research axes have been considered :

- the resolution of the unsteady sensitivity equations derived from the state equations, to exhibit the dependency of the flow dynamics with respect to the control and apply an unsteady gradient-based approach[67] ;
- the optimization of control parameters using a statistical metamodel-based strategy [39].

In this perspective, unsteady Reynolds Averaged Navier-Stokes equations are solved, with some turbulence closures. Different models for synthetic jet have been implemented to simulate the actuation, and then validated for different turbulence closures .

Specific developments have be carried out in the metamodel-based optimizer to include a noise term into Gaussian Process model, which is used to filter errors arising from unsteady simulations. A systematic assessment of modeling and numerical errors has been archived for a backward facing step test-case, with the objective of controlling the re-attachment point location[46], [58].

This activity is conducted in collaboration with the CFD team of Ecole Centrale de Nantes.

6.3.3. *Adjoint-based mesh quality control*

Participants: Jean-Antoine Desideri, Maxime Nguyen-Dinh [ONERA doctoral student], Jacques Peter [Research Engineer, ONERA/DSNA], Renaud Sauvage [Airbus France], Mathieu Meaux [EADS IW].

In his doctoral thesis [29], Nguyen Dinh has studied mesh adaptation methods based on the total derivatives of aerodynamic outputs with respect to mesh coordinates by the discrete adjoint method. Firstly, mesh adaptation methods have been devised for Eulerian flows. Zones to be refined are detected using a sensor based on the total derivative, and numerical experiments confirmed the adequacy of the approach. Secondly, the method was extended to the Reynolds-averaged Navier equations (RANS) and thirdly demonstrated for 3D industrial configurations [53].

6.3.4. *Helicopter rotor blade optimization in both situations of hovering and forward flight*

Participants: David Alfano [Airbus Helicopter], Michel Costes [Research Engineer, ONERA/DAAP], Jean-Antoine Désideri, Arnaud Le Pape [Research Engineer, ONERA/DAAP], Enric Roca Leon.

E. Roca Leon has conducted a CIFRE thesis at ONERA DAAP supported by Airbus Helicopter (Marignane) [34]. This thesis follows the doctoral thesis of A. Dumont in which the adjoint-equation approach was used to optimize a rotor blade in hovering flight. The goal of this new thesis is to solve a two-objective optimization problem in which the hovering-flight criterion is considered preponderant, but a new criterion that takes into account the forward-flight situation is also introduced, concurrently. The second criterion is the power necessary to maintain the forward motion. The first phase of thesis work has been devoted to the set up of a hierarchy of models from low to high fidelity, in order to calibrate appropriate functional criteria. Then, actual two-objective optimizations are conducted via our Nash game approach to competitive optimization with territory splitting based on reduced Hessian diagonalization. Successful optimization has been realized involving 16 geometrical parameters to reduce the power in forward motion while maintaining sub-optimality of the drag in hover [55] [64] [65].

6.4. Isogeometric analysis and design

Participants: Régis Duvigneau, Asma Gdhami, Bernard Mourrain [Galaad Project-Team], Bernd Simeon [Tech. Univ. of Kaiserslautern].

Design optimization stands at the crossroad of different scientific fields (and related software): Computer-Aided Design (CAD), Computational Fluid Dynamics (CFD) or Computational Structural Dynamics (CSM), parametric optimization. However, these different fields are usually not based on the same geometrical representations. CAD software relies on Splines or NURBS representations, CFD and CSM software uses grid-based geometric descriptions (structured or unstructured), optimization algorithms handle specific shape parameters. Therefore, in conventional approaches, several information transfers occur during the design phase, yielding approximations that can significantly deteriorate the overall efficiency of the design optimization procedure. Moreover, software coupling is often cumbersome in this context.

The isogeometric approach proposes to definitely overcome this difficulty by using CAD standards as a unique representation for all disciplines. The isogeometric analysis consists in developing methods that use NURBS representations for geometric modeling, computational domain description and solution basis functions. Using such a unique data structure allows to compute the solution on the exact geometry (not a discretized geometry), obtain a more accurate solution (high-order approximation), reduce spurious numerical sources of noise that deteriorate convergence, avoid data transfers between the software. Moreover, NURBS representations are naturally hierarchical and allows to define multi-level algorithms for solvers as well as optimizers.

In this context, a collaborative work has also been carried out with the Technical University of Kaiserslautern, concerning the computation of shape gradients for linear elasticity problems[42], [68]. Moreover, the doctoral thesis of Asma Gdhami, in collaboration with ENIT in Tunisia, has started and concerns the development of isogeometric schemes for hyperbolic systems.

6.5. Optimum design in structural mechanics

6.5.1. *Shape Optimization in Multidisciplinary Non-Linear Mechanics*

Participants: Aalae Benki, Jean-Antoine Désidéri, Abderrahmane Habbal, Gael Mathis [ArcelorMittal, CRAA].

In collaboration with the ArcelorMittal's Center for Research in Automotive and Applications (CRAA), we study the multidisciplinary shape and parameter design of highly non linear mechanical 2D and 3D structures. We have developed methods adapted to the approximation of Pareto Fronts such as Normal Boundary Intersection NBI and Normalized Normal Constraint Method NNCM. Due to the time consuming cost evaluation, the use of cheap to evaluate surrogate models is mandatory. We have studied the consistency of the approach NBI or NNCM plus surrogates, which turned out to be successful for a broad panel of standard mathematical benchmarks. The coupling is successfully applied to a small scale industrial case, namely the shape optimization of a can bottom vis à vis dome reversal pressure and dome growth criteria. We have then defined a Nash game between criteria where the latter are approximated by the RBF metamodels. First, we validated the computation of a Nash equilibrium for mathematical functions, then we computed Nash equilibria for the small scale industrial case of the shape optimization of the can bottom.

Then, we considered the 3D problem of an automotive twist beam. In this 3D case, we aim to Pareto-optimal shapes for two objectives, the first being to minimize the Von-Mises strain to guarantee the formability of the twist beam, and the second being to maximize the stiffness. For solution with higher stiffness than the initial one, we could decrease the thickness to obtain a mass reduction with the same end-user properties.

We also introduced, to our knowledge for the first time in the structural optimization area, the notion of Kalai-Smorodinky equilibria which is aimed at the selection of equilibria among Pareto-optimal solutions. We applied this notion of equilibria to both industrial cases, and compared the results to Nash equilibria.

6.5.2. *Optimization of Addendum Surfaces in Stamping*

Participants: Fatima Zahra Oujebbour, Rachid Ellaia, Abderrahmane Habbal, Ziheng Zhao.

Within the OASIS Consortium (ArcelorMittal, ErDF, Inria, UTC, EURODECISION, ESILV, NECS, Delta-CAD, SCILAB-DIGITEO), Opale Project leads the Optimization task. Our aim is to develop decentralized decision-making algorithms dedicated to find efficient solutions (Pareto optimal) in a complex multidisciplinary framework (forming, stamping, welding non-linear processes, spring-back, vibration, in-function linear processes, crash and fatigue non linear and non differentiable processes) for several (between three and five) criteria. An important difficulty when trying to identify the Pareto Front, even when using adapted methods such the Normal Boundary Intersection, is that the criteria involved (thanks to the high nonlinearity in the mechanical models) exhibit many local optima. So one must use global optimization methods. We have studied the hybrid approach Simulated Annealing with Simultaneous Perturbation SASP for a suite of mathematical test-cases. To envisage the application of our method to the complex CPU time consuming stamping process, we lead an intermediate phase dedicated to the validation of the SASP method for the minimization of the spring-back that follows the stamping of a metal sheet, the design variable being the process parameters (two then four parameters). Then, we considered the more complex shape design of the initial blank. The initial blank design is a critical step in stamping design procedure, therefore it should be optimally designed. Our aim is to find the optimal initial blank shape that avoids or at least minimizes the springback and failure flaws. For this study, the geometry of the blank contour is described by parametric spline curves. Seven control points (P1,...,P7) are used to define the spline curves in order to have a wide variety of geometries. The exact computational evaluation of our criteria, springback and failure, is very expensive (the FE model request around 45 min to predict these two criteria) and the design space is of quite high dimension. Therefore, we considered the recourse to the sparse grid interpolation. Optimization process based on sparse grid interpolation is an optimal alternative in which criteria can be approximated with a suitable interpolation formula that needs significantly less points than the full grid. the obtained metamodel using sparse grid interpolation needs less than 1s to predict springback and failure on the same computation machine. To find the optimal initial blank shape, it was decided to perform the optimization process using the obtained metamodel. The construction of the sparse grid interpolant was based on the Chebyshev Gauss-Lobatto grid type and using the polynomial basis functions. This technique achieves a good accuracy with a competitive number of grid points. The comparison of the obtained fronts shows that we can capture Pareto solutions by NBI and NNCM with fewer points than NSGAII which requires a large number of populations and several generations to obtain the Pareto front. [48] [49] [50]

6.6. Application of shape and topology design to biology and medicine

6.6.1. *Assessing the ability of the 2D Fisher-KPP equation to model cell-sheet wound closure*

Participants: Abderrahmane Habbal, Hélène Barelli [Univ. Nice Sophia Antipolis, CNRS, IPMC], Grégoire Malandain [Inria, EPI Morpheme].

We address in this joint collaboration the ability of the widely used Fisher-KPP equations to render some of the dynamical features of epithelial cell-sheets during wound closure.

Our approach is based on nonlinear parameter identification, in a two-dimensional setting, and using advanced 2D image processing of the video acquired sequences. As original contribution, we lead a detailed study of the profiles of the classically used cost functions, and we address the "wound constant speed" assumption, showing that it should be handled with care.

We study five MDCK cell monolayer assays in a reference, activated and inhibited migration conditions. Modulo the inherent variability of biological assays, we show that in the assay where migration is not exogeneously activated or inhibited, the wound velocity is constant. The Fisher-KPP equation is able to accurately predict, until the final closure of the wound, the evolution of the wound area, the mean velocity of the cell front, and the time at which the closure occurred. We also show that for activated as well as for inhibited migration assays, many of the cell-sheet dynamics cannot be well captured by the Fisher-KPP model. Original unexplored utilizations of the model such as wound assays classification based on the calibrated diffusion and proliferation rate parameters is ongoing. [47]

6.7. Distributed Systems

6.7.1. *High-Performance manipulation and storage of e-Science data*

Participants: Benoit Lange, Toan Nguyen.

The work carried in previous years on distributed High-Performance Computing for e-Science workflows has enlightened the need for appropriate tools and methods to manage petabyte and exabyte volumes of data. This has been the focus of the work carried by Benoit Lange during his Post-Doc position in 2014. It was dedicated to the definition and prototyping of a large-scale HPC platform to support the execution of application solvers, efficient storage and management of large-volumes of data produced by the simulation applications and the visualization of their results on high-end graphics workstations. This platform also includes analytics software to produce specific results corresponding to the user queries. This is based on the Hadoop ecosystem [59]. Is is central for the communication between the dedicated HPC nodes running the solvers and the visualization nodes interfacing the end-users. It includes high-speed storage with dedicated file systems on specific nodes, and long-term storage for reference data using magnetic juke-boxes that store petabytes of application data. This work is supported by an FP7 project in which Inria is responsible for the Data Management work-package (Call FP7-2013-ICT-11, Grant 619439, 2014-2016). The partners of the project, named VELaSSCo (Visualization for Extremely Large Scale Scientific Computing), are : CIMNE (SP, coordinator), JOTNE and SINTEF (No), ATOS (SP), Fraunhofer IGD (D) and the University of Edinburg (UK).

<p style="text-align:center">**BAMBOO Project-Team**</p>

6. New Results

6.1. Evolution of the genomes of endosymbionts in insects: the case of Hamiltonella defensa interacting with its various partners

Insect cells host many endosymbiotic bacteria, which are in general classified according to their importance for the host: "primary" symbionts are by definition mandatory and synthesize essential nutrients for the insects that feed on poor or unbalanced food sources, while "secondary" symbionts are optional and use mutualistic strategies and/or manipulation of reproduction to invade and persist within insect populations. *Hamiltonella defensa* is a secondary endosymbiont that established two distinct associations with phloemophagous insects. In aphids, it protects the host against parasitoid attacks. Its ability to infect many host tissues, notably the hemolymph, could promote its contact with parasitoid eggs. Despite this protective phenotype, the high costs associated with its presence within the host prevent its fixation in the population. In the whitefly *Bemisia tabaci* however, this symbiont is found only in cells specialised in hosting endosymbionts, the bacteriocytes. In these cells, it cohabits with other symbiotic species, such as the primary symbiont *Portiera aleyrodidarum*, a proximity that favors potential exchanges between the two symbionts. It is fixed in populations of *B. tabaci*, which suggests an important role for the consortium, probably nutritious.

We studied the specificities of each of these systems. First, in the bacteriocytes of *B. tabaci*, we identified a partitioning of the synthetic capacities of two endosymbionts, *H. defensa* and *P. aleyrodidarum*, in addition to a potential metabolic complementation between the symbionts and their host for the synthesis of essential amino acids. We proposed a key nutritive role for *H. defensa*, which would indicate a transition to a mandatory status in relation to the host and would explain its fixation in the population.

We also focused on the genomic evolution of the genus *Hamiltonella*, by comparing the strains infecting *B. tabaci* with a strain infecting the aphids. We highlighted the specialization of the symbionts to their hosts, and found that the genomes of the endosymbionts reflected their respective ecology. The aphid strain thus possesses many virulence factors and is associated with two partners, a bacteriophage and a recombination plasmid. These systems, inactive in the symbiont of *B. tabaci*, are directly related to the protection against and arms race with parasitoids. Conversely, the presumed avirulence of whitefly endosymbionts is consistent with their nutritional phenotype and a transition to a mandatory status to the host.

Finally, we studied the phenomenon of "accelerated mutation rate" in *H. defensa*, compared to its sister species *Regiella insecticola*, which is also a clade of protective endosymbionts of aphids. After excluding the assumption that the transition to the intracellular life occurred independently in the two lineages, we tried to establish a link between these differences in terms of evolvability in the endosymbionts and of their gene contents, particularly for genes involved in ecology and DNA repair. All the results obtained have provided insight into the evolution of the species *H. defensa*, since the last ancestor to the present species, by establishing a link between bacterial phenotype and genomic evolution.

The publications related to this area of research are either submitted or in preparation (to be submitted in the first months of the year).

6.2. Cardinium cBtQ1 providing insights into genome reduction, symbiont motility, and its settlement in Bemisia tabaci

Many insects harbor inherited bacterial endosymbionts. Although some of them are not strictly essential and are considered facultative, they can be a key to host survival under specific environmental conditions, such as parasitoid attacks, climate changes, or insecticide pressures. The whitefly *Bemisia tabaci* is at the top of the list of organisms inflicting agricultural damage and outbreaks, and changes in its distribution may be

associated to global warming. In partnership with the group of Andrès Moya at the ICBiBE (Institut Cavanilles de Biodiversitat i Biologia Evolutiva), the genome of *Cardinium* cBtQ1, a facultative bacterial endosymbiont of *B. tabaci*, was sequenced and analysed [23].

6.3. Mitochondrial respiration and genomic analysis provide insight into the influence of the symbiotic bacterium on host trypanosomatid oxygen consumption

Certain trypanosomatids, such as *Angomonas deanei*, co-evolve with an endosymbiotic bacterium in a mutualistic relationship that is characterised by intense metabolic exchanges. We were able to show that the symbionts were able to respire for up to 4 h after isolation from the host. Moreover, our work suggests that the symbiont influences the mitochondrial respiration of the host protozoan [5].

6.4. Telling metabolic stories to explore metabolomics data

The increasing availability of metabolomics data enables to better understand the metabolic processes involved in the immediate response of an organism to environmental changes and stress. The data usually come in the form of a list of metabolites whose concentrations significantly changed under some conditions, and are thus not easy to interpret without being able to precisely visualize how such metabolites are interconnected.

We presented a method that enables to organize the data from any metabolomics experiment into metabolic stories [18]. Each story corresponds to a possible scenario explaining the flow of matter between the metabolites of interest. These scenarios may then be ranked in different ways depending on which interpretation one wishes to emphasize for the causal link between two affected metabolites: enzyme activation, enzyme inhibition or domino effect on the concentration changes of substrates and products. Equally probable stories under any selected ranking scheme can be further grouped into a single anthology that summarizes, in a unique subnetwork, all equivalently plausible alternative stories. An anthology is simply a union of such stories. We detailed an application of the method to 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. We discussed several interpretations for the changes we see, and we suggested hypotheses that could in principle be experimentally tested. Noticeably, our method requires simple input data and could be used in a wide variety of applications.

6.5. MiRNA and co: Methodologically exploring the world of small RNAs

We developed a reliable, robust, and much faster method for the prediction of pre-miRNAs. With this method, we aimed mainly at two goals: efficiency and flexibility. Efficiency was made possible by means of a quadratic algorithm. Since the majority of the predictors use a cubic algorithm to verify the pre-miRNA hairpin structure, they may take too long when the input is large. Flexibility relies on two aspects, the input type and the organism clade. MIRINHO can receive as input both a genome sequence and small RNA sequencing (sRNA-seq) data of both animal and plant species. To change from one clade to another, it suffices to change the lengths of the stem-arms and of the terminal loop. Concerning the prediction of plant miRNAs, because their premiRNAs are longer, the methods for extracting the hairpin secondary structure are not as accurate as for shorter sequences. With MIRINHO, we also addressed this problem, which enabled to provide premiRNA secondary structures more similar to the ones in MIRBASE than the other available methods.

Mirinho served as the basis to two other issues we addressed. The first issue led to the treatment and analysis of sRNA-seq data of *Acyrthosiphon pisum*, the pea aphid. The goal was to identify the miRNAs that are expressed during the four developmental stages of this species, allowing further biological conclusions concerning the regulatory system of such an organism. For this analysis, we developed a whole pipeline, called MIRINHOPIPE, at the end of which MIRINHO was aggregated.

We then moved on to the second issue, that involved problems related to the prediction and analysis of non-coding RNAs (ncRNAs) in the bacterium *Mycoplasma hyopneumoniae*. A method, called ALVINHO, was thus developed for the prediction of targets in this bacterium, together with a pipeline for the segmentation of a numerical sequence and detection of conservation among ncRNA sequences using a k-partite graph.

We finally addressed a problem related to motifs, that is to patterns, that may be composed of one or more parts, that appear conserved in a set of sequences and may correspond to functional elements. This had already been addressed in a robust method called Smile. However, depending on the input parameters, the output may be too large to be tractable, as was realized in other works of the team. We then presented some clustering solutions to group the motifs that may correspond to a same biological element, and thus to better distinguish the biologically significant ones from noise that may be present in what often are large outputs from many motif extraction algorithms.

The publications related to this area of research are either submitted or in preparation (to be submitted in the first months of the year).

6.6. Efficient Algorithms for analysing RNA-seq Data

In the last years, we had addressed the problem of identifying and quantifying variants (alternative splicing and genomic polymorphism) in RNA-seq data when no reference genome is available, without assembling the full transcripts. Based on the fundamental idea that each variant corresponds to a recognizable pattern, a bubble, in a de Bruijn graph constructed from the RNA-seq reads, we propose a general model for all variants in such graphs. We then introduced an exact algorithm, called KISSPLICE, to extract alternative splicing events. We had showed that it enables to identify more correct events than general purpose transcriptome assemblers.

The main time bottleneck in the KISSPLICE algorithm is the bubble enumeration step. Thus, in an effort to make our method as scalable as possible, we had modified Johnson's cycle listing algorithm (Johnson (1975)) to enumerate bubbles in general directed graphs, while maintaining the same time complexity. We now proposed, using a different enumeration technique, an algorithm to list bubbles with path length constraints in weighted directed graphs [29]. For a graph with n vertices and m edges, the method we propose lists all bubbles with a given source in $O(n(m + nlogn))$ delay. Moreover, we experimentally showed that this algorithm is several orders of magnitude faster than the listing algorithm of KISSPLICE to identify bubbles corresponding to alternative splicing events.

Additionally, we showed that the same techniques used to list bubbles can be applied to one classical enumeration problem: K-shortest paths problems [29]. We considered a different parameterisation of the K-shortest paths problem: instead of bounding the number of st-paths, we bound the weight of the st-paths. We present a general scheme to list bounded length st-paths in weighted graphs that takes $O(nt(n, m))$ time per path, where $t(n, m)$ is the time for a single source shortest path computation. This algorithm uses memory linear in the size of the graphs, independent of the number of paths output. For undirected non-negatively weighted graphs, we also show an improved algorithm that lists all st-paths with bounded length in $O((m + t(n, m)))$ time per path.

The main memory bottleneck in KISSPLICE is the construction and representation of the de Bruijn graph. Thus, again with the goal to make our method as scalable as possible, we propose a new compact way to build and represent a de Bruijn graph improving over the state of the art [22]. We show both theoretically and experimentally that our approach uses 30% to 40% less memory than such state of the art, with an insignificant impact on the construction time. Our de Bruijn graph representation is general, in other words it is not restricted to the variation finding or RNA-seq context, and can be used as part of any algorithm that represents NGS data with de Bruijn graphs.

A major issue when analysing transcriptomes using short sequencing reads is to be able to deal with repeats that are longer than the reads. We proposed a first explicit model for large families of inexact repeats in the de Bruijn Graphs generated from RNA-seq data [21]. Taking advantage of this modelling, we also proposed an efficient algorithm which enumerates alternative splicing events without traversing repeat-induced subgraphs, therefore offering a first answer to one the main question left open at the end of Gustavo Sacomoto's PhD [4].

Motivated by previous work on the classical problem of listing cycles, we also studied from a more purely theoretical point of view how to list chordless cycles [28]. We thus developed an amortized $\tilde{O}(|V|)$-delay algorithm for listing chordless cycles in undirected graphs. Chordless cycles are very natural structures in undirected graphs, with an important history and distinguished role in graph theory. The best known solution to list all the C chordless cycles contained in an undirected graph $G = (V, E)$ takes $O(|E|2 + |E| \cdot C)$ time. In this paper we provide an algorithm taking $\tilde{O}(|E| + |V| \cdot C)$ time. We also show how to obtain the same complexity for listing all the P chordless st-paths in G (where C is replaced by P).

6.7. Reference-free detection of isolated SNPs

Detecting single nucleotide polymorphisms (SNPs) between genomes is becoming a routine task with next-generation sequencing. Generally, SNP detec- tion methods use a reference genome. As non-model organisms are increasingly investigated, the need for reference-free methods has been amplified. Most of the existing reference-free methods have fundamental limitations: they can only call SNPs between exactly two datasets, and / or they require a prohibitive amount of computational resources. V. Lacroix participated in the developement of a method, called DISCOSNP to detect both heterozygous and homozygous isolated SNPs from any number of read datasets, without a reference genome, and with very low memory and time footprints (billions of reads can be analyzed with a standard desktop computer) [25]. To facilitate downstream genotyping analyses, DISCOSNP ranks predictions and outputs quality and coverage per allele. Compared to finding isolated SNPs using a state-of-the-art assembly and mapping approach, DISCOSNP requires significantly less computational resources, shows similar precision / recall values, and highly ranked predictions are less likely to be false positives. An experimental validation was conducted on an arthropod species (the tick *Ixodes ricinus*) on which de novo sequencing was performed. Among the predicted SNPs that were tested, 96% were successfully genotyped and truly exhibited polymorphism.

6.8. Endothelial, epithelial, and fibroblast cells exhibit specific splicing programs independently of their tissue of origin

Alternative splicing is the main mechanism of increasing the proteome diversity coded by a limited number of genes. It is well established that different tissues or organs express different splicing variants. However, organs are composed of common major cell types, including fibroblasts, epithelial, and endothelial cells. By analysing large-scale data sets generated by The ENCODE Project Consortium and after extensive RT-PCR validation, we demonstrated that each of the three major cell types expresses a specific splicing program independently of its organ origin [17]. Furthermore, by analysing splicing factor expression across samples, publicly available splicing factor binding site data sets (CLIP-seq), and exon array data sets after splicing factor depletion, we identified several splicing factors that contribute to establishing these cell type-specific splicing programs.

6.9. Length and symmetry on the sorting by weighted inversions problem

Large-scale mutational events that occur when stretches of DNA sequence move throughout genomes are called genome rearrangement events. In bacteria, inversions are one of the most frequently observed rearrangements. In some bacterial families, inversions are biased in favour of symmetry as shown by recent research. In addition, several results suggest that short segment inversions are more frequent in the evolution of microbial genomes. Despite the fact that symmetry and length of the reversed segments seem very important, they have not been considered together in any problem in the genome rearrangement field. We defined the problem of sorting genomes (or permutations) using inversions whose costs are assigned based on their lengths and asymmetries [27]. We presented five procedures and assessed their performance on small sized permutations. The ideas presented in the paper provide insights to solve the problem and set the stage for a proper theoretical analysis.

6.10. Efficient tree reconciliation enumerator plus cophylogeny reconstruction algorithm via an Approximate Bayesian Computation

Phylogenetic tree reconciliation is the approach of choice for investigating the co-evolution of sets of organisms such as hosts and parasites. It consists in a mapping between the parasite tree and the host tree using event-based maximum parsimony. Given a cost model for the events, many optimal reconciliations are however possible. Any further biological interpretation of them must therefore take this into account, making the capacity to enumerate all optimal solutions a crucial point. Only two algorithms currently exist that attempt such enumeration; in one case not all possible solutions are produced while in the other not all cost vectors are currently handled. Our objective in addressing this problem was two-fold. The first was to fill this gap, and the second was to test whether the number of solutions generally observed can be an issue in terms of interpretation.

We presented a polynomial-delay algorithm called EUCALYPT for enumerating all optimal reconciliations [12]. We showed that in general many solutions exist. We gave an example where, for two pairs of host-parasite trees having each less than 41 leaves, the number of solutions is 5120, even when only time-feasible ones are kept. To facilitate their interpretation, those solutions were also classified in terms of how many of each event they contain. The number of different classes of solutions may thus be notably smaller than the number of solutions, yet they may remain high enough, in particular for the cases where losses have cost 0. In fact, depending on the cost vector, both numbers of solutions and of classes thereof may increase considerably (for the same instance, to respectively 4080384 and 275). To further deal with this problem, we introduced and analysed a restricted version where host-switches are allowed to happen only between species that are within some fixed distance along the host tree. This restriction allows us to reduce the number of time-feasible solutions while preserving the same optimal cost, as well as to find time-feasible solutions with a cost close to the optimal in the cases where no time-feasible solution is found.

Despite an increasingly vast literature on cophylogenetic reconstructions for studying host-parasite associations, understanding the common evolutionary history of such systems remains a problem that is far from being solved. Most algorithms for host-parasite reconciliation use an event-based model, where the events include in general (a subset of) cospeciation, duplication, loss, and host-switch. All known parsimonious event-based methods then assign a cost to each type of event in order to find a reconstruction of minimum cost. This is what we did ourselves in EUCALYPT. The main problem with this approach is that the cost of the events strongly influences the reconciliation obtained.

To deal with this problem, we developed an algorithm, called COALA, for estimating the frequency of the events based on an approximate Bayesian computation approach [8]. 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 estimation of the frequency of the events in cases where the dataset consists of trees with a large number of taxa.

We evaluated our method on simulated and on biological datasets. We showed that in both cases, for the same pair of host and parasite trees, different sets of frequencies for the events lead to equally probable solutions. Moreover, often these solutions differ greatly in terms of the number of inferred events. It appears crucial to take this into account before attempting any further biological interpretation of such reconciliations. More generally, we also showed 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.

6.11. Others

Other works, often experimental were also developed during 2014 and published in a number of papers [6], [7], [9], [10], [11], [13], [16], [19], [20], [24], [26].

BEAGLE Project-Team

5. New Results

5.1. Highlights of the Year

We organized two satellite workshops of international conferences:

- The Aevol tutorial during ALife 2014 (July 30 - August 2, New York) http://www.aevol.fr/alifeTutorial

- The "Computational Methods and Modeling of Astrocyte Physiology and Neuron-Glia Interactions" workshop during the Computational NeuroScience 2014 conference (July 26 - 31, Quebec City, Canada)

These highlight our active presence in the scientific life of our two sub-domains in major conferences.

5.2. Sparse short-distance connections enhance calcium wave propagation in a 3D model of astrocyte networks

Participants: H. Berry, J. Lallouette, M. De Pittá

Traditionally, astrocytes have been considered to couple via gap-junctions into a syncytium with only rudimentary spatial organization. However, this view is challenged by growing experimental evidence that astrocytes organize as a proper gap-junction mediated network with more complex region-dependent properties. On the other hand, the propagation range of intercellular calcium waves (ICW) within astrocyte populations is as well highly variable, depending on the brain region considered. This suggests that the variability of the topology of gap-junction couplings could play a role in the variability of the ICW propagation range. Since this hypothesis is very difficult to investigate with current experimental approaches, we explored it using a biophysically realistic model of three-dimensional astrocyte networks in which we varied the topology of the astrocyte network, while keeping intracellular properties and spatial cell distribution and density constant. Computer simulations of the model suggest that changing the topology of the network is indeed sufficient to reproduce the distinct ranges of ICW propagation reported experimentally. Unexpectedly, our simulations also predict that sparse connectivity and restriction of gap-junction couplings to short distances should favor propagation while long–distance or dense connectivity should impair it. Altogether, those results provide support to recent experimental findings that point towards a significant functional role of the organization of gap-junction couplings into proper astroglial networks. Dynamic control of this topology by neurons and signaling molecules could thus constitute a new type of regulation of neuron-glia and glia-glia interactions.

This result has been published in [18] and as conference talks. It is based on J. Lallouette's PhD thesis work in collaboration with M. De Pittà (postdoc in the team) and E Ben-Jacob, Tel Aviv University, Israel.

5.3. Glutamate Mediated Astrocytic Filtering of Neuronal Activity

Participants: H. Berry, J. Lallouette, M. De Pittá

Neuron-astrocyte communication is an important regulatory mechanism in various brain functions but its complexity and role are yet to be fully understood. In particular, the temporal pattern of astrocyte response to neuronal firing has not been fully characterized. Here, we used neuron-astrocyte cultures on multi-electrode arrays coupled to Ca2+ imaging and explored the range of neuronal stimulation frequencies while keeping constant the amount of stimulation. Our results reveal that astrocytes specifically respond to the frequency of neuronal stimulation by intracellular Ca2+ transients, with a clear onset of astrocytic activation at neuron firing rates around 3-5 Hz. The cell-to-cell heterogeneity of the astrocyte Ca2+ response was however large and increasing with stimulation frequency. Astrocytic activation by neurons was abolished with antagonists of type I metabotropic glutamate receptor, validating the glutamate-dependence of this neuron-to-astrocyte

pathway. Using a realistic biophysical model of glutamate-based intracellular calcium signaling in astrocytes, we suggest that the stepwise response is due to the supralinear dynamics of intracellular IP3 and that the heterogeneity of the responses may be due to the heterogeneity of the astrocyte-to-astrocyte couplings via gap junction channels. Therefore our results present astrocyte intracellular Ca2+ activity as a nonlinear integrator of glutamate-dependent neuronal activity.

This result has been published in a paper currently in press, [26] and is a direct result from J. Lallouette's PhD thesis in collaboration with Y. Hanein's group, in Tel Aviv University (for the experimental measurements), M. De Pittà (postdoc in the team), and E Ben-Jacob, Tel Aviv University, Israel.

5.4. Space-induced bifurcation in repression-based transcriptional circuits

Participants: H. Berry, A. Lo Van

Experimental measurements of the mobility of macromolecules, especially proteins, in cells and their membranes consistently report transient subdiffusion with possibly position-dependent —non-homogeneous— properties. However, the spatiotemporal dynamics of protein mobility when transient subdiffusion is restricted to a subregion of space is still unclear. We have investigated the spatial distribution at equilibrium of proteins undergoing transient subdiffusion due to continuous-time random walks (CTRW) in a restricted subregion of a two-dimensional space. Our Monte-Carlo simulations suggest that this process leads to a non-homogeneous spatial distribution of the proteins at equilibrium, where proteins increasingly accumulate in the CTRW subregion as its anomalous properties are increasingly marked. These results suggest that, even though they exhibit the same time-dependence of the mean-squared displacement, the different scenarios proposed to account for subdiffusion in the cell lead to different protein distribution in space, even at equilibrium and without coupling with reaction. We also we assessed the influence of the spatial distribution of the genes on the dynamics of 3-gene transcriptional ring networks regulated by repression, i.e. repressilator circuits. Our simulations suggest that variations of spatial parameters – namely the degree of demixing of the positions of the gene or the spatial range of the mRNA and proteins (i.e. the typical distance they travel before degradation) – have dramatic effects by switching the dynamical regime from spontaneous oscillations to a stationary state where each species fluctuates around a constant value. By analogy with the bifurcations arising from the variation of kinetic parameters, we referred to those transitions as space-induced bifurcations. Therefore, our results strongly support the idea that the spatial organization of the molecular actors of transcriptional networks is crucial for the dynamics of gene expression and suggest that the spatial localization of the synthetic genes in the cell could be used as an additional toggle to control the dynamics of the inserted construct in synthetic biology experiments.

This group of results has been published in [20], [13], [12] and [23]. It consists in the PhD and Master works of B. Caré and A. Lo Van, respectively, and a collaboration with H Chaté, CEA, Saclay.

5.5. Modeling interaction of transcription processes in neighbour genes

Participants: G. Beslon, S. Meyer

During the transcription process, the genetic sequence encoded in the DNA molecule is expressed by an enzymatic complex. This process is often considered as independent for each gene, despite numerous reported cases of one transcribed gene perturbing a neighbour gene's expression, which is then regarded as a side-effect. Here, we suggest in the contrary that such interactions are a widespread feature, resulting from the propagation along the DNA molecule of mechanical stress generated during gene transcription. This torsional stress modifies the facility with which the transcription machinery separates the two strands of the double-helix in order to access the bases, and thus the expression level of any gene located nearby. We develop a quantitative model of this effect, showing that it depends strongly on the orientation of the genes, which is confirmed by the analysis of in vivo expression levels in the drosophila genome. This observation suggests that torsional coupling may play an important role in genetic regulation, and might favor the orientation-dependent co-localization of genes involved in similar functions, which need to be expressed together.

Publication: [21]

5.6. A model of genome size evolution

Participants: G. Beslon, C. Knibbe, S. Fisher

Even though numerous genomic 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.

In collaboration with Samuel Bernard (Inria Dracula Team and Institut Camille Jordan, UMR CNRS 5208, Lyon), we have developed a model for genome size evolution that takes into account both local mutations such as small insertions and small deletions, and large chromosomal rearrangements such as duplications and large deletions. We introduced the possibility of undergoing several mutations within one generation. The model, albeit minimalist, revealed a non-trivial spontaneous dynamics of genome size: in the absence of selection, an arbitrary large part of genomes remains beneath a finite size, even for a duplication rate 2.6-fold higher than the rate of large deletions, and even if there is also a systematic bias toward small insertions compared to small deletions. Specifically, we showed that the condition of existence of an asymptotic stationary distribution for genome size non-trivially depends on the rates and mean sizes of the different mutation types. We also gave upper bounds for the median and other quantiles of the genome size distribution, and argue that these bounds cannot be overcome by selection. Taken together, these results show that the spontaneous dynamics of genome size naturally prevents it from growing infinitely, even in cases where intuition would suggest an infinite growth. This work was part of Stephan Fischer's PhD thesis, which was defended in December 2013.
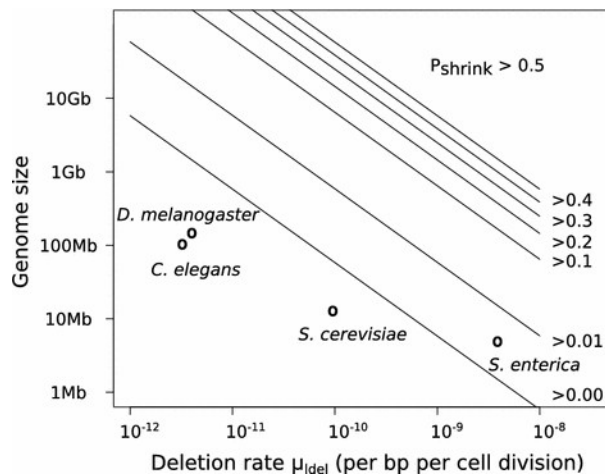
Figure 2. Comparison of the bounds on genome size with the genome size for four organisms. Spontaneous deletion rates were computed per base pair and per cell division from experimental data on mutation accumulations for the bacterium Salmonella enterica, the budding yeast Saccharomyces cerevisiae, the worm Caenorhabditis elegans and the fruit fly Drosophila melanogaster. The value next to each line is the lower bound for the probability that a genome located along this line will shrink at the next step in our model for equal duplication and deletion rates.

This year, using quantitative numerical examples with parameters taken from biological data, we showed that, in practice, a shrinkage bias appears very quickly in genomes undergoing mutation accumulation, even though DNA gains and losses appear to be perfectly symmetrical at first sight. This spontaneous dynamics provides the genome with a stability-related size limit below which it can be influenced by other evolutionary forces (selection, drift, biases, ...).

All this work has been published this year [15], and is already mentioned as "most read article" by Springer.

5.7. A novel view on reductive evolution

Participants: G. Beslon, C. Knibbe, B. Batut

Bacterial genomes show substantial variations in size. The smallest bacterial genomes are those of endocellular symbionts of eukaryotic hosts, which have undergone massive genome reduction and show patterns that are consistent with the degenerative processes that are predicted to occur in species with small effective population sizes. However, similar genome reduction is found in some free-living marine cyanobacteria that are characterized by extremely large populations. Using a combination of bioinformatics approaches and of in silico experimental evolution (with the aevol model), we have been able to propose a scenario that explains the reductive evolution of marine bacteria.

This work was part of Bérénice Batut's PhD thesis [10], which was defended in November 2014. Bérénice was co-supervised by Guillaume Beslon and Carole Knibbe (Inria BEAGLE team) for the simulations and by Gabriel Marais and Vincent Daubin (Laboratoire de Biométrie et Biologie Evolutive, UMR CNRS 5558) for the genomic analyses. This work had already yielded a publication in 2013 [34]. This year, we published a review in the high-level journal [11]. The scenario proposed in the PhD manuscript, as well as the simulations and analyses done this year to support it, should be published in 2015.

5.8. Genome evolution aware gene trees

Participant: E. Tannier

Traditionally the inference of a gene tree is made from a multiple alignment of homologous sequences according to a model of molecular evolution. Trees for several gene families are thus constructed one by one, independently from each other. Constructed this way trees often carry unresolutions or bad resolutions. Information for their full resolution may lie in the poorly exploited dependency between gene families, each bringing information for the resolution of the others. We used several kinds of such dependencies in the construction of gene trees: information from a species tree through a model of gene content evolution, information from extant synteny through ortholog predictions, and information from ancestral synteny through a model of gene neighborhood evolution. We developed, improved, implemented and gave a user interface to several "correction" techniques, yielding a series of correction modules called "RefineTree". We tested its parts on simulated data and apply it on the full set of gene families from the Ensembl Compara database. We showed that according to several measures including the tree likelihood computed from sequence evolution, the stability of genome content and the linearity of ancestral chromosomes, trees corrected by refineTree are arguably more plausible than the ones stored by Ensembl.

This work has been achieved by Magali Semeria, Laurent Gueguen (LBBE) and Eric Tannier in Lyon, in collaboration with Nadia El-Mabrouk's group from the computer science department of the university of Montreal. This collaboration started when Nadia El-Mabrouk was an Inria visiting professor in our team in 2012 and 2013. An article has been submitted.

5.9. Variable food availability increases weight: a mathematical prediction

Participant: H. Soula

Due to the conservation of energy, the energy storage in adipose tissue reflect the difference of energy expenditure and energy intake. Without change in physical activity, the main paradigm has always been that this storage does not depend on the timing of intake but on its whole temporal integration: the overall food intake. However, mammal and especially rats can compensate energy expenditure to save energy in case of starving. This adaptation should provoke variation in energy expenditure when food availability varies in time. Using animal experiments and mathematical modelling, we showed that indeed food availability variation - while conserving the same amount of energy - can disrupt and perturb energy balance. Submitted to variation in availability with a period above 4 weeks, rats where bigger with higher fat mass than control. Even so these

rats had eaten the same amount of food as the control group during the same period. Our mathematical model uses delay equations and can predict both the food intake and the body weight variations. We showed that delay in energy saving adaptation cause this variation and estimate the lag at 1 week. This result could very well apply to humans in the so called 'yoyo regime'. Regime that are stopped are a typical case of food intake variation and could cause greater fat accretion instead of body weight reduction. We show that this should happen if the regime lasts longer than one week.

This result has been the subject of an article in the weekly journal of Inserm Rhônes-Alpes with an interview of author H. Soula.

5.10. Insights on gene family dynamics from digital genetics experiments

Participants: C. Knibbe

Gene families are sets of homologous genes formed by duplications of a single original gene. Inferring their history in terms of gene duplications, gene losses and gene mutations yields fundamental insights into the molecular basis of evolution. However, the traditional approach, the phylogenetic inference of gene family evolution, faces two difficulties: (i) the delimitation of gene families based on sequence similarity, and (ii) the fact that the models of evolution used for reconstruction are tested against simulated data that are produced by the model itself. This year, we showed that digital genetics, or in silico experimental evolution, can provide thought-provoking synthetic gene family data, robust to rearrangements in gene sequences and, most importantly, not biased by where and how we think natural selection should act. Using Aevol, we analyzed the evolution of 3,512 synthetic gene families under directional selection. The turnover of gene families in evolutionary runs was such that only 21% of those families would be accessible for classical phylogenetic inference. Extinct families showed patterns different from the final, observable ones, both in terms of dynamics of gene gains and losses and in terms of gene sequence evolution. This study also reveals that gene sequence evolution, and thus evolutionary innovation, occurred not only through local mutations, but also through chromosomal rearrangements that re-assembled parts of existing genes.

This work was published in the international conference ALIFE 2014 [28].

<p align="center" style="color:red">DRACULA Project-Team</p>

6. New Results

6.1. Highlights of the Year

- Marine Jacquier and Fabien Crauste (in collaboration with C.O. Soulage and H.A. Soul) published a paper ([18], see also § 6.7) in PLoS ONE 2014.

- Sotiris Prokopiou, Loic Barbarroux, Samuel Bernard, Olivier Gandrillon and Fabien Crauste (in collaboration with J. Mafille, Y. Leverrier, C. Arpin and J. Marvel) published a paper ([21], see also § 6.2) in Computation 2014.

- We organized a session "Deterministic and stochastic models in biology and medicine" at 10th AIMS Conference on Dynamical Systems, Differential Equations and Applications, Madrid (Spain), 7 - 11 July 2014 http://www.aimsciences.org/conferences/2014/.

- Our project entitled "Prion and Alzheimer: mathematical modeling and experiments dealing with a dangerous liaison" has been granted by the French Association France Alzheimer, and has been selected with 3 other projects amongst 14 supported works to be part of a scientific popularizing broadcasting campaign through a short scientific cartoon http://www.francealzheimer.org/projets-soutenus-cette-ann%C3%A9e/lab-alz-comprendre-enjeux-recherche/964 and https://www.youtube.com/watch?v=X0mLf8IJhV4&list=PLCq-e7n2r6Wgo3kaseDHetNAPAG7y9B-d.

6.2. Multi-scale model of the CD8 T cell immune response

We presented in [21] the first multi-scale model of CD8 T cell activation in a lymph node, following an acute infection. CD8 T cell dynamics are described using a cellular Potts model (hence cells are discrete interacting objects), whereas intracellular regulation is associated with a continuous system of nonlinear ordinary differential equations focusing on the dynamics of key proteins. This model allows to reproduce the dynamics of CD8 T cells over a five days period (corresponding to the activation and differentiation into effector cells) and is currently used to characterize the generation of memory cells.

6.3. Mathematical model of hematopoiesis

We investigate in [5] a mathematical model of blood cell production in the bone marrow (hematopoiesis). The model 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.

6.4. The role of spatial organization of cells in erythropoiesis

Erythropoiesis, the process of red blood cell production occurs mainly in the bone marrow. The functional unit of mammalian erythropoiesis, the erythroblastic island, consists of a central macrophage surrounded by adherent erythroid progenitor cells (CFU-E/Pro-EBs) and their differentiating progeny, the erythroblasts. Central macrophages display on their surface or secrete various growth or inhibitory factors that influence the fate of the surrounding erythroid cells. CFU-E/Pro-EBs have three possible fates : a) expansion of their numbers without differentiation, b) differentiation into reticulocytes that are released into the blood, c) death by apoptosis. CFU-E/Pro-EB fate is under the control of a complex molecular network, that is highly dependent upon environmental conditions in the erythroblastic island. In order to assess the functional role of space coupled with the complex network behavior in erythroblastic islands, we developed hybrid discrete-continuous models of erythropoiesis. In [13], a model was developed in which cells are considered as individual physical objects, intracellular regulatory networks are modeled with ordinary differential equations and extracellular concentrations by partial differential equations. We used this model to investigate the impact of an important difference between humans and mice in which mature late-stage erythroblasts produce the most Fas-ligand in humans, whereas early-stage erythroblasts produce the most Fas-ligand in mice. Although the global behaviors of the erythroblastic islands in both species were similar, differences were found, including a relatively slower response time to acute anemia in humans. Also, our modeling approach was very consistent with in vitro culture data, where the central macrophage in reconstituted erythroblastic islands has a strong impact on the dynamics of red blood cell production. Conclusions: The specific spatial organization of erythroblastic islands is key to the normal, stable functioning of mammalian erythropoiesis, both in vitro and in vivo. Our model of a simplified molecular network controlling cell decision provides a realistic functional unit of mammalian erythropoiesis that integrates multiple microenvironmental influences within the erythroblastic island with those of circulating regulators of erythropoiesis, such as EPO and glucocorticosteroids, that are produced at remote sites.

6.5. Mathematical modelling of cell polarization

In [19], a fine description of the behaviour of a nonlinear drift diffusion model inspired from spontaneous cell polarization was performed. This model has Keller Segel type properties and in particular, quantitative proofs were obtained for the convergence to steady state or self similar profile or blow up. The behaviour depends on the mass of the initial data.

6.6. Numerical modelling of cell distribution in blood flow

Properties of blood cells and their interaction determine their distribution in flow. It is observed experimentally that erythrocytes migrate to the ow axis, platelets to the vessel wall, and leucocytes roll along the vessel wall. In [2], a three-dimensional model based on Dissipative Particle Dynamics method and a new hybrid (discrete-continuous) model for blood cells is used to study the interaction of erythrocytes with platelets and leucocytes in flow. Erythrocytes are modelled as elastic highly deformable membranes, while platelets and leucocytes as elastic membranes with their shape close to a sphere. Separation of erythrocytes and platelets in flow is shown for different values of hematocrit. Erythrocyte and platelet distributions are in a good qualitative agreement with the existing experimental results. Migration of leucocyte to the vessel wall and its rolling along the wall is observed.

6.7. Mathematical model of food intake dynamics

In [18], we propose a nonlinear mathematical model of food intake dynamics and body weight dynamics, involving the description of several regulating hormones (leptin, ghrelin, insulin). Using a temporal perturbation of food availability in groups of rats, this model is able to predict body weight and food intake variations by taking into account energy expenditure dynamics based on a memory of the previous food intake. This model also allowed us to estimate the memory lag to approximately 8 days. It also explains how important variations in food availability during periods longer than these 8 days can induce body weight gains.

6.8. Long time existence of weak solutions to cross diffusion models

We pointed out a general entropy structure in cross diffusion systems. We used this structure with duality arguments to build a general framework in which weak solutions exist for a long time. This led to two research articles (one [10] in Siam Journal of Mathematical Analysis and one and [32] recently accepted in Comm. In PDE). This was conducted with the help of the ANR KIBORD.

6.9. Mathematics of Darwin's diagram

Darwin illustrated his theory about emergence and evolution of biological species with a diagram. It shows how species exist, evolve, appear and disappear. Our goal in [8] is to give a mathematical interpretation of this diagram and to show how it can be reproduced in mathematical models. It appears that conventional models in population dynamics are not sufficient, and we introduce a number of new models which take into account local, nonlocal and global consumption of resources, and models with space and time dependent coefficients.

6.10. A micellar on-pathway intermediate step explains the kinetics of prion amyloid formation

In [16], we used a strong interdisciplinary collaboration between mathematicians and biologists to exhibit a new element taking an important role in the development of the pathological prion formation. Indeed, in a previous work by Alvarez-Martinez et al. (2011), the authors pointed out some fallacies in the mainstream interpretation of the prion amyloid formation. It appeared necessary to propose an original hypothesis able to reconcile the in vitro data with the predictions of a mathematical model describing the problem. Here, a model is developed accordingly with the hypothesis that an intermediate on-pathway leads to the conformation of the prion protein into an amyloid competent isoform thanks to a structure, called micelles, formed from hydrodynamic interaction. The authors also compared data to the prediction of their model and proposed a new hypothesis for the formation of infectious prion amyloids.

<p style="text-align:center; color:red">**IBIS Project-Team**</p>

5. New Results

5.1. Highlights of the Year

A paper based on the PhD thesis of Diana Stefan was accepted for *PLoS Computational Biology* this year [7].

5.2. 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 8.7 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 8.7 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.

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 the BIPOP project-team, we have shown how tools developed for the simulation of non-smooth mechanical, electrical and control systems can be adapted for this purpose, in a paper published in *Physica D* [1] and presented at the 21st International Symposium on Mathematical Theory of Networks and Systems (MTNS 2014) [12]. 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.

5.3. 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.

Valentin Zulkower, in the framework of his PhD thesis, has developed novel methods for the analysis of reporter gene data, based on the use of regularized linear inversion. This allows a range of estimation problems in the analysis of reporter gene data, notably the inference of growth rate, promoter activity, and protein concentration profiles, to be solved in a mathematically sound and practical manner. We have evaluated the validity of the approach using *in-silico* simulation studies, and observed that the methods are more robust and less biased than indirect approaches usually encountered in the experimental literature based on smoothing and subsequent processing of the primary data, like in WELLREADER. We have applied the methods to the analysis of fluorescent reporter gene data acquired in kinetic experiments with *Escherichia coli*. The methods were shown capable of reliably reconstructing time-course profiles of growth rate, promoter activity, and protein concentration from weak and noisy signals at low population volumes. Moreover, they captured critical features of those profiles, notably rapid changes in gene expression during growth transitions. The linear inversion methods have been implemented in the Python package WELLFARE, and integrated by Michel Page in the web application WELLINVERTER (Section 4.2). This work was submitted for publication early 2015.

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 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* [2]. Stéphan Lacour also reviewed this and other work on RpoS in a publication in *Environmental Microbiology Reports* [4].

A second example derives from the PhD thesis of Diana Stefan. Although from a biological point of view the motility network of *E. coli* is also central in this work, its main thrust lies in clarifying and solving methodological issues in the automated inference of quantitative models of gene regulatory networks from time-series gene expression data, also called reverse engineering in the bioinformatics literature. The application of existing reverse engineering methods is commonly based on implicit assumptions on the biological processes under study. First, the measurements of mRNA abundance obtained in transcriptomics experiments are taken to be representative of protein concentrations. Second, the observed changes in gene expression are assumed to be solely due to transcription factors and other specific regulators, while changes in the activity of the gene expression machinery and other global physiological effects are neglected. While convenient in practice, these assumptions are often not valid and bias the reverse engineering process. In her PhD thesis, Diana Stefan systematically investigated, using a combination of models and experiments, the importance of this bias and possible corrections. She measured with the help of fluorescent reporter genes the activity of genes involved in the FliA-FlgM module of the *E. coli* motility network. From these data, protein concentrations and global physiological effects were estimated by means of kinetic models of gene expression. The results indicate that correcting for the bias of commonly-made assumptions improves the quality of the models inferred from the data. Moreover, it was shown by simulation that these improvements are expected to be even stronger for systems in which protein concentrations have longer half-lives and the activity of the gene expression machinery varies more strongly across conditions than in the FliA-FlgM module. The approach proposed in this study is broadly applicable when using time-series transcriptome data to learn about the structure and dynamics of regulatory networks. The paper describing the work was published in *PLoS Computational Biology* [7].

5.4. Models of carbon metabolism in bacteria

All free-living bacteria have to adapt to a changing environment. Specific regulatory systems respond to particular stresses, but the most common decision bacteria have to make is the choice between alternative carbon sources, each sustaining a specific, maximal growth rate. Many bacteria have evolved a strategy that consists in utilizing carbon sources sequentially, in general favouring carbon sources that sustain a higher growth rate. As long as a preferred carbon source is present in sufficient amounts, the synthesis of enzymes necessary for the uptake and metabolism of less favourable carbon sources is repressed. This phenomenon is called Carbon Catabolite Repression (CCR) and the most salient manifestation of this regulatory choice is diauxic growth, a phenomenon discovered by Jacques Monod more than 70 years ago. Although this system is one of the paradigms of the regulation of gene expression in bacteria, the underlying mechanisms remain controversial. Carbon catabolite repression involves the coordination of different subsystems of the cell - responsible for the uptake of carbon sources, their breakdown for the production of energy and precursors, and the conversion of the latter to biomass.

The complexity of this integrated system, with regulatory mechanisms cutting across metabolism, gene expression, signaling and subject to global physical and physiological constraints, has motivated important modeling efforts over the past four decades, especially in the enterobacterium *Escherichia coli*. Different hypotheses concerning the dynamic functioning of the system have been explored by a variety of modeling approaches. In an article in *Trends in Microbiology* [3], which was initiated during the sabbatical of Andreas Kremling in Grenoble in 2013, we have reviewed these studies and summarized their contributions to the quantitative understanding of carbon catabolite repression, focusing on diauxic growth in E. coli. Moreover, we have proposed a highly simplified representation of diauxic growth that makes it possible to bring out the salient features of the models proposed in the literature and confront and compare the explanations they provide.

A bottleneck in the development of dynamic and quantitatively predictive models of bacterial metabolism, explicitly accounting for the different regulatory mechanisms on the molecular level, is information on the kinetic parameters describing the enzymatic reactions and other molecular interactions. One particularly important piece of information is knowledge of enzyme concentrations. Recent technological advances in quantitative proteomics have made mass spectrometry-based quantitative assays an interesting alternative to more traditional immuno- affinity based approaches for quantifying enzyme concentrations. In particular, these advances have improved specificity and multiplexing capabilities. In a study carried out at CEA Grenoble, a quantification workflow to analyze enzymes involved in central metabolism in *E. coli* was developed. This workflow combined full-length isotopically labeled standards with selected reaction monitoring analysis. The workflow was used to accurately quantify 22 enzymes involved in *E. coli* central metabolism in a wild-type reference strain and two derived strains, optimized for higher NADPH production. Delphine Ropers and Hidde de Jong participated in the analysis of these data. In combination with measurements of metabolic fluxes, we showed that proteomics data can be used to assess different levels of regulation, in particular enzyme abundance and catalytic rate. This is key to the development of predictive kinetic models, but also provides information that can be used for strain design in biotechnology. An article based on this work was published in *Molecular and Cellular Proteomics* [8].

Other ongoing work on the analysis of bacterial metabolism is carried out by Delphine Ropers in collaboration with Inra/INSA in Toulouse, in the framework of the PhD thesis of Manon Morin, supported by a Contrat Jeune Scientifique Inra-Inria. In their respective PhD theses, Stéphane Pinhal and Valentin Zulkower also study specific aspects of carbon metabolism, using both models and experimental data.

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

At the single-cell level, the processes that govern single-cell dynamics in general and gene expression in particular are 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 (computer assisted or genetically synthesized) control of cell populations and even of single cells.

Work in IBIS on the probabilistic gene expression and interaction dynamics at the level of individual cells is addressed in terms of identification of intrinsic noise models from population snapshot data, on the one hand, and the inference of models focusing on cellular variability within isogenic populations from individual cell fluorescence microscopy gene expression profiles, on the other hand. Along with modelling and inference comes analysis of the inferred models in various respects, notably in terms of single-cell state estimation and control. Other problems related with single-cell modelling and extracellular variability are considered in high-eukariotic cells through external collaborations.

In the context of yeast cell response to osmotic shocks, in collaboration with the CONTRAINTES project-team, and colleagues from Université Paris Descartes and University of Pavia (Italy), Eugenio Cinquemani has investigated the use of mixed effects-modelling and identification techniques to characterize individual cell dynamics 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. Starting from identification approaches in pharmacokinetics, we have developed and applied inference methods to the context of microfluidics data, with focus on the budding yeast response to osmotic shocks. First results presented at conference in 2013 have been taken further, both in terms of mathematical analysis of the models developed and in terms of biological interpretation. Model identification and validation were performed together with Andres Gonzales, PhD student at the University of Pavia, who has visited IBIS for six months in 2014. A journal publication is currently being prepared for publication.

In a second line of work, starting from the models inferred in the above collaboration, the problem of real-time state estimation and control of single yeast cells has been considered. Together with the BIOCORE project-team, we have put in place algorithms for state estimation in presence of hybrid random switching and continuous dynamics, and integrated them with a feedback control approach developed by collaborators at TU Delft (the Netherlands). The whole monitoring, estimation and control chain has been deployed and applied *in silico* to the stochastic control of osmosensitive genes in single yeast cells. Methods and results have been presented at the 12th international conference on Computational Methods for Systems Biology (CMSB 2014), whose proceedings have been published as a volume of the LNCS series [14]. It is shown in particular that stochastic model-based estimation and control outperforms existing methods of single-cell control based on deterministic approximations.

Additional work on identification and estimation of hidden states for intrinsic noise models of gene expression/regulation in single bacterial cells, started with reference to arabinose uptake dynamics but also applicable to other regulatory networks in *E. coli*, is being developed. In parallel, collaboration of Eugenio Cinquemani with Marianna Rapsomaniki, PhD student affiliated with the University of Patras (Greece) and ETH Zürich (Switzerland), has been devoted to the analysis of data from Fluorescence Recovery After Photobleaching (FRAP) experiments and the inference of kinetic parameters of protein dynamics in single high-eukariotic cells. As an alternative to current approximate analytical methods, we have explored inference methods based on simulation of biological processes in realistic environments at a particle level. We introduced and demonstrated a new method for the inference of kinetic parameters of protein dynamics, where a limited number of *in-silico* FRAP experiments is used to construct a mapping from FRAP recovery curves to the parameters sought. Parameter estimates from experimental data are then computed by applying the mapping to the observed recovery curves, at virtually no additional price for any number of experiments, along with the application of a bootstrap procedure for determining identifiability of the parameters and confidence intervals for their estimates. After validation on synthetic data, the method was successfully applied to the analysis of the nuclear proteins Cdt1, PCNA and GFPnls in mammalian cells, also shedding light on cell-to-cell variability of the protein kinetics. Method and results have recently been published in *Bioinformatics* [6].

5.6. Growth control in bacteria and biotechnological applications

A bacterial cell adapts its growth rate and the level of gene expression required to sustain growth to the environment, notably to the availability of nutrients providing the molecular building blocks and the energy required for growth. This adaptive response involves the global physiological state of the cell, in particular the activity of the gene expression machinery, and 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. There is 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.

The quantitative models adopting the bottom-up pespective describe the molecular mechanisms controlling the activity of the gene expression machinery. The calibration and analysis of these models is made difficult by their complexity, the nonidentifiability of many parameter values, and the heterogeneity of experimental data sources. To overcome these difficulties, Delphine Ropers and Edith Grac are developing model ensembles with the same structure but different parameter values that are consistent with the experimental data. In collaboration with Jean-Luc Gouzé and Ismail Belgacem from the BIOCORE project-team at Inria Sophia-Antipolis-Méditerranée, they have analysed the dynamical behavior of a central module of these models, which controls the cellular concentration of the RNA polymerase, the key player of the transcriptional machinery. By means of model reduction approaches and monotone system theory, they have analyzed the equilibria of the system and their stability, which they could relate to biological observations on *E. coli*. This work has been published in the proceedings of the 21st International Symposium on Mathematical Theory of Networks and Systems (MTNS 2014) [9] and the 53rd IEEE Conference on Decision and Control (CDC 2014) [10]. A journal article is in preparation.

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 is under the tight control of an inducible promoter. By adjusting the inducer concentration in the medium we can adjust the activity of the gene expression machinery and thereby reversibly switch the growth rate of the bacterium between zero and the maximal growth rate. Our modified *E. coli* strain, described in a paper prepared for submission, 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 7.1).

KALIFFE Project-Team

6. New Results

6.1. Mixed semi-Lagrangian/finite difference methods for plasma simulations

We present an efficient algorithm for the long time behavior of plasma simulations. We will focus on 4D drift-kinetic model, where the plasma's motion occurs in the plane perpendicular to the magnetic field and can be governed by the 2D guiding-center model.

Hermite WENO reconstructions are applied for solving the Vlasov equation. Here we consider an arbitrary computational domain with an appropriate numerical method for the treatment of boundary conditions.

Then we apply this algorithm for plasma turbulence simulations. We first solve the 2D guiding-center model in a D-shape domain and investigate the numerical stability of the steady state. Then, the 4D drift-kinetic model is studied with a mixed method, *i.e.* the semi-Lagrangian method in linear phase and finite difference method during the nonlinear phase. Numerical results show that the mixed method is efficient and accurate in linear phase and it is much stable during the nonlinear phase. Moreover, in practice it has better conservation properties [9].

6.2. High order semi implicit schemes for PDEs

We consider a new formulation of implicit-explicit (IMEX) methods for the numerical discretization of time dependent partial differential equations. We construct several semi-implicit Runge-Kutta methods up to order three. This method is particularly well suited for problems where the stiff and non-stiff components cannot be well separated. We present different numerical simulations for reaction-diffusion, convection diffusion and nonlinear diffusion system of equations. Finally, we conclude by a stability analysis of the schemes for linear problems.

6.3. A Hierarchy of Hybrid Numerical Methods for Multi-Scale Kinetic Equations

We construct a hierarchy of hybrid numerical methods for multi-scale kinetic equations based on moment realizability matrices, a concept introduced by Levermore, Morokoff and Nadiga. Following such a criterion, one can consider hybrid scheme where the hydrodynamic part is given either by the compressible Euler or Navier-Stokes equations, or even with more general models, such as the Burnett or super-Burnett systems [8].

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

We present high order absorbing boundary conditions (ABC)for the Helmholtz equation in 2D, that can adapt to any regular shapedsurfaces. The new ABCs are derived by using the technique ofmicro-diagonalisation to approximate the Dirichlet-to-Neumann map.Numerical results on different shapes illustrate the behavior of thenew ABCs along with high-order finite elements [5].

MOISE Project-Team

6. New Results

6.1. Mathematical Modelling of the Ocean Dynamics

6.1.1. *Numerical Schemes for Ocean Modelling*
Participants: Eric Blayo, Laurent Debreu, Jérémie Demange, Florian Lemarié.

In his PhD, Jérémie Demange has worked 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 ([55]). A special focus has also been put on the numerical representation of internal gravity waves (IGW). The normal mode decomposition also allows the computation of IGW characteristic variables and speeds and thus enables the derivation of monotonic advection schemes ([54]).

In 2014, we worked on the stability constraints for oceanic numerical models ([56]). The idea is to carry a deep analysis of these constraints in order to propose new time stepping algorithms for ocean models. Except for vertical diffusion (and possibly the external mode and bottom drag), oceanic models usually rely on explicit time-stepping algorithms subject to Courant-Friedrichs-Lewy (CFL) stability criteria. Implicit methods could be unconditionally stable, but an algebraic system must be solved at each time step and other considerations such as accuracy and efficiency are less straightforward to achieve. Depending on the target application, the process limiting the maximum allowed time-step is generally different. In this paper, we introduce offline diagnostics to predict stability limits associated with internal gravity waves, advection, diffusion, and rotation. This suite of diagnostics is applied to a set of global, regional and coastal numerical simulations with several horizontal/vertical resolutions and different numerical models. We show that, for resolutions finer that $1/2°$, models with an Eulerian vertical coordinate are generally constrained by vertical advection in a few hot spots and that numerics must be extremely robust to changes in Courant number. Based on those results, we review the stability and accuracy of existing numerical kernels in vogue in primitive equations oceanic models with a focus on advective processes and the dynamics of internal waves. We emphasize the additional value of studying the numerical kernel of oceanic models in the light of coupled space-time approaches instead of studying the time schemes independently from spatial discretizations. From this study, we suggest some guidelines for the development of temporal schemes in future generation multi-purpose oceanic models.

6.1.2. *Coupling Methods for Oceanic and Atmospheric Models*
Participants: Eric Blayo, Mehdi-Pierre Daou, Laurent Debreu, Florian Lemarié, Antoine Rousseau.

6.1.2.1. *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 complex global 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, now extended with ARTELIA Group.

Following the work done by Manel Tayachi in her PhD, Medhi Pierre Daou has started implementing and analyzing a coupling between 1D shallow water equations and 3D Navier Stokes equations. In the context of our partnership with ARTELIA, he uses industrial codes (Mascaret, Telemac and OpenFoam). A first implementation has been realized in an academic testcase, and a second one is presently under implementation in a much more realistic context, in the framework of the European project CRISMA.

6.1.2.2. 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 [57]. For the last few years we have been actively working on the analysis of Schwarz waveform relaxation to apply this type of iterative coupling method to air-sea coupling [95], [96], [94]. In the context of the simulation of tropical cyclone, sensitivity tests to the coupling method have 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. In 2014, this work has been the subject of several invited conferences [23], [24], [25], [26] and collaborations with geophysicists [41], [47], [48].

Past year has also been dedicated to the establishment of strong collaborations between the applied mathematics and the climate community to assess the impact of our work on IPCC-like climate models and to go further in the theoretical work by including the formulation of physical parameterizations. As a results, a PhD-thesis (C. Pelletier) funded by Inria has started in fall 2014 in collaboration with the LSCE (Laboratoire des Sciences du Climat et de l'Environnement). Moreover a PPR (*Projet à partenariat renforcé*) called SIMBAD (SIMplified Boundary Atmospheric layer moDel for ocean modeling purposes) is funded by Mercator-Ocean for the next three years. The aim of this project in collaboration with Meteo-France, Ifremer, LMD, and LOCEAN is to derive a 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.

6.2. Development of New Methods for Data Assimilation

6.2.1. *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 ([84]). A book chapter on multiresolution methods for data assimilation has also been published ([51]).

6.2.2. *Variational Data Assimilation with Control of Model Error*

Participant: 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. The plan is to extend this study focusing on the parallel aspects of such approach.

6.2.3. *Assimilation of Images*

Participants: François-Xavier Le Dimet, Maëlle Nodet, Arthur Vidard, Nelson Feyeux, Vincent Chabot, Nicolas Papadakis.

6.2.3.1. *Direct assimilation of image sequences*

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 [102]. 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 [76]. 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 paper accepted late 2014 [8] and presented in several national [37], [38], [27] and international conferences [30], [28], [29].

Vincent Chabot also defended his PhD in July 2014 [2].

In recent developments [17] we have also used "Level Sets" methods to describe the evolution of the images. The advantage of this approach is that it permits, thanks to the level sets function, to consider the images as a state variable of the problem. We have derived an Optimality System including the level sets of the images.

6.2.3.2. *Optimal transport for image assimilation*

Within the optimal transport project TOMMI funded by the ANR white program (started mid 2011), a new optimization scheme based on proximal splitting method has been proposed to solve the dynamic optimal transport problem. We investigate the use of optimal transport based distances for data assimilation. N. Feyeux started his PhD on this subject last year. The study is still under investigation, but preliminary encouraging results have already been presented twice, in France [68] and Austria [69].

6.2.4. *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 ICIPE conference [31].

6.2.5. *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 a simplified configuration of the NEMO model. A rectangular box model placed in mid-latitudes, and subject 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 in [15] that optimal boundary has a rather complicated geometry which is neither a staircase, nor a straight line. The boundary conditions found in the data assimilation procedure bring the solution toward the reference solution allowing to correct the influence of the rotated grid (see fig. 1).

Adjoint models, necessary to variational data assimilation, have been produced by the TAPENADE software, developed by the SCIPORT 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.
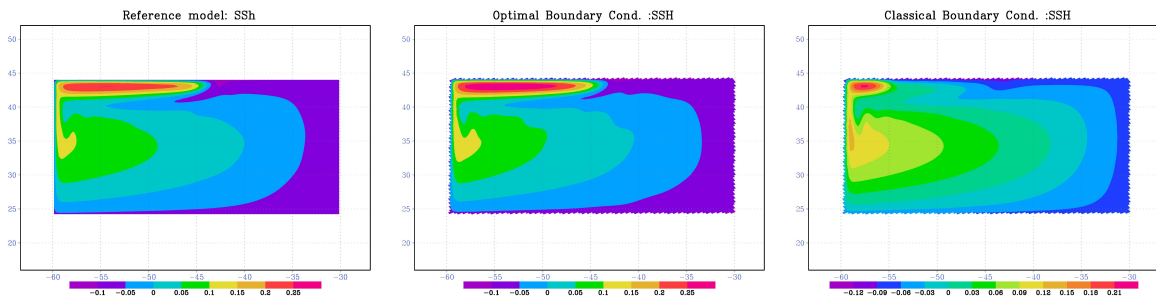
Figure 1. Sea surface elevation: reference solution on the aligned grid (left), solutions on the $30°$ rotated grid with optimal (center) and classical (right) boundary conditions.

6.3. Data Assimilation for Geophysical Models

6.3.1. *Development of a Variational Data Assimilation System for OPA9/NEMO*

Participants: Arthur Vidard, 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 just started in january 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 2). 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 was defended early 2014 [1].
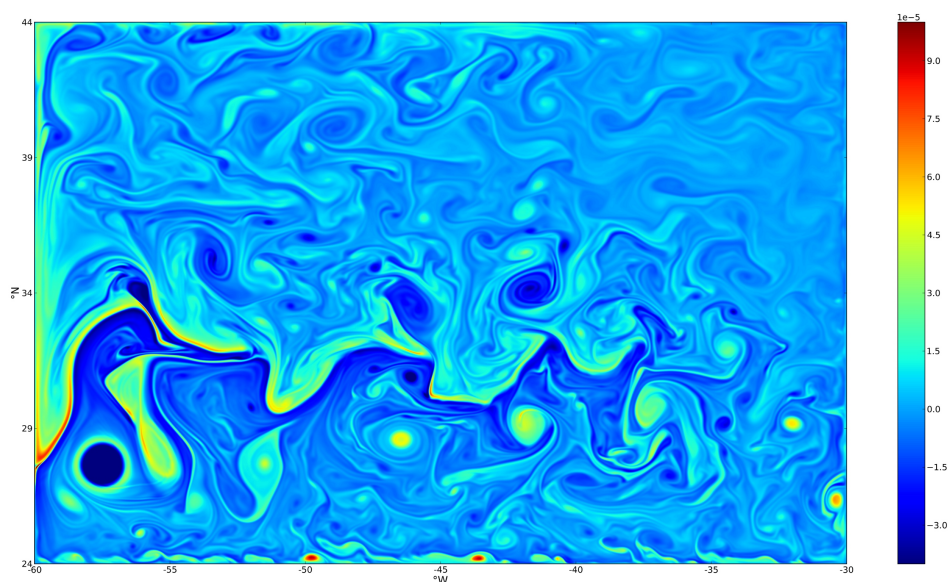
Figure 2. Snapshot of the relative vorticity field (1/s) for an academic oceanic basin model at 1/100° horizontal resolution.

Lastly, multi resolution algorithms have been developed to solve the variational problem. An EU-ITN (International Training Network) project is going to be submitted early 2015 to continue working in this particular aspect.

6.3.2. 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 Géophysique 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 gives poor results.

B. Bonan defended his PhD 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, very promising for the future, as we want to implement this method into a full 3D model. A journal paper has published on this subject [5], and the results have been presented in the conference [46].

6.4. Quantifying Uncertainty

6.4.1. Sensitivity Analysis for Forecasting Ocean Models

Participants: Eric Blayo, Laurent Gilquin, Céline Helbert, François-Xavier Le Dimet, Simon Nanty, Maëlle Nodet, Clémentine Prieur, Laurence Viry, Federico Zertuche.

6.4.1.1. Scientific context

Forecasting geophysical 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' [101] 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 [99] and more recently by Owen [97] but the quantities they estimate still require $O(d)$ samples.

In a recent work [21] we introduce a new approach to estimate all first-order Sobol' indices by using only two samples based on replicated latin hypercubes and all second-order Sobol' indices by using only two samples based on replicated randomized orthogonal arrays. 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. As for these models, some groups of inputs are correlated, Laurent Gilquin extended the approach based on replicated designs for the estimation of grouped Sobol' indices [70].

We can now wonder what are the asymptotic properties of these new estimators, or also of more classical ones. In [10], the authors deal with asymptotic properties of the estimators. In [89], 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 [92], 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 [14]. In [91], 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. These results were also adapted to problems related to more general models such as Shallow-Water equations, in the context of the control of an open channel [72].

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 [71].

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 [90] in the case where input parameters are correlated. Clémentine Prieur supervised Gaëlle Chastaing's PhD thesis on the topic (defended in September 2013) [78]. We obtained first results [79], 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 recently accepted [66] and [80]. We also considered (see the paragraph 6.4.1) the estimation of groups Sobol' indices, with a procedure based on replicated designs. These indices provide information at the level of groups, and not at a finer level, but their interpretation is still rigorous.

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, and a paper has been submitted [74].

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 ([93], [98]): a new estimation 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 part of the thesis 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 part of the thesis was dedicated to the development of a new sequential approach based on a course to fine wavelets algorithm. Federico Zertuche presented his work at the annual meeting of the GDR Mascot Num in 2014 [36].

6.4.1.6. Data assimilation and second order sensitivity analysis

A main advantage of Variational Methods in Data Assimilation is to exhibit a so-called Optimality System (OS) that contains all the available information : model, data, statistics. Therefore a sensitivity analysis (i.e. the evaluation of the gradient) with respect to the inputs of the model has to be carried out on the OS. With iMECH and INM we have applied sensitivity analysis in the framework of a pollution problem in a lake. The application of second order analysis for sensitivity permits to evaluate the sensitivity with respect to observations and furthermore to determine the optimal location of new sensors at points with the highest sensitivity [16], [52].

This methodology has been applied to

- **Oil Spill**. These last years have known several disasters produced by wrecking of ships and drifting platforms with severe consequences on the physical and biological environments. In order to minimize the impact of these oil spills its necessary to predict the evolution of oil spot. Some basic models are available and some satellites provide images on the evolution of oil spots. Clearly this topic is a combination of the two previous one: data assimilation for pollution and assimilation of images. A theoretical framework has been developed with Dr. Tran Thu Ha (iMech).

- **Data Assimilation in Supercavitation (with iMech)**. Some self propelled submarine devices can reach a high speed thanks to phenomenon of supercavitation: an air bubble is created on the nose of the device and reduces drag forces. Some models of supercavitation already exist but are working on two applications of variational methods to supercavitation:
  - Parameter identification : the models have some parameters that can not be directly measured. From observations we retrieve the unknown parameters using a classical formalism of inverse problems.
  - Shape Optimization. The question is to determine an optimum design of the shape of the engine in order to reach a maximum speed.

6.5. Tracking of Mesoscale Convective Systems

Participant: Clémentine Prieur.

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).

A first approach involves adapting the multiple hypothesis tracking (MHT) model originally designed by the NCAR (National Centre for Atmospheric Research) for tracking storms [103] to the data for West Africa. With A. Makris (working on a post-doctoral position), we proposed a Bayesian approach [18], 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 [18] and investigated in the specific context of real data on West Africa (submitted paper). Using PHD (probability hypothesis density) filters adapted to our problem, generalising recent developments in particle filtering for spatio-temporal branching processes (e.g. [77]) 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.

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.

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 the 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 [81] for a review of recent extensions of the notion of return level to the multivariate framework. In the context of environmental risk, [100] 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 [82] 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 [86], [88], 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 [87]. 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 2013. A first paper on data reconstruction has been submitted. It was a necessary step as the initial series contained many missing data.

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 coercivity, 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 [6] and when only partial observational data are available. A paper on the non parametric estimation of the drift has been accepted recently [7] (see Samson et al., 2012, for results for parametric models). As far as the estimation of the diffusion term is concerned, a paper has been submitted [7], in collaboration with J.R. León (Caracas, Venezuela) and P. Cattiaux (Toulouse). Recursive estimators have been also proposed by the same authors in [64] recently submitted.

Note that Professor Jose R. León (Caracas, Venezuela) is now funded by an international Inria Chair and will spend one year in our team, allowing to collaborate further on parameter estimation.

6.8. Land Use and Transport Models Calibration

Participants: Thomas Capelle, Laurent Gilquin, Clémentine Prieur, Arthur Vidard, Peter Sturm, Elise Arnaud.

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 led by the STEEP EPI. This project started early 2013 and two PhD about sensitivity analysis and calibration were launched late 2013. This work led to conference papers [32], [33] and a submitted journal paper [70]

<p style="text-align:center; color:red;">**NUMED Project-Team**</p>

5. New Results

5.1. Highlights of the Year

Vincent Calvez has been awareded an ERC Grant and the prestigious Bronze medal CNRS.

<p style="text-align:center; color:red">STEEP Team</p>

6. New Results

6.1. Highlights of the Year

This year has seen a number of major advances in the team research projects, on several fronts. The first one concerns the most important and time consuming project, namely integrated land use, activity and transport modelling (LUTI modelling). In this respect, the results described in 6.8 below constitute probably the first set of works contributing sophisticated numerical procedures to the calibration and validation of the TRANUS LUTI model.

The second significant breakthrough concerns the completion of a downscaling methof for Material Flow Analysis (MFA), a key aspect in the characterization and understanding of territorial metabolism for decision-help purposes (section 6.2).

Finally, the modelling effort on land use change for the ESNET project has now been mostly completed, and an operational LUCC model has been calibrated and validated for this project (section 6.3).

6.2. Downscaling Material Flow Analysis: the case of the cereals supply chain in France

The spatial reconstruction of the production, trade, transformation and consumption flows of a specific material can become an important decision-help tool for improving resource management and for studying environmental pressures from the producer's to the consumer's viewpoint. One of the obstacles preventing its actual use in the decision-making process is that building such studies at various geographical scales proves to be costly both in time and manpower. We propose a semi-automatic methodology to overcome this issue. First a supply chain model at the national level has to be designed. Supply and use tables are used to handle the data consistently. Finding the appropriate level of detail for both products and industries is an iterative process: with a small number of highly aggregated product categories, the study isn't likely to provide useful information while with a very detailed list of products and industries, finding input data, especially at local scales, won't be feasible. Secondly, national production, transformation, trade and consumption data have to be reconciled in order to respect the law of mass conservation: this is done through constraint optimization. Thirdly, regional supply and use tables are generated (either with direct data or through downscaling of national data using local proxies, e.g. employment statistics) and reconciled, taking into account the additional constraint that regional data must add up to national one.

We applied the methodology on the case of cereals and reconstructed the supply chain flows of the 22 French regions as well as the flows of four nested territories: France, the Rhône-Alpes région, the Isère département and the territory of the SCOT of Grenoble. Uncertainties of output data were estimated via Monte-Carlo simulations. We display the results using our Sankey diagram vizualisation tool. A research paper is in the reviewing process for one of the major journals in this field.

Future steps include coupling this model with economic (added value), social (local employment) and environmental (environmental pressures) aspects in order to provide new information to decision-makers at various administrative levels (from a group of cities to the national level).

6.3. 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 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 now operational, thanks to a major effort on this front in 2014. Both transitions from non urban to urban and use and changes of agricultural practices are now calibrated and validated. The first scenario has been successfully simulated in terms of land use. The three other scenairios of the project are in the final stage of elaboration before simulation, so that the land use change simulation phase of the ESNET project should be completed by the end of April, 2015.

Two resarch papers are in the process of being written on the question of land use practices and their evolution in the study area, and a third one on issues of principle in land use modelling is also underway.

6.4. 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. It has been published in [5].

6.5. On the acceptability of land use transport integrated models by French end users as operational tools: from understanding to daily use

Land Use and Transport Integrated models (LUTIs) are promising approaches for urban planning. There is large literature describing their technical architectures or using them in various scientific contexts. Yet little attention has been paid to expectations of practitioners (planners) and to the daily use of such models. There is clearly an important gap between research and practice: a daily use of LUTIs for the simulation of regional planning policies is still an exception in France, despite important research investments and recent interest of planning agencies., and this situation does not seem to be specific to France. We worked on sheding light on what would make them definitely accepted and more used by planners to evaluate a range of urban and transport policies. To do so, we have interviewed different types of end users in France to identify their motivations and barriers to use LUTI models, in addition to literature study and our own experience dealing with urban planning agencies. We have analysed the main obstacles that prevent LUTIs from being widely used by local authorities. It is important to identify that there are two main issues: 1) Do current LUTIs really answer the questions and practical issues territorial agencies are confronted with on a day-to-day basis? Do they match their interests and expectations? 2) Are current LUTIs suitable with respect to the constraints and limitations of local agencies? The main obstacles associated with these issues are: first, it is difficult to match rather generic models with very specific and varied end users questions; second, it is costly and heavy

to implement and use a LUTI (capacity obstacles); third, there is no guarantee that results of a dedicated LUTI will have any impact on the policy design (decision making obstacles). The results of our analysis show demand for a far more bottom-up oriented approach: the models should consider objectives and general needs of end users to live up to their expectations. Only a closer collaboration between modelers and end users, and more efforts to integrate modeling into urban planning, will make LUTIs considered as relevant approaches.

This work has been done in collaboration with Mathieu Saujot (IDDRI) and Mathieu De Lapparent (IF-SSTAR), and belongs to the work program of CiTIES project.

6.6. Replication procedure for grouped Sobol' indices estimation in dependent uncertainty spaces

Sensitivity analysis studies how the uncertainty on an output of a mathematical model can be attributed to sources of uncertainty among the inputs. Global sensitivity analysis of complex and expensive mathematical models is a common practice to identify influent inputs and detect the potential interactions between them. Among the large number of available approaches, the variance-based method introduced by Sobol' allows to calculate sensitivity indices called Sobol' indices. Each index gives an estimation of the influence of an individual input or a group of inputs. These indices give an estimation of how the output uncertainty can be apportioned to the uncertainty in the inputs. One can distinguish first-order indices that estimate the main effect from each input or group of inputs from higher-order indices that estimate the corresponding order of interactions between inputs. This estimation procedure requires a significant number of model runs, number that has a polynomial growth rate with respect to the input space dimension. This cost can be prohibitive for time consuming models and only a few number of runs is not enough to retrieve accurate informations about the model inputs.

The use of replicated designs to estimate first-order Sobol' indices has the major advantage of reducing drastically the estimation cost as the number of runs becomes independent of the input space dimension. The generalization to closed second-order Sobol' indices relies on the replication of randomized orthogonal arrays. The motivation of this work is to extend this methodology in presence of dependent inputs. Indeed, the case of correlated parameters has to be tackled with caution, as the calculation of single input indices does not provide anymore a proper information, that can be easily interpreted. One strategy is thus to define grouped indices for groups of correlated variables. We address this issue by proposing an approach based on replicated designs and randomized orthogonal arrays that enables to take into account dependency within inputs. We suppose that this dependency can be expressed through constraints. This approach can be used facing any set of constraints at the condition that one is able to provide points in the input space that verify the considered constraints. Guided by our application on a land-use and transport integrated model (LUTI) where some economical parameters are linked by order relations, we focus on the case of sets of linear ordered constraints. Thus we propose a sampling strategy based on the simplex geometric structure, that ensures a proper input space filling.

This work has been done in collaboration with Laurent Gilquin and Clementine Prieur (members of Moise Team), and belongs to the work program of CiTIES project. It is described in [18]. The proposed procedure will be soon applied to study the sensitivity of TRANUS model.

6.7. Specifications for the calibration of Simbad model

"Simbad" is a LUTI model developed by LET. In the context of the CITiESANR project, we have done a comprehensive and detailed study of the parameters of the model in order to fully specify the calibration process of the model. For example, we have specified the objects of interest and indicators, as well as satisfaction criteria. This work has been done in close collaboration with LET.

6.8. Calibration of the TRANUS Land Use Module

The setting up of a LUTI model requires, like most numerical models, at least one phase of parameter estimation. This is concisely referred to here as calibration, although the calibration of a LUTI model also entails

other aspects such as the definition of spatial zones, of economic sectors, etc. The TRANUS LUTI model plus software, like many other existing models, come along with a relatively simple calibration methodology. Most LUTI models indeed perform parameter estimation in a piecewise fashion, by sequentially estimating subsets of parameters. While this reduces the mathematical and computational complexity of calibration, neglecting the interactions across different modules and their parameters, may result in a significant loss of a model's quality. A second issue is that TRANUS, like several other LUTI softwares, employs rudimentary numerical routines for parameter estimation. We aim at reducing these weaknesses.

To do so, we first defined a particular parameter estimation problem for TRANUS properly as an optimisation problem, based on an explicit cost function that is to be minimised (something lacking in many articles on LUTI calibration). Next, we developed a series of numerical estimation schemes to solve this optimisation problem. The main difficulty here was that the model is dynamic; by delving into the model's equations and structure, we were able to unwind the model's dynamics and to make it amenable to standard numerical optimisation by gradient descent type methods [4]. This was first done for the estimation of a particular subset of model parameters (the so-called shadow prices). We have recently started to work on the simultaneous estimation of these and other model parameters.

This work is done in collaboration with Arthur Vidard from the MOISE Inria project-team and Brian Morton from the University of North Carolina at Chapel Hill.

6.9. State of the Art on the Calibration and Validation of LUTI Models

One of the tasks of the CITiES project is to construct an extensive state of the art report on the calibration and validation of LUTI models. We coordinate this effort, which involves all partners of CITiES, together with the project partner LVMT (Nicolas Coulombel). It consists of the definition of a taxonomy, of an extensive literature research and of a critical analysis of this literature. A short publication that explains the goals of this effort and some intermediate findings, has been presented in [3]. The completion of this task is expected for the first semester of 2015.

<p align="center" style="color:red;">AVALON Project-Team</p>

6. New Results

6.1. Energy efficiency of large scale distributed systems

Participants: Laurent Lefèvre, Daniel Balouek Thomert, Eddy Caron, Radu Carpa, Ghislain Landry Tsafack Chetsa, Marcos Dias de Assunçao, Jean-Patrick Gelas, Olivier Glück, Jean-Christophe Mignot, François Rossigneux, Violaine Villebonnet.

6.1.1. *Improving Energy Efficiency of Large Scale Systems without a priori Knowledge of Applications and Services*

Unlike their hardware counterpart, software solutions to the energy reduction problem in large scale and distributed infrastructures hardly result in real deployments. At the one hand, this can be justified by the fact that they are application oriented. At the other hand, their failure can be attributed to their complex nature which often requires vast technical knowledge behind proposed solutions and/or thorough understanding of applications at hand. This restricts their use to a limited number of experts, because users usually lack adequate skills. In addition, although subsystems including the memory and the storage are becoming more and more power hungry, current software energy reduction techniques fail to take them into account. We propose a methodology for reducing the energy consumption of large scale and distributed infrastructures. Broken into three steps known as (i) phase identification, (ii) phase characterization, and (iii) phase identification and system reconfiguration; our methodology abstracts away from any individual applications as it focuses on the infrastructure, which it analyses the runtime behaviour and takes reconfiguration decisions accordingly.

The proposed methodology is implemented and evaluated in high performance computing (HPC) clusters of varied sizes through a Multi-Resource Energy Efficient Framework (MREEF). MREEF implements the proposed energy reduction methodology so as to leave users with the choice of implementing their own system reconfiguration decisions depending on their needs. Experimental results show that our methodology reduces the energy consumption of the overall infrastructure of up to 24% with less than 7% performance degradation. By taking into account all subsystems, our experiments demonstrate that the energy reduction problem in large scale and distributed infrastructures can benefit from more than "the traditional" processor frequency scaling. Experiments in clusters of varied sizes demonstrate that MREEF and therefore our methodology can easily be extended to a large number of energy aware clusters. The extension of MREEF to virtualized environments like cloud shows that the proposed methodology goes beyond HPC systems and can be used in many other computing environments.

6.1.2. *Reservation based Usage for Energy Efficient Clouds: the Climate/Blazar Architecture*

The FSN XLcloud project (cf Section 8.1) strives to establish the demonstration of a High Performance Cloud Computing (HPCC) platform based on OpenStask, that is designed to run a representative set of compute intensive workloads, including more specifically interactive games, interactive simulations and 3D graphics. XLcloud is based on OpenStack, and Avalon is contributing to the energy efficiency part of this project. We have proposed and brought our contribution to Climate, a new resource reservation framework for OpenStack, developed in collaboration with Bull, Mirantis and other OpenStack contributors. Climate allows the reservation of both physical and virtual resources, in order to provide a mono-tenancy environment suitable for HPC applications. Climate chooses the most efficient hosts (flop/W). This metric is computed from the CPU / GPU informations, mixed with real power consumption measurements provided by the Kwapi framework. The user requirements may be loose, allowing Climate to choose the best time slot to place the reservation. Climate has been improved with standby mode features, to shut down automatically the unused hosts. The first release of Climate was done in January 2014. Through the OpenStack process, Climate is now named Blazar.

6.1.3. Clustered Virtual Home Gateway (vHGW)

This result is a joint work between Avalon team (J.P. Gelas, L. Lefevre) and Addis Abeba University (M. Tsibie and T. Assefa). The customer premises equipment (CPE), which provides the interworking functions between the access network and the home network, consumes more than 80% of the total power in a wireline access network. In the GreenTouch initiative (cf Section 8.3), we aim at a drastic reduction of the power consumption by means of a passive or quasi-passive CPE. Such approach requires that typical home gateway functions, such as routing, security, and home network management, are moved to a virtual home gateway (vHGW) server in the network. In our first prototype virtual home gateways of the subscribers were put in LXC containers on a unique GNU/Linux server. The container approach is more scalable than separating subscribers by virtual machines. We demonstrated a sharing factor of 500 to 1000 virtual home gateways on one server, which consumes about 150 W, or 150 to 300 mW per subscriber. Comparing this power consumption with the power of about 2 W for the processor in a thick client home gateway, we achieved an efficiency gain of 5-10x. The prototype was integrated and demonstrated at TIA 2012 in Dallas. In our current work, we propose the Clustered vHGWs Data center architecture to yield optimal energy conservation through virtual machine's migration among physical nodes based on the current subscriber's service access state, while ensuring SLA respective subscribers. Thus, optimized energy utilization of the data center is assured without compromising the availability of service connectivity and QoS preferences of respective subscribers. The last prototype including those new features was integrated and demonstrated recently to the GreenTouch consortium members at Melbourne University.

6.1.4. Energy proportionality with heterogeneous computing resources

This work [16] focuses on improving energy proportionality of large scale virtualized environments. The main problem of such infrastructures is their high static costs due to high idle power consumption of idle servers. Our goal is to reach an infrastructure able to adapt its energy consumption to the current working load. Therefore we propose an original infrastructure composed of heterogeneous computing resources. We consider the heterogeneity at the level of the architecture, and we gather in our platform low power ARM processors together with powerful x86 servers. Around this infrastructure, we are developing a decisional framework to schedule applications on the architecture, or combination of architectures, most suitable to their current needs. The framework reacts dynamically to the resource needs evolutions by migrating the applications to the chosen destinations, and switching off unused nodes to save energy. We validate our scheduling policies by building a simulator based on a set of experimental inputs about power and performance hardware profiles and applications load profiles. This work is jointly done with IRIT Lab. (Toulouse) under the support of Inria Large Scale Initiative Hemera.

6.1.5. Energy efficient Core Networks

This work [11] seeks to improve the energy efficiency of backbone networks by providing an intra-domain Software Defined Network (SDN) approach to selectively turn off a subset of links. To do this, we change the status of router ports and transponders on the two extremities of a link. The status of these components is set to sleep mode whenever a link is not required to transfer data, and brought back to operational state when needed. We have analyzed the implementation issues of an energy-efficient SDN-based traffic engineering in core networks. We propose the STREETE framework (SegmenT Routing based Energy Efficient Traffic Engineering) that represents an online method to switch some links off/on dynamically according to the network load. We have implemented our proposed algorithms in the OMNET++ packet-based discrete event simulator. Experiments considering real network topologies (Germany50 and Ge´ant) and real dynamic traffic matrices allowed us to quantify the trade-off between energy saving and impact of our solution on network performance. As mean to reroute the traffic we use a promising new protocol, SPRING. This comes in contrast with other works, which use classical IP link weights changes or MPLS+RSVP-TE for this purpose. SPRING proved itself well suited for dynamic reconfiguration of the network. Experimental results show that the consumption of 44% of links can be reduced while preserving good quality of service.

6.1.6. Energy aware scheduling for multi data centers clouds

Our work tackles the challenge of improving the energy efficiency of server provisioning and workload management [17]. It introduces a metric allowing infrastructure administrators to specify their preferences between performance and energy savings. We describe a framework for resource management which provides control for informed and automated provisioning at the scheduler level while providing developers (administrator or end-user) with an abstract layer to implement aggregation and resource ranking based on contextual information such as infrastructure status, users' preferences and energy- related external events occurring over time. We integrate our solution in DIET which allows for managing heterogeneous nodes at the middleware layer. The evaluation is performed by means of simulations and real-life experiments on the GRID'5000 testbed. Results show improvements in energy efficiency with minimal impact on application and system performance. Implementation has been used within the industrial project Nu@ge in the context of a federation of modular datacenters.

6.2. Simulation of Large Scale Distributed Systems

Participants: Frédéric Desprez, Jonathan Rouzaud-Cornabas, Frédéric Suter.

6.2.1. *Versatile, Scalable, and Accurate Simulation of Distributed Applications and Platforms*

The study of parallel and distributed applications and platforms, whether in the cluster, grid, peer-to-peer, volunteer, or cloud computing domain, often mandates empirical evaluation of proposed algorithmic and system solutions via *simulation*. Unlike direct experimentation via an application deployment on a real-world testbed, simulation enables fully repeatable and configurable experiments for arbitrary hypothetical scenarios. Two key concerns are accuracy (so that simulation results are scientifically sound) and scalability (so that simulation experiments can be fast and memory-efficient). While the scalability of a simulator is easily measured, the accuracy of many state-of-the-art simulators is largely unknown because they have not been sufficiently validated. In this work we describe recent accuracy and scalability advances made in the context of the SIMGRID simulation framework. A design goal of SIMGRID is that it should be versatile, i.e., applicable across all aforementioned domains. We present quantitative results that show that SIMGRID compares favorably to state-of-the-art domain-specific simulators in terms of scalability, accuracy, or the trade-off between the two. An important implication is that, contrary to popular wisdom, striving for versatility in a simulator is not an impediment but instead is conducive to improving both accuracy and scalability.

6.2.2. *Simulation of MPI Applications with Time-Independent Traces*

Analyzing and understanding the performance behavior of parallel applications on parallel computing platforms is a long-standing concern in the High Performance Computing community. When the targeted platforms are not available, simulation is a reasonable approach to obtain objective performance indicators and explore various hypothetical scenarios. In the context of applications implemented with the Message Passing Interface, two simulation methods have been proposed, on-line simulation and off-line simulation, both with their own drawbacks and advantages. In this work we present an off-line simulation framework, i.e., one that simulates the execution of an application based on event traces obtained from an actual execution. The main novelty of this work, when compared to previously proposed off-line simulators, is that traces that drive the simulation can be acquired on large, distributed, heterogeneous, and non-dedicated platforms. As a result the scalability of trace acquisition is increased, which is achieved by enforcing that traces contain no time-related information. Moreover, our framework is based on an state-of-the-art scalable, fast, and validated simulation kernel.

6.2.3. *Adding Storage Simulation Capacities to the SimGrid Toolkit*

For each kind of distributed computing infrastructures, i.e., clusters, grids, clouds, data centers or supercomputers, storage is a essential component to cope with the tremendous increase in scientific data production and the ever-growing need for data analysis and preservation. Understanding the performance of a storage subsystem or dimensioning it properly is an important concern for which simulation can help by allowing for fast, fully repeatable, and configurable experiments for arbitrary hypothetical scenarios. However, most simulation frameworks tailored for the study of distributed systems offer no or little abstractions or models of storage resources.

In this work we extend SimGrid, a versatile toolkit for the simulation of large-scale distributed computing systems, with storage simulation capacities. We define the required abstractions and propose a new API to handle storage components and their contents in SimGrid-based simulators. Then we characterize the performance of the fundamental storage component that are disks and derive models of these resources. Finally we list several concrete use cases of storage simulations in clusters, grids, clouds, and data centers for which the proposed extension would be beneficial.

6.3. MapReduce Computations on Hybrid Distributed Computations Infrastructures

Participants: Gilles Fedak, Julio Anjos, Asma Ben Cheikh Ahmed.

In this section we report on our efforts to provide MapReduce Computing environments on Hybrid infrastructures, i.e composed of Desktop Grids and Cloud computing environments.

6.3.1. *BIGhybrid - A Toolkit for Simulating MapReduce in Hybrid Infrastructures*

Cloud computing has increasingly been used as a platform for running large business and data processing applications. Although clouds have become extremely popular, when it comes to data processing, their use incurs high costs. Conversely, Desktop Grids, have been used in a wide range of projects, and are able to take advantage of the large number of resources provided by volunteers, free of charge. Merging cloud computing and desktop grids into a hybrid infrastructure can provide a feasible low-cost solution for big data analysis. Although frameworks like MapReduce have been devised to exploit commodity hardware, their use in a hybrid infrastructure raise some challenges due to their large resource heterogeneity and high churn rate. This study introduces BIGhybrid, a toolkit that is used to simulate MapReduce in hybrid environments. Its main goal is to provide a framework for developers and system designers that can enable them to address the issues of Hybrid MapReduce. In this paper, we describe the framework which simulates the assembly of two existing middleware: BitDew- MapReduce for Desktop Grids and Hadoop-BlobSeer for Cloud Computing. The experimental results that are included in this work demonstrate the feasibility of our approach.

6.3.2. *Parallel Data Processing in Dynamic Hybrid Computing Environment Using MapReduce*

In this work, we propose a novel MapReduce computation model in hybrid computing environment called HybridMR is proposed. Using this model, high performance cluster nodes and heterogeneous desktop PCs in Internet or Intranet can be integrated to form a hybrid computing environment. In this way, the computation and storage capability of large-scale desktop PCs can be fully utilized to process large-scale datasets. HybridMR relies on a hybrid distributed file system called HybridDFS, and a time-out method has been used in HybridDFS to prevent volatility of desktop PCs, and file replication mechanism is used to realize reliable storage. A new node priority-based fair scheduling (NPBFS) algorithm has been developed in HybridMR to achieve both data storage balance and job assignment balance by assigning each node a priority through quantifying CPU speed, memory size and I/O bandwidth. Performance evaluation results show that the proposed hybrid computation model not only achieves reliable MapReduce computation, reduces task response time and improves the performance of MapReduce, but also reduces the computation cost and achieves a greener computing mode.

6.3.3. *Ensuring Privacy for MapReduce on Hybrid Clouds Using Information Dispersal Algorithm*

MapReduce is a powerful model for parallel data processing. The motivation of this work is to allow running map-reduce jobs partially on untrusted infrastructures, such as public Clouds and Desktop Grid, while using a trusted infrastructure, such as private cloud, to ensure that no outsider could get the 'entire' information. Our idea is to break data into meaningless chunks and spread them on a combination of public and private clouds so that the compromise would not allow the attacker to reconstruct the whole data-set. To realize this, we use the Information Dispersion Algorithms (IDA), which allows to split a file into pieces so that, by carefully

dispersing the pieces, there is no method for a single node to reconstruct the data if it cannot collaborate with other nodes. We propose a protocol that allows MapReduce computing nodes to exchange the data and perform IDA-aware MapReduce computation. We conduct experiments on the Grid'5000 testbed and report on performance evaluation of the prototype.

6.4. Using Active Data to Provide Smart Data Surveillance to E-Science Users

Participants: Gilles Fedak, Anthony Simonet.

Large scientific experiments drive scientists to use many storage and computing platforms as well as different applications, tools and analysis scripts. The resulting heterogeneous environments make data management operations challenging; the significant number of events and the absence of data integration makes it difficult to track data provenance, manage sophisticated analysis processes, and recover from unexpected situations. Current approaches often require costly human intervention and are inherently error prone. The difficulty managing and manipulating such large and highly distributed datasets also limits automated sharing and collaboration. In this collaboration with Kyle Chard and Ian Foster from Argonne National Lab and University of Chicago, we study a real world e-Science application involving terabytes of data, using three different analysis and storage platforms, and a number of applications and analysis processes. We demon- strate that using a specialized data life cycle and programming model—Active Data—we can easily implement global progress monitoring, sharing and recovery from unexpected events in heterogeneous environments and automate human tasks.

6.5. HPC Component Model

Participants: Hélène Coullon, Vincent Lanore, Christian Perez, Jérôme Richard.

6.5.1. 3D FFT and L^2C

We have studied the relevance of dealing with 3D FFT parallel algorithms with the software component model L^2C [31]. We have implemented several existing 3D FFT algorithms, and we have evaluated their performance, their scalability, and their reuse rate. Experiments made on clusters of Grid'5000 and on the Curie supercomputer up to 8192 cores show that L2C based 3D assemblies are scalable and have the same kind of performance than existing 3D libraries such as FFTW or 2DECOMP. This work confirms than components can be used for optimized HPC applications

6.5.2. Stencil Skeletons in L^2C

Mesh-based scientific simulation is an important class of scientific application which could benefit from component models. Therefore, we have studied and designed a first adaptation of the SIPSim model [33] (Structured Implicit Parallelism for scientific Simulations) to handle HPC component models. The heat equation application has been implemented on top of L^2C following this adapted SIPSim model. First experiments on clusters of Grid'5000 and on the Curie supercomputer show promising results, of which a complete analysis is still ongoing. This work is a first step toward a complete implicit parallelism stencil skeleton using L^2C.

6.5.3. Reconfigurable HPC component model

High-performance applications whose structure changes dynamically during execution are extremely complex and costly to develop, maintain and adapt to new hardware. Such applications would greatly benefit from easy reuse and separation of concerns which are typical advantages of component models. Unfortunately, no existing component model is both HPC-ready (in terms of scalability and overhead) and able to easily handle dynamic reconfiguration.

We aim at addressing performance, scalability and programmability by separating locking and synchronization concerns from reconfiguration code. To this end, we have defined *directMOD*, a component model which provides on one hand a flexible mechanism to lock subassemblies with a very small overhead and high scalability, and on the other hand a set of well-defined mechanisms to easily plug various independently-written reconfiguration components to lockable subassemblies. We evaluate both the model itself and a C++/MPI implementation called *directL2C* based on L^2C.

6.6. Security for Virtualization and Clouds

Participants: Eddy Caron, Arnaud Lefray, Jonathan Rouzaud-Cornabas.

Our framework Security Aware Models for Clouds has two purposes. The first one is, for a client, to model an IaaS application composed of virtual machines, applications, datas and communications and specify the associated security requirements. The whole modelization is contained into a XML file. The second one is the scheduling. It takes as inputs application models (XML) and the infrastructure of the cloud (currently in XML) i.e. a hierarchical set of physical machines. The scheduler encapsulates applications into virtual machines when needed and then maps virtual machines onto physical machines. The result of this scheduling is a file with the mapping i.e. a list of (VM, PM) couples.

The scheduler, as a standalone engine, can be used as simulator. But it can be interfaced with a Cloud stack (e.g. OpenStack, OpenNebula) to act as a production scheduler. This inferfacing is achieved by dynamically inferring the infrastructure model from the Cloud database and applying the decision i.e the output mapping list. Furthermore, the security policies (as input) are splitted for local security enforcement on each physical machine.

Sam4C (Security-Aware Models For Clouds) is a twofold framework, namely Sam4C-Modeler and Sam4C-Scheduler. The first is dedicated to modeling an application with the tenant's virtual machines and network interconnection. The second is is a security-aware scheduler, meaning it overrides the basic default scheduler with mainly the following enhanced capabilities

We have designed a scheduling module called SPS. This module is designed to support all the operations concerning the Cloud. It is based on the OpenStack and extends OpenStack with security aspects to fulfil the requirements of Seed4C.

6.7. Locality-aware Cooperation for VM Scheduling in Distributed Clouds

Participant: Frédéric Desprez.

In collaboration with the Ascola team (A. Lèbre, J. Pastor), ASAP team (Marin Bertier), and the Myriads team (C. Tedeschi), we worked on the design of a distributed Cloud Computing infrastructure [23]. The promotion of such infrastructures as the next platform to deliver the Utility Computing paradigm, leads to new virtual machines (VMs) scheduling algorithms leveraging peer-to-peer approaches. Although these proposals considerably improve the scalability, leading to the management of hundreds of thousands of VMs over thousands of physical machines (PMs), they do not consider the network overhead introduced by multi-site infrastructures. This overhead can have a dramatic impact on the performance if there is no mechanism favoring intra-site versus inter-site manipulations.

In 2014, we designed a new building block designed on top of a network with Vivaldi coordinates maximizing the locality criterion (*i.e.,* efficient collaborations between PMs) [12]. We combined such a mechanism with DVMS, a large-scale virtual machine scheduler and showed its benefit by discussing several experiments performed on four distinct sites of the Grid'5000 testbed. With our proposal and without changing the scheduling decision algorithm, the number of inter-site operations has been reduced by 72%. This result provides a glimpse of the promising future of using locality properties to improve the performance of massive distributed Cloud platforms.

CTRL-A Exploratory Action

6. New Results

6.1. Highlights of the Year

We have been invited to participate to the organization of events, which highlight our active presence in the scientific life in the two domains which we are bridging :

- autonomic computing: Eric Rutten is PC member, as well as workshops chair, of the 12th IEEE International Conference on Autonomic Computing, ICAC 2015 (http://icac2015.imag.fr/), and PC co-chair of the 3rd IEEE International Conference on Cloud and Autonomic Computing, CAC 2015 (http://autonomic-conference.org/), the two major conferences on the topic.
- control: Eric Rutten is organizer of a special session on discrete control for computing at the 12th IFAC - IEEE International Workshop on Discrete Event Systems, WODES 2014 (http://wodes2014.lurpa.ens-cachan.fr/), the main conference specialized in Discrete Event Systems, ; he is on the IFAC Technical Committee 1.3 on Discrete Event and Hybrid Systems, (http://tc.ifac-control.org/1/3/) and on the IEEE Control Systems Society Discrete Event Systems Technical Committee (http://discrete-event-systems.ieeecss.org).

6.2. Discrete control and reactive language support

Participants: Gwenaël Delaval, Eric Rutten, Stéphane Mocanu.

Concerning language support, we have designed and implemented BZR, a mixed imperative/declarative programming language: declarative contracts are enforced upon imperatively described behaviors (see 5.1). The semantics of the language uses the notion of Discrete Controller Synthesis (DCS) [5]. We target the application domain of adaptive and reconfigurable systems: our language can serve programming closed-loop adaptation controllers, enabling flexible execution of functionalities w.r.t. changing resource and environment conditions. DCS is integrated into a programming language compiler, which facilitates its use by users and programmers, performing executable code generation. The tool is concretely built upon the basis of a reactive programming language compiler, where the nodes describe behaviors that can be modeled in terms of transition systems. Our compiler integrates this with a DCS tool [3]. This work is done in close cooperation with the Inria team Sumo at Inria Rennes (H. Marchand). Ongoing work concerns aspects of compilation and debugging and logico-numeric extension of BZR based on the ReaX tool developed at Inria Rennes in the framework of the ANR Ctrl-Green project (see 8.2.1).

We are also currently working on on combining maximally permissive discrete control with runtime mechanisms for choosing between valid control values, involving e.g. a classical controller or stochastic aspects ; and on exploring the notion of adaptive discrete control, which is yet an open question in discrete control in contrast to the well-known adaptive continuous control.

Another activity related to discrete control is or work with Leiden University and CWI (N. Khakpour, now at Linnaeus U., and F. Arbab) on enforcing correctness of the behavior of an adaptive software system during dynamic adaptation is an important challenge along the way to realize correct adaptive systems. In this research, we model adaptation as a supervisory control problem and synthesize a controller that guides the behavior of a software system during adaptation. The system during adaptation is modeled using a graph transition system and properties to be enforced are specified using an automaton. To ensure correctness, we then synthesize a controller that imposes constraints on the system during adaptation [14].

6.3. Design and programming

6.3.1. *Component-based approaches*

Participants: Frederico Alvares, Eric Rutten.

Component-based architectures have shown to be very suited for self-adaptation purposes, not only because of their intrinsic characteristics like reusability and modularity, but also as virtue of their dynamical reconfiguration capabilities. The issue, nevertheless, remains that adaptation behaviors are generally conceived by means of fine-grained reconfiguration actions from the very initial configurations. This way, besides the complexity in managing large-sized architectures, the space of reachable configurations is not know in advance, which prevents ensuring well-mastered adaptive behaviours. We address this problem by designing Ctrl-F, a domain-specific language which objective is to provide high-level support for describing adaptation behaviors and policies in component-based architectures. The proposed language lies on synchronous reactive programming, which means that it benefits of an entire environment and formal tooling allowing for the verification and control of reconfigurations. We show the applicability of Ctrl-F by first integrating it to FraSCAti, a Service Component Architecture middleware platform, and then by applying it to Znn.com, a well known self-adaptive case study.

We work on the topic in cooperation with the Spirals Inria team at Inria Lille (L. Seinturier). It constitutes a follow-up on previous work in the ANR Minalogic project MIND, industrializing the Fractal component-based framework, with a continuation of contacts with ST Microelectronics (V. Bertin). Our integration of BZR and Fractal [4], [2] is at the basis of our current work. On a related topic, we are also starting a cooperation on introducing reactive control in hierarchical autonomic architectures, with A. Diaconescu and E. Najm at TelecomParisTech.

6.3.2. *Rule-based systems*
Participants: Julio Cano, Adja Sylla, Gwenaël Delaval, Eric Rutten.

Event-Condition-Action (ECA) rules are a widely used language for the high level specification of controllers in adaptive systems, such as Cyber-Physical Systems and smart environments, where devices equipped with sensors and actuators are controlled according to a set of rules. The evaluation and execution of every ECA rule is considered to be independent from the others, but interactions of rule actions can cause the system behaviors to be unpredictable or unsafe. Typical problems are in redundancy of rules, inconsistencies, circularity, or application- dependent safety issues. Hence, there is a need for coordination of ECA rule-based systems in order to ensure safety objectives. We propose a tool-supported method for verifying and controlling the correct interactions of rules, relying on formal models related to reactive systems, and Discrete Controller Synthesis (DCS) to generate correct rule controllers [12].

We work on this topic in cooperation with CEA LETI/DACLE (L. Gurgen) and target the application and experimentation domain of smart environment in the Internet of Things [11].

Another complementary direction on which we are starting a cooperation with CEA LETI/DACLE is the topic of a high-level language for safe rule-based programming in the LINC platform: the PhD of Adja Sylla on this topic will be co-advised with F. Pacull and M. Louvel at CEA.

6.4. Infrastructure-level support

6.4.1. *Autonomic Cloud and Big-Data systems*

This activity continues work started several years ago in the Sardes Inria-team, before it split into Erods (at LIG) and Ctrl-A (at Inria).

6.4.1.1. *Coordination in multiple-loop autonomic Cloud systems*
Participants: Soguy Gueye, Gwenaël Delaval, Stéphane Mocanu, Bogdan Robu, Eric Rutten.

Complex computing systems are increasingly self-adaptive, with an autonomic computing approach for their administration. Real systems require the co-existence of multiple autonomic management loops, each complex to design. However their uncoordinated co-existence leads to performance degradation and possibly to inconsistency. There is a need for methodological supports facilitating the coordination of multiple autonomic managers. We address this problem in the context of the ANR project Ctrl-Green (see 8.2.1), in cooperation with LIG (N. de Palma) in the framework of the PhD of S. Gueye. We propose a method focusing on the

discrete control of the interactions of managers [7] [9]. We follow a component-based approach and explore modular discrete control, allowing to break down the combinatorial complexity inherent to the state-space exploration technique [13]. This improves scalability of the approach and allows constructing a hierarchical control. It also allows re-using complex managers in different contexts without modifying their control specifications. We build a component-based coordination of managers, with introspection, adaptivity and reconfiguration. We validate our method on a multiple-loop multi-tier system.

We are currently working on the distributed execution of modular controllers and on considering more control objectives, beyond purely discrete or logical ones, evaluating the new tool ReaX developed at Inria Rennes (Sumo) (see 6.2) and exploring continuous or stochastic control of servers provisioning.

6.4.1.2. Control for Big data
Participants: Bogdan Robu, Mihaly Berekmeri, Nicolas Marchand.

To deal with the issue of ensuring performance constraints while also minimizing costs in systems for Big Data analytics based on the parallel programming paradigm MapReduce, we propose a control theoretical approach, based on techniques that have already proved their usefulness for the control community. We develop an algorithm to create the first linear dynamic model for a Big Data MapReduce system, running a concurrent workload. Furthermore we identify two major performance constraint use cases: relaxed-minimal resource and strict performance constraints. For the first case we developed a feedback control mechanism and, to minimize the number of control actuations, an event-based feedback controller. For the second case we add a feedforward controller that efficiently suppresses the effects of large workload size variations. The work is validated in a simulated Matlab environment build at GIPSA-lab and online on a real 60 node MapReduce cluster (part of GRID 500), running a data intensive Business Intelligence workload. Our experiments demonstrate the success of the control strategies employed in assuring service time constraints [17], [18].

This work is performed in cooperation with LIG (S. Bouchenak) in the framework of the PhD of M. Berekmeri.

6.4.2. *Reconfiguration control in DPR FPGA*
Participant: Eric Rutten.

Dynamically reconfigurable hardware has been identified as a promising solution for the design of energy efficient embedded systems. However, its adoption is limited by the costly design effort including verification and validation, which is even more complex than for non dynamically reconfigurable systems. We work on this topic in the context of a ensign environment, developed in the framework of the ANR project Famous, in cooperation with LabSticc in Lorient and Inria Lille (DaRT team) [10]. We propose a tool-supported formal method to automatically design a correct-by-construction control of the reconfiguration. By representing system behaviors with automata, we exploit automated algorithms to synthesize controllers that safely enforce reconfiguration strategies formulated as properties to be satisfied by control. We design generic modeling patterns for a class of reconfigurable architectures, taking into account both hardware architecture and applications, as well as relevant control objectives. We validate our approach on two case studies implemented on FPGAs [1].

We are currently valorizing results in more publications, and extending the use of control techniques by evaluating the new tool ReaX developed at Inria Rennes (Sumo) in the framework of the ANR Ctrl-Green project (see 6.2 and 8.2.1).

6.4.3. *Autonomic memory management in HPC*
Participants: Naweiluo Zhou, Gwenaël Delaval, Bogdan Robu, Eric Rutten.

Concurrent programs need to manage the time trade-off between synchronization and computing. A high concurrency level may decrease computing time but at the same time increase synchronization cost among threads. The traditional way to handle synchronization problems is through implementing locks. However locks suffer from the likelihood of deadlocks, vulnerability to failures, faults etc.. Software Transactional Memory (STM) has emerged as a promising technique to address synchronization issues through transactions. In STM, blocks of instructions accessing the shared data are wrapped into transactions. In STM each

transaction executes speculatively, and conflicts may be aroused when two transactions are trying to modify the same area simultaneously. A way to reduce conflicts is by adjusting concurrency levels. A suitable concurrency level can maximize program performance. However, there is no universal rule to decide the best concurrency level for a program from an offline view. Hence, it becomes necessary to adopt a dynamical tuning strategy to better manage a STM system, so that a program can achieve a better performance. In the context of the action-team HPES of the Labex Persyval-lab[0] (see 8.1), we explore the autonomic computing approach and control techniques to address these runtime tuning problems as a feedback control loop to automate the choices of concurrency levels, conflict management policies, and other parameters, with the objective of optimizing program execution time. This work is performed in cooperation with LIG (J.F. Méhaut) in the framework of the PhD of N. Zhou.

6.4.4. *Control of smart environments*

Participants: Julio Angel Cano Romero, Mengxuan Zhao, Eric Rutten, Hassane Alla [Gipsa-lab].

6.4.4.1. *Generic supervision architecture*

New application domains of control, such as in the Internet of Things (IoT) and Smart Environments, require generic control rules enabling the systematization and the automation of the controller synthesis. We are working on an approach for the generation of Discrete Supervisory Controllers for these applications. A general modeling framework is proposed for the application domain of smart home. We formalize the design of the environment manager as a Discrete Controller Synthesis (DCS) problem, w.r.t. multiple constraints and objectives, for example logical issues of mutual exclusion, bounding of power peaks. We validate our models and manager computations with the BZR language and an experimental simulator [15]. This work is performed in cooperation with Orange labs (G. Privat) in the framework of the Cifre PhD of M. Zhao.

6.4.4.2. *Rule-based specification*

In the Internet of things, Event - Condition - Action (ECA) are used as a flexible tool to govern the relations between sensors and actuators. Runtime coordination and formal analysis becomes a necessity to avoid side effects mainly when applications are critical. In cooperation with CEA LETI/DACLE, we have worked on a case study for safe applications development in IoT and smart home environments [11].

---

[0]https://persyval-lab.org/en/sites/hpes

DANTE Team

6. New Results

6.1. Highlights of the Year

6.1.1. *The Internet of Things: A new equipments of excellence*

Inaugurated last autumn, the very large scale IoT-LAB platform (https://www.iot-lab.info) is strengthening the capabilities of the FIT equipment of excellence dedicated to the Internet of Things. Offering a unique wide-ranging collection of equipment, these laboratories are available to both researchers and commercial companies alike.

IoT-LAB is a large-scale experimental platform for communicating objects and networks of sensors. It enables the rapid deployment of experiments and the collection of large amounts of data. It includes over 2700 sensor nodes, distributed over six sites in France, offering a wide range of different processor architectures and radio components. IoT-LAB is available for use on line. It is already used by over 300 users in forty countries, including around ten commercial companies. As of the end of October 2014, some 10 000 experiments had already been carried out.

6.1.2. *Graph-based signal processing*

Our first results towards the definition of a digital framework for signal processing on graphs constitutes an important outcome of DANTE's activity in 2014. Our participation to this emerging discipline was marked with several scientific recognitions: publication in the main DSP conference [14], involvement in the first ANR project focusing on this theme and retained for funding (2015-2019), we are in charge of the organisation of a Special Session dedicated to "Methodologies for signal processing on graphs" at Eusipco conference (2015).

6.1.3. *Complex contagion process*

Diffusion of innovation can be interpreted as a social spreading phenomena governed by the impact of media and social interactions. Although these mechanisms have been identified by quantitative theories, their role and relative importance are not entirely understood, since empirical verification has so far been hindered by the lack of appropriate data. Here we analyse a dataset recording the spreading dynamics of the world's largest Voice over Internet Protocol service to empirically support the assumptions behind models of social contagion. We show that the rate of spontaneous service adoption is constant, the probability of adoption via social influence is linearly proportional to the fraction of adopting neighbors, and the rate of service termination is time-invariant and independent of the behavior of peers. By implementing the detected diffusion mechanisms into a dynamical agent-based model, we are able to emulate the adoption dynamics of the service in several countries worldwide. This approach enables us to make medium-term predictions of service adoption and disclose dependencies between the dynamics of innovation spreading and the socioeconomic development of a country. This work was recently published in the Journal of the Royal Society Interface.

6.2. Diffusion and dynamic of complex networks

Participants: Márton Karsai [correspondant], Éric Fleury, Christophe Crespelle.

Time varying networks and the weakness of strong ties   We analyse a mobile call dataset and find a simple statistical law that characterize the temporal evolution of users' egocentric networks. We encode this observation in a reinforcement process defining a time-varying network model that exhibits the emergence of strong and weak ties. We study the effect of time-varying and heterogeneous interactions on the classic rumor spreading model in both synthetic, and real-world networks. We observe that strong ties severely inhibit information diffusion by confining the spreading process among agents with recurrent communication patterns. This provides the counterintuitive evidence that strong ties may have a negative role in the spreading of information across networks.

Complex contagion process in spreading of online innovation [8].   Here we analyse a dataset recording the spreading dynamics of the world's largest Voice over Internet Protocol service to empirically support the assumptions behind models of social contagion. We show that the rate of spontaneous service adoption is constant, the probability of adoption via social influence is linearly proportional to the fraction of adopting neighbors, and the rate of service termination is time-invariant and independent of the behavior of peers. By implementing the detected diffusion mechanisms into a dynamical agent-based model, we are able to emulate the adoption dynamics of the service in several countries worldwide. This approach enables us to make medium-term predictions of service adoption and disclose dependencies between the dynamics of innovation spreading and the socio-economic development of a country.

The role of endogenous and exogenous mechanisms in the formation of R&D networks [10].   Here   we propose a general modeling framework that includes both endogenous and exogenous mechanisms of in link formations in networks with tunable relative importance. The model contains additional ingredients derived from empirical observations, such as the heterogeneous propensity to form alliances and the presence of circles of influence, i.e. clusters of firms in the network. We test our model against the Thomson Reuters SDC Platinum dataset, one of the most complete datasets available nowadays, listing cross-country R&D alliances from 1984 to 2009. Interestingly, by fitting only three macroscopic properties of the network, this framework is able to reproduce a number of microscopic measures characterizing the network topology, including the distributions of degree, local clustering, path length and component size, and the emergence of network clusters. Furthermore, by estimating the link probabilities towards newcomers and established firms from the available data, we find that endogenous mechanisms are predominant over the exogenous ones in the network formation. This quantifies the importance of existing network structures in selecting partners for R&D alliances.

Controlling Contagion Processes in Time-Varying Networks [9].   In this project we derive an analytical framework for the study of control strategies specifically devised for time-varying networks. We consider the removal/immunization of individual nodes according the their activity in the network and develop a block variable mean-field approach that allows the derivation of the equations describing the evolution of the contagion process concurrently to the network dynamic. We derive the critical immunization threshold and assess the effectiveness of the control strategies. Finally, we validate the theoretical picture by simulating numerically the information spreading process and control strategies in both synthetic networks and a large-scale, real-world mobile telephone call dataset.

Data-driven spreading for the detection of weak ties [24].   In this work we propose a new method to infer the strength of social ties by using new data-driven simulation techniques. We qualify links by the importance they play during the propagation of information in the social structure. We apply data-driven spreading processes combined with a river-basin algorithmic method to identify links, which are the responsible to bring the information to large number of nodes. We investigate the correlations of the new importance measure with other conventional characteristics and identify their best combination through a percolation analysis to sophisticate further the assignment of social tie strengths. Finally we explore the role of the identified high importance links in control of globally spreading processes through data-driven SIR model simulations. These results point out that the size of infected population can be reduced considerably by weakening interactions through ties with high importance but zero overlap compared to strategies based on dyadic communications.

Dynamic Contact Network Analysis in Hospital Wards [18].   We analyse a huge and very precise trace of contact data collected during 6 months on the entire population of a rehabilitation hospital. We investigate the graph structure of the average daily contact network. Our main results are to unveil striking properties of this structure in the considered hospital, and to present a methodology that can be used for analyzing any dynamic complex network where nodes are classified into groups.

6.3. Performance analysis and networks protocols

Participants: Anthony Busson [correspondant], Thomas Begin, Isabelle Guérin Lassous.

Modeling and optimization of CSMA/CA in VANET [7]. We propose a simple theoretical model to compute the maximum spatial reuse feasible in a VANET. We focus on the ad hoc mode of the IEEE 802.11p standard. Our model offers simple and closed-form formulas on the maximum number of simultaneous transmitters, and on the distribution of the distance between them. It leads to an accurate upper bound on the maximum capacity. In order to validate our approach, results from the analytical models are compared to simulations performed with the network simulator NS-3. We take into account different traffic distributions (traffic of vehicles), and study the impact of this traffic on capacity. An application of this work is the parameterization of the CSMA/CA mechanism.

Fast and accurate approximate performance analysis of multi-server facilities [4]. Systems with multiple servers are common in many areas and their correct dimensioning is in general a difficult problem under realistic assumptions on the pattern of user arrivals and service time distribution. We present an approximate solution for the underlying $Ph/Ph/c/N$ queueing model. Our approximation decomposes the solution of the Ph/Ph/c/N queue into solutions of simpler $M/Ph/c/N$ and $Ph/M/c/N$ queues. To further mitigate dimensionality issues, for larger numbers of servers and/or service time phases, we use a reduced state approximation to solve the $M/Ph/c/N$ queue. The proposed approach is conceptually simple, easy to implement and produces generally accurate results for the mean number in the system, as well as the loss probability. Typical relative errors for these two quantities are below 5%. A very significant speed advantage compared to the numerical solution of the full $Ph/Ph/c/N$ queue can be gained as the number of phases representing the arrival process and/or the number of servers increases.

Interference and throughput in spectrum sensing cognitive radio networks using point processes .
Spectrum sensing is vital for secondary unlicensed nodes to coexist and avoid interference with the primary licensed users in cognitive wireless networks. In this paper, we develop models for bounding interference levels from secondary network to the primary nodes within a spectrum sensing framework. Instead of classical stochastic approaches where Poisson point processes are used to model transmitters, we consider a more practical model which takes into account the medium access control regulations and where the secondary Poisson process is judiciously thinned in two phases to avoid interference with the secondary as well as the primary nodes. The resulting process will be a modified version of the Maté´rn point process. For this model, we obtain bounds for the complementary cumulative distribution function of interference and present simulation results which show the developed analytical bounds are quite tight. Moreover, we use these bounds to find the operation regions of the secondary network such that the interference constraint is satisfied on receiving primary nodes. We then obtain theoretical results on the primary and secondary throughputs and find the throughput limits under the interference constraint.

Modeling of IEEE 802.11 Multi-hop Wireless Chains with Hidden Nodes [11]. We follow up an existing modeling framework to analytically evaluate the performance of multi-hop flows along a wireless chain of four nodes. The proposed model accounts for a non-perfect physical layer, handles the hidden node problem, and is applicable under workload conditions ranging from flow(s) with low intensity to flow(s) causing the network to saturate. Its solution is easily and quickly obtained and delivers estimates for the expected throughput and for the datagram loss probability of the chain with a good accuracy.

Anticipation of ETX Metric to manage Mobility in Ad Hoc Wireless Networks [19]. When a node is moving in a wireless network, the routing metrics associated to its wireless links may reflect link quality degradations and help the routing process to adapt its routes. Unfortunately, an important delay between the metric estimation and its inclusion in the routing process makes this approach inefficient. In this paper, we introduce an algorithm that predicts metric values a few seconds in advance, in order to compensate the delay involved by the link quality measurement and their dissemination by the routing protocol. We consider classical metrics, in particular ETX (Expected Transmission Count) and ETT (Expected Transmission Time), but we combine their computations

to our prediction algorithm. Extensive simulations show the route enhancement as the Packet Delivery Ratio (PDR) is close to 1 in presence of mobility.

6.4. Graphs & Signal Processing

Participants: Paulo Gonçalves [correspondant], Éric Fleury, Christophe Crespelle.

6.4.1. Signal Processing on Graphs

Semi-Supervised Learning for Graph to Signal Mapping: a Graph Signal Wiener Filter Interpretation [14]. We investigate a graph to signal mapping with the objective of analyzing intricate structural properties of graphs with tools borrowed from signal processing. We successfully use a graph-based semi-supervised learning approach to map nodes of a graph to signal amplitudes such that the resulting time series is smooth and the procedure efficient and scalable. Theoretical analysis of this method reveals that it essentially amounts to a linear graph-shift-invariant filter with the a priori knowledge put into the training set as input. Further analysis shows that we can interpret this filter as a Wiener filter on graphs. We finally build upon this interpretation to improve our results.

6.4.2. Graphs

(Nearly-)tight bounds on the contiguity and linearity of cographs [6]. In this paper we show that the contiguity and linearity of cographs on n vertices are both $O(\log n)$. Moreover, we show that this bound is tight for contiguity as there exists a family of cographs on n vertices whose contiguity is $\Omega(\log n)$. We also provide an $\Omega(\log n / \log \log n)$ lower bound on the maximum linearity of cographs on n vertices. As a by-product of our proofs, we obtain a min-max theorem, which is worth of interest in itself, stating equality between the rank of a tree and the minimum height of one of its path partitions.

6.4.3. Signal processing

Analysis of intrapartum foetal heart rate (FHR), enabling early detection of foetal acidosis to prevent asphyxia and labour adverse outcomes, remains a challenging signal processing task. In this direction, we carried out a series of works to characterize the fetal heart rate variability with specific attributes able to discriminate between healthy fetuses and fetuses presenting a risk of brain injury. Last year, we investigated two different approaches:

Nearest-Neighbor based Wavelet Entropy Rate Measures for Intrapartum Fetal Heart Rate Variability [23]. Firstly, we showed that a k-nearest neighbor procedure yields estimates for entropy rates that are robust and well-suited to FHR variability. Secondly, we experimentally proved that entropy rates measured on multiresolution wavelet coefficients permit to improve classification performance.

Impacts of labour first and second stages on Hurst parameter based intrapartum FHR analysis [22].   In this study, we proposed to quantify the FHR temporal dynamics with a Hurst exponent estimated within a wavelet framework. Analyses performed over a large (3049 records) and well documented database revealed that the evolution of the Hurst exponent during delivery, is significantly different for healthy fetuses and for acidotic fetuses.

6.5. Complex network metrology

Participant: Christophe Crespelle.

Measuring the Degree Distribution of Routers in the Core Internet [15].   Most current models of the internet rely on knowledge of the degree distribution of its core routers, which plays a key role for simulation purposes. In practice, this distribution is usually observed directly on maps known to be partial, biased and erroneous. This raises serious concerns on the true knowledge one may have of this key property. Here, we design an original measurement approach targeting reliable estimation of the degree distribution of core routers, without resorting to any map. It consists in sampling random core routers and precisely estimate their degree thanks to probes sent from many distributed monitors. We run and assess a large-scale measurement following this approach, carefully controlling and

correcting bias and errors encountered in practice. The estimate we obtain is much more reliable than previous knowledge, and it shows that the true degree distribution is very different from all current assumptions.

Measuring Routing Tables in the Internet [21].   The most basic function of an Internet router is to decide, for a given packet, which of its interfaces it will use to forward it to its next hop. To do so, routers maintain a routing table, in which they look up for a prefix of the destination address. The routing table associates an interface of the router to this prefix, and this interface is used to forward the packet. We explore here a new measurement method based upon distributed UDP probing to estimate this routing table for Internet routers.

<p style="text-align:center;color:red;">**MESCAL Project-Team**</p>

6. New Results

6.1. Simulation of distributed architectures

- Simgrid is a toolkit providing core functionalities for the simulation of distributed applications in heterogeneous distributed environments. It models fine-grain detail of the studied platform. In [6], we present quantitative results that show that SimGrid compares favorably to state-of-the-art domain-specific simulators in terms of scalability, accuracy, or the trade-off between the two. In [37], [29], we develop an hybrid approach of simulation and emulation of applications that use starPU. By using this approach, Simgrid calibrates the time to run specific subtasks at runtime and simulates all system calls of the application. This approach allows us to obtain performance results that are within one percent of measured results.

- In [33], [18], we study the problem of sampling the stationary distribution of a random walker in $\{0 \cdots N\}^d$ using simulation. This algorithm combines the rejection method and coupling from the past of a set of trajectories of the Markov chain that generalizes the classical sandwich approach. We also provide a complexity analysis of this approach in several cases showing a coupling time in $O(N^2 d \log d)$ when no arc is forbidden and an experimental study of its performance.

6.2. Interactive Analysis and Visualization of Large Distributed Systems

- In [13], we review the methodology that we use to visualize information for large-scale data-set. Our approach uses tools from information theory to define a trade-off between the loss of information and the compactness of the representation. This methodology is applied to spatio-temporal representation of traces of execution in [30], [16], [17], [32]. In these papers, we show how to build a concise overview of the trace behavior as the result of a spatio-temporal data aggregation process. The experimental results show that this approach can help the quick and accurate detection of anomalies in traces containing up to two hundred million events.

- Trace analysis graphical user environments have to provide different views on trace data, to really help provide insights on the traced application behavior. In [22], [35], we propose an open and modular software architecture, the FrameSoC workbench, that defines clear principles for view engineering and for view consistency management. The FrameSoC workbench has been successfully applied in real trace analysis use-cases. This work has also been tested on real scenario coming from a collaboration with ST Microelectronic [25].

- In [7], we design a novel prediction method with Bayes model to predict a load fluctuation pattern over a long-term interval, in the context of Google data centers. All of the prediction methods are evaluated using Google trace with 10,000+ heterogeneous hosts. Experiments show that our Bayes method improves the long-term load prediction accuracy by up 5 to 50%, compared to other state-of-the-art methods.

6.3. Management of Parallel Architectures

- In [12], we present a topology-aware load balancing algorithm for parallel multi-core machines and its proof of asymptotic convergence to an optimal solution. The algorithm, named HwTopoLB, takes into account the properties of current parallel systems composed of multi-core compute nodes, namely their network interconnection, and their complex and hierarchical core topology. We have implemented HwTopoLB using the Charm++ Parallel Runtime System and evaluated its performance with two different benchmarks and one application. Our experimental results confirms that HwTopoLB outperform existing load balancing strategies on different multi-core systems.

- Large scale distributed systems typically comprise hundreds to millions of entities that have only a partial view of resources. How to fairly and efficiently share such resources between entities in a distributed way has thus become a critical question. In [31], we develop a possible answer based on Lagrangian optimization and distributed gradient descent. Under certain conditions, the resource sharing problem can be formulated as a global optimization problem, which can be solved by a distributed self-stabilizing demand and response algorithm.

- The management of resources on testbeds, including their description, reservation and verification, is a challenging issue, especially on of large scale testbeds such as those used for research on High Performance Computing or Clouds. In [23], we present the solution designed for the Grid'5000 testbed in order to: (1) provide users with an in-depth and machine-parsable description of the testbed's resources; (2) enable multi-criteria selection and reservation of resources using a HPC resource manager; (3) ensure that the description of the resources remains accurate. In [24], we present Kascade, a solution for the broadcast of data to a large set of compute nodes. We evaluate Kascade using a set of large scale experiments in a variety of experimental settings, and show that Kascade: (1) achieves very high scalability by organizing nodes in a pipeline; (2) can almost saturate a 1 Gbit/s network, even at large scale; (3) handles failures of nodes during the transfer seamlessly because of its fault-tolerant design.

6.4. Reproducible experiments and papers

- In the field of large-scale distributed systems, experimentation is particularly difficult. The studied systems are complex, often nondeterministic and unreliable, software is plagued with bugs, whereas the experiment workflows are unclear and hard to reproduce. In [5], we provide an extensive list of features offered by general-purpose experiment management tools dedicated to distributed systems research on real platforms. We then use it to assess existing solutions and compare them, outlining possible future paths for improvements.

- Experiment reproducibility is a milestone of the scientific method. Reproducibility of experiments in computer science would bring several advantages such as code re-usability and technology transfer. The reproducibility problem in computer science has been solved partially, addressing particular class of applications or single machine setups. In [26], we present our approach oriented to setup complex environments for experimentation, environments that require a lot of configuration and the installation of several software packages. The main objective of our approach is to enable the exact and independent reconstruction of a given software environment and the reuse of code. We present a simple and small software appliance generator that helps an experimenter to construct a specific software stack that can be deployed on different available testbeds. [14],

- In [28], [45], we address the question of developing a lightweight and effective workflow for conducting experimental research on modern parallel computer systems in a reproducible way. Our workflow simply builds on two well-known tools (Org-mode and Git) and enables us to address issues such as provenance tracking, experimental setup reconstruction, replicable analysis. Although this workflow is perfectible and cannot be seen as a final solution, we have been using git for two years now and we have recently published a fully reproducible article, which demonstrates the effectiveness of our proposal.

6.5. Game Theory and Distributed Optimization

- In wireless networks, channel conditions of and user quality of service (QoS) requirements vary, often quite arbitrarily, with time (e.g. due to user mobility, fading, etc.) In this dynamic setting, static solution concepts (such as Nash equilibrium) are no longer relevant. Hence, we focus on the concept of no-regret : policies that perform at least as well as the best fixed transmit profile in hindsight. In [21], we examine the performance of the seminal Foschini–Miljanic (FM) power control scheme in a random environment. We provide a formulation of power control as an online optimization problem and we show that the FM dynamics lead to no regret in this dynamic context. We introduce

an adjusted version of the FM algorithm which retains the convergence and no-regret properties of the original algorithm in this constrained setting. In [27], we examine the problem of cost / energy-efficient power allocation in uplink multi-carrier orthogonal frequency-division multiple access wireless networks. We use tools from stochastic convex programming to develop a learning scheme that retains its convergence properties irrespective of the magnitude of the observational errors. In [11], we consider a cognitive radio network where wireless users with multiple antennas communicate over several non-interfering frequency bands. We draw on the method of matrix exponential learning and online mirror descent techniques to derive a no-regret policy that relies only on local channel state information.

- In game theory, the best-response strategy of a player is a strategy that maximizes the selfish payoff of this player. A natural and popular question is, when players update their strategy over time, do they converge to a Nash equilibrium. In [15], we characterize the revision sets in different variants of the best response algorithm that guarantee convergence to pure Nash Equilibria in potential games. We prove that if the revision protocol is separable, then the greedy version as well as smoothed versions of the algorithm converge to pure Nash equilibria. If the revision protocol is not separable, then convergence to Nash Equilibria may fail in both cases. In [43], we investigate a class of reinforcement learning dynamics in which each player plays a "regularized best response" to a score vector consisting of his actions' cumulative payoffs. Our main results extend several properties of the replicator dynamics such as the elimination of dominated strategies, the asymptotic stability of strict Nash equilibria and the convergence of time-averaged trajectories to interior Nash equilibria in zero-sum games.

6.6. Agent-based modeling and applications to Smart Energy and Transportation Systems

- Renewable energy sources, such as wind, are characterized by non-dispatchability, high volatility, and non-perfect forecasts. Energy storage or electric loads that have a flexible consumption are viewed as a way to mitigate these effects. In [9], [19], we study centralized and distributed algorithms for solving this problem. We provide theoretical bounds on the trade-off between energy loss and the use of reserves. We develop a centralized algorithm that attains this bound in [9]. In [19], we study a distributed optimization problem by modeling a two-stage electricity market. We show that the market is efficient: the players' selfish responses to prices coincide with a socially optimal policy. We develop a distributed solution technique based on the Alternating Direction Method of Multipliers (ADMM) and trajectorial forecasts to compute the Nash-equilibrium.

- Bike-sharing systems are becoming important for urban transportation. In these systems, users arrive at a station, pick up a bike, use it for a while, and then return it to another station of their choice. In [8], we propose a stochastic model of an homogeneous bike-sharing system and study the effect of the randomness of user choices on the number of problematic stations. Even in a homogeneous city, the system exhibits a poor performance: the minimal proportion of problematic stations is of the order of the inverse of the capacity. We show that simple incentives, such as suggesting users to return to the least loaded station among two stations, improve the situation by an exponential factor.

- In [10], we discuss the validation of an agent-based model of emergent city systems with heterogeneous agents. We transform our model into an analytically tractable discrete Markov model, and we examine the city size distribution. We show that the Markov chains lead to a power-law distribution when the ranges of migration options are randomly distributed across the agent population. We also identify sufficient conditions under which the Markov chains produce the Zipf's Law, which has never been done within a discrete framework. The conditions under which our simplified model yields the Zipf's Law are in agreement with, and thus validate, the configurations of the original heterogeneous agent-based model.

MOAIS Project-Team

5. New Results

5.1. Scheduling semi-malleable jobs to minimize mean flow time

This paper [9] deals with the problem of scheduling n_A malleable and n_B non-malleable jobs to be executed together on two parallel identical machines to minimize mean flow time. We propose a set of dominant schedules for this problem, and a dynamic programming algorithm that finds an optimal schedule in this dominant set in time $O(n_A^2 n_B)$.

5.2. Elements of Design for Containers and Solutions in the LinBox Library

We describe in this paper [12] new design techniques used in the exact linear algebra library LinBox, intended to make the library safer and easier to use, while keeping it generic and efficient. First, we review the new simplified structure for containers, based on our *founding scope allocation* model. We explain design choices and their impact on coding: unification of our matrix classes, clearer model for matrices and submatrices,... Then we present a variation of the *strategy* design pattern that is comprised of a controller-plugin system: the controller (solution) chooses among plug-ins (algorithms) that always call back the controllers for subtasks. We give examples using the solution `mul`. Finally we present a benchmark architecture that serves two purposes: Providing the user with easier ways to produce graphs; Creating a framework for automatically tuning the library and supporting regression testing.

5.3. Scheduling Data Flow Program in XKaapi: A New Affinity Based Algorithm for Heterogeneous Architectures

Efficient implementations of parallel applications on heterogeneous hybrid architectures require a careful balance between computations and communications with accelerator devices. Even if most of the communication time can be overlapped by computations, it is essential to reduce the total volume of communicated data. The literature therefore abounds with ad hoc methods to reach that balance, but these are architecture and application dependent. We propose [12] here a generic mechanism to automatically optimize the scheduling between CPUs and GPUs, and compare two strategies within this mechanism: the classical Heterogeneous Earliest Finish Time (HEFT) algorithm and our new, parametrized, Distributed Affinity Dual Approximation algorithm (DADA), which consists in grouping the tasks by affinity before running a fast dual approximation. We ran experiments on a heterogeneous parallel machine with twelve CPU cores and eight NVIDIA Fermi GPUs. Three standard dense linear algebra kernels from the PLASMA library have been ported on top of the XKaapi runtime system. We report their performances. It results that HEFT and DADA perform well for various experimental conditions, but that DADA performs better for larger systems and number of GPUs, and, in most cases, generates much lower data transfers than HEFT to achieve the same performance.

5.4. Evaluation of OpenMP Dependent Tasks with the KASTORS Benchmark Suite

The recent introduction of task dependencies in the OpenMP specification provides new ways of synchronizing tasks. Application programmers can now describe the data a task will read as input and write as output, letting the runtime system resolve fine-grain dependencies between tasks to decide which task should execute next. Such an approach should scale better than the excessive global synchronization found in most OpenMP 3.0 applications. As promising as it looks however, any new feature needs proper evaluation to encourage application programmers to embrace it. This paper [26] introduces the KASTORS benchmark suite designed to evaluate OpenMP tasks dependencies. We modified state-of-the-art OpenMP 3.0 benchmarks and data-flow parallel linear algebra kernels to make use of tasks dependencies. Learning from this experience, we propose extensions to the current OpenMP specification to improve the expressiveness of dependencies. We eventually evaluate both the GCC/libGOMP and the CLANG/libIOMP implementations of OpenMP 4.0 on our KASTORS suite, demonstrating the interest of task dependencies compared to taskwait-based approaches.

5.5. Sparse Polynomial Interpolation Codes and their decoding beyond half the minimal distance

We present [21] algorithms performing sparse univariate polynomial interpolation with errors in the evaluations of the polynomial. Based on the initial work by Comer, Kaltofen and Pernet [Proc. ISSAC 2012], we define the sparse polynomial interpolation codes and state that their minimal distance is precisely the length divided by twice the sparsity. At ISSAC 2012, we have given a decoding algorithm for as much as half the minimal distance and a list decoding algorithm up to the minimal distance. Our new polynomial-time list decoding algorithm uses sub-sequences of the received evaluations indexed by a linear progression, allowing the decoding for a larger radius, that is, more errors in the evaluations while returning a list of candidate sparse polynomials. We quantify this improvement for all typically small values of number of terms and number of errors, and provide a worst case asymptotic analysis of this improvement. For instance, for sparsity T = 5 with up to 10 errors we can list decode in polynomial-time from 74 values of the polynomial with unknown terms, whereas our earlier algorithm required 2T (E + 1) = 110 evaluations. We then propose two variations of these codes in characteristic zero, where appropriate choices of values for the variable yield a much larger minimal distance: the length minus twice the sparsity.

5.6. A Spatiotemporal Data Aggregation Technique for Performance Analysis of Large-scale Execution Traces

Analysts commonly use execution traces collected at runtime to understand the behavior of an application running on distributed and parallel systems. These traces are inspected post mortem using various visualization techniques that, however, do not scale properly for a large number of events. This issue, mainly due to human perception limitations, is also the result of bounded screen resolutions preventing the proper drawing of many graphical objects. This paper [21] proposes a new visualization technique overcoming such limitations by providing a concise overview of the trace behavior as the result of a spatiotemporal data aggregation process. The experimental results show that this approach can help the quick and accurate detection of anomalies in traces containing up to two hundred million events.

5.7. Scheduling independent tasks on multi-cores with GPU accelerators

More and more computers use hybrid architectures combining multi-core processors and hardware accelerators like GPUs (Graphics Processing Units). We present in this paper [3] a new method for scheduling efficiently parallel applications with m CPUs and k GPUs, where each task of the application can be processed either on a core (CPU) or on a GPU. The objective is to minimize the maximum completion time (makespan). The corresponding scheduling problem is NP-hard, we propose an efficient approximation algorithm which achieves an approximation ratio of $4/3 + 1/3k$. We first detail and analyze the method, based on a dual approximation scheme, that uses dynamic programming to balance evenly the load between the heterogeneous resources. Then, we present a faster approximation algorithm for a special case of the previous problem, where all the tasks are accelerated when affected to GPU, with a performance guarantee of $3/2$ for any number of GPUs. We run some simulations based on realistic benchmarks and compare the solutions obtained by a relaxed version of the generic method to the one provided by a classical scheduling algorithm (HEFT). Finally, we present an implementation of the $4/3$-approximation and its relaxed version on a classical linear algebra kernel into the scheduler of the XKaapi runtime system.

5.8. A Flexible Framework for Asynchronous In Situ and In Transit Analytics for Scientific Simulations

High performance computing systems are today composed of tens of thousands of processors and deep memory hierarchies. The next generation of machines will further increase the unbalance between I/O capabilities and processing power. To reduce the pressure on I/Os, the in situ analytics paradigm proposes to process the data as closely as possible to where and when the data are produced. Processing can be embedded

in the simulation code, executed asynchronously on helper cores on the same nodes, or performed in transit on staging nodes dedicated to analytics. Today, software environments as well as usage scenarios still need to be investigated before in situ analytics become a standard practice. In this paper [3] we introduce a framework for designing, deploying and executing in situ scenarios. Based on a component model, the scientist designs analytics workflows by first developing processing components that are next assembled in a dataflow graph through a Python script. At runtime the graph is instantiated according to the execution context, the framework taking care of deploying the application on the target architecture and coordinating the analytics workflows with the simulation execution. Component coordination, zero-copy intra-node communications or inter-nodes data transfers rely on per-node distributed daemons. We evaluate various scenarios performing in situ and in transit analytics on large molecular dynamics systems simulated with Gromacs using up to 1664 cores. We show in particular that analytics processing can be performed on the fraction of resources the simulation does not use well, resulting in a limited impact on the simulation performance (less than 6%). Our more advanced scenario combines in situ and in transit processing to compute a molecular surface based on the Quicksurf algorithm.

5.9. Generic Deterministic Random Number Generation in Dynamic-Multithreaded Platforms

On dynamic multithreaded platforms with on-line scheduling such as work-stealing, randomized computations raise the issue of reproducibility. Compliant with de facto standard sequential Deterministic Random Number Generators (DRNGs) noted R, we propose [23] a parallel DRNG implementation for finite computations that provides deterministic parallel execution. It uses the stateless sub-stream approach, enabling the use of efficient DRNG such as Mersenne Twister or Linear Congruential. We demonstrate that if R provides fast jump ahead in the random sequence, the re-seeding overhead is small, polylog in expectation, independently from the parallel computation's depth. Experiments benchmark the performance of randomized algorithms employing our solution against the stateful DRNG DotMix, tailored to the Cilk Plus dynamic multithreading runtime. The overhead of our implementation ParDRNG compares favorably to the linear overhead of DotMix re-seedings.

ROMA Team

6. New Results

6.1. Highlights of the Year

Yves Robert was awarded the 2014 IEEE Technical Committee on Scalable Computing (TCSC) Award for Excellence.

In October 2014, CERFACS, ENS Lyon, INPT, Inria and University of Bordeaux launched a consortium around the software package MUMPS (see http://mumps-consortium.org).

6.2. Cost-Optimal Execution of Boolean DNF Trees with Shared Streams

Several applications process queries expressed as trees of Boolean operators applied to predicates on sensor data streams, e.g., mobile apps and automotive apps. Sensor data must be retrieved from the sensors, which incurs a cost, e.g., an energy expense that depletes the battery of a mobile device, a bandwidth usage. The objective is to determine the order in which predicates should be evaluated so as to shortcut part of the query evaluation and minimize the expected cost. This problem has been studied assuming that each data stream occurs at a single predicate. In this work [17], [27] we study the case in which a data stream occurs in multiple predicates, either when each predicate references a single stream or when a predicate can reference multiple streams. In the single-stream case we give an optimal algorithm for a single-level tree and show that the problem is NP-complete for DNF trees. For DNF trees we show that there exists an optimal predicate evaluation order that is depth-first, which provides a basis for designing a range of heuristics. In the multi-stream case we show that the problem is NP-complete even for single-level trees. As in the single stream case, for DNF trees we show that there exists a depth-first leaf evaluation order that is optimal and we design efficient heuristics.

6.3. Efficient checkpoint/verification patterns for silent error detection

Errors have become a critical problem for high performance computing. Checkpointing protocols are often used for error recovery after fail-stop failures. However, silent errors cannot be ignored, and their peculiarity is that such errors are identified only when the corrupted data is activated. To cope with silent errors, we need a verification mechanism to check whether the application state is correct. Checkpoints should be supplemented with verifications to detect silent errors. When a verification is successful, only the last checkpoint needs to be kept in memory because it is known to be correct. In this work (RR UT-EECS-14-729), we analytically determine the best balance of verifications and checkpoints so as to optimize platform throughput. We introduce a balanced algorithm using a pattern with p checkpoints and q verifications, which regularly interleaves both checkpoints and verifications across same-size computational chunks. We show how to compute the waste of an arbitrary pattern, and we prove that the balanced algorithm is optimal when the platform MTBF (Mean Time Between Failures) is large in front of the other parameters (checkpointing, verification and recovery costs). We conduct several simulations to show the gain achieved by this balanced algorithm for well-chosen values of p and q, compared to the base algorithm that always perform a verification just before taking a checkpoint ($p = q = 1$), and we exhibit gains of up to 19%.

6.4. Assessing general-purpose algorithms to cope with fail-stop and silent errors

In this work (RR-Inria-8599), we combine the traditional checkpointing and rollback recovery strategies with verification mechanisms to address both fail-stop and silent errors. The objective is to minimize either makespan or energy consumption. While DVFS is a popular approach for reducing the energy consumption, using lower speeds/voltages can increase the number of errors, thereby complicating the problem. We consider

an application workflow whose dependence graph is a chain of tasks, and we study three execution scenarios: (i) a single speed is used during the whole execution; (ii) a second, possibly higher speed is used for any potential re-execution; (iii) different pairs of speeds can be used throughout the execution. For each scenario, we determine the optimal checkpointing and verification locations (and the optimal speeds for the third scenario) to minimize either objective. The different execution scenarios are then assessed and compared through an extensive set of experiments.

6.5. Scheduling the I/O of HPC applications under congestion

A significant percentage of the computing capacity of large-scale platforms is wasted due to interferences incurred by multiple applications that access a shared parallel file system concurrently. One solution to handling I/O bursts in large-scale HPC systems is to absorb them at an intermediate storage layer consisting of burst buffers. However, our analysis of the Argonne's Mira system shows that burst buffers cannot prevent congestion at all times. As a consequence, I/O performance is dramatically degraded, showing in some cases a decrease in I/O throughput of 67%. In this work (RR-Inria-8519), we analyze the effects of interference on application I/O bandwidth, and propose several scheduling techniques to mitigate congestion. We show through extensive experiments that our global I/O scheduler is able to reduce the effects of congestion, even on systems where burst buffers are used, and can increase the overall system throughput up to 56%. We also show that it outperforms current Mira I/O schedulers.

6.6. Power-aware replica placement in tree networks with multiple servers per client

In this work (RR-Inria-8474), we revisit the well-studied problem of replica placement in tree networks. Rather than minimizing the number of servers needed to serve all client requests, we aim at minimizing the total power consumed by these servers. In addition, we use the most general (and powerful) server assignment policy, where the requests of a client can be served by multiple servers located in the (unique) path from this client to the root of the tree. We consider multi-modal servers that can operate at a set of discrete speeds, using the dynamic voltage and frequency scaling (DVFS) technique. The optimization problem is to determine an optimal location of the servers in the tree, as well as the speed at which each server is operated. A major result is the NP-completeness of this problem, to be contrasted with the minimization of the number of servers, which has polynomial complexity. Another important contribution is the formulation of a Mixed Integer Linear Program (MILP) for the problem, together with the design of several polynomial-time heuristics. We assess the efficiency of these heuristics by simulation. For mid-size instances (up to 30 nodes in the tree), we evaluate their absolute performance by comparison with the optimal solution (obtained via the MILP). The most efficient heuristics provide satisfactory results, within 20% of the optimal solution.

6.7. Parallel scheduling of task trees with limited memory

This work [28] investigates the execution of tree-shaped task graphs using multiple processors. Each edge of such a tree represents some large data. A task can only be executed if all input and output data fit into memory, and a data can only be removed from memory after the completion of the task that uses it as an input data. Such trees arise, for instance, in the multifrontal method of sparse matrix factorization. The peak memory needed for the processing of the entire tree depends on the execution order of the tasks. With one processor the objective of the tree traversal is to minimize the required memory. This problem was well studied and optimal polynomial algorithms were proposed.

Here, we extend the problem by considering multiple processors, which is of obvious interest in the application area of matrix factorization. With multiple processors comes the additional objective to minimize the time needed to traverse the tree, i.e., to minimize the makespan. Not surprisingly, this problem proves to be much harder than the sequential one. We study the computational complexity of this problem and provide inapproximability results even for unit weight trees. We design a series of practical heuristics achieving different trade-offs between the minimization of peak memory usage and makespan. Some of these heuristics are able to process a tree while keeping the memory usage under a given memory limit. The different heuristics are evaluated in an extensive experimental evaluation using realistic trees.

6.8. Scheduling Trees of Malleable Tasks for Sparse Linear Algebra

Scientific workloads are often described as directed acyclic task graphs. In this work [30], we focus on the multifrontal factorization of sparse matrices, whose task graph is structured as a tree of parallel tasks. Among the existing models for parallel tasks, the concept of *malleable* tasks is especially powerful as it allows each task to be processed on a time-varying number of processors. Following the model advocated by Prasanna and Musicus [62], [63] for matrix computations, we consider malleable tasks whose speedup is p^α, where p is the fractional share of processors on which a task executes, and α $(0 < \alpha \leq 1)$ is a parameter which does not depend on the task. We first motivate the relevance of this model for our application with actual experiments on multicore platforms. Then, we study the optimal allocation proposed by Prasanna and Musicus for makespan minimization using optimal control theory. We largely simplify their proofs by resorting only to pure scheduling arguments. Building on the insight gained thanks to these new proofs, we extend the study to distributed multicore platforms. There, a task cannot be distributed among several distributed nodes. In such a distributed setting (homogeneous or heterogeneous), we prove the NP-completeness of the corresponding scheduling problem, and propose some approximation algorithms. We finally assess the relevance of our approach by simulations on realistic trees. We show that the average performance gain of our allocations with respect to existing solutions (that are thus unaware of the actual speedup functions) is up to 16% for $\alpha = 0.9$ (the value observed in the real experiments).

6.9. Non-clairvoyant reduction algorithms for heterogeneous platforms

In this work [6], we have revisited the classical problem of the reduction collective operation in a heterogeneous environment. We have discussed and evaluated four algorithms that are non-clairvoyant, i.e., they do not know in advance the computation and communication costs. On the one hand, Binomial-stat and Fibonacci-stat are static algorithms that decide in advance which operations will be reduced, without adapting to the environment; they were originally defined for homogeneous settings. On the other hand, Tree-dyn and Non-Commut-Tree-dyn are fully dynamic algorithms, for commutative or non-commutative reductions. We have shown that these algorithms are approximation algorithms with constant or asymptotic ratios. We assessed the relative performance of all four non-clairvoyant algorithms with heterogeneous costs through a set of simulations. Our conclusions hold for a variety of distributions.

6.10. Memory-aware tree traversals with pre-assigned tasks

We have studied the complexity of traversing tree-shaped workflows whose tasks require large I/O files. We target a heterogeneous architecture with two resource types, each with a different memory, such as a multicore node equipped with a dedicated accelerator (FPGA or GPU). The tasks in the workflow are colored according to their type and can be processed if all there input and output files can be stored in the corresponding memory. The amount of used memory of each type at a given execution step strongly depends upon the ordering in which the tasks are executed, and upon when communications between both memories are scheduled. The objective is to determine an efficient traversal that minimizes the maximum amount of memory of each type needed to traverse the whole tree. In this study [11], we establish the complexity of this two-memory scheduling problem, and provide inapproximability results. In addition, we design several heuristics, based on both post-order and general traversals, and we evaluate them on a comprehensive set of tree graphs, including random trees as well as assembly trees arising in the context of sparse matrix factorizations.

6.11. Analysis of Dynamic Scheduling Strategies for Matrix Multiplication on Heterogeneous Platforms

The tremendous increase in the size and heterogeneity of supercomputers makes it very difficult to predict the performance of a scheduling algorithm. Therefore, dynamic solutions, where scheduling decisions are made at runtime have overpassed static allocation strategies. The simplicity and efficiency of dynamic schedulers such as Hadoop are a key of the success of the MapReduce framework. Dynamic schedulers such as StarPU, PaRSEC or StarSs are also developed for more constrained computations, e.g. task graphs coming from linear

algebra. To make their decisions, these runtime systems make use of some static information, such as the distance of tasks to the critical path or the affinity between tasks and computing resources (CPU, GPU,. . .) and of dynamic information, such as where input data are actually located. In this study [16], we concentrate on two elementary linear algebra kernels, namely the outer product and the matrix multiplication. For each problem, we propose several dynamic strategies that can be used at runtime and we provide an analytic study of their theoretical performance. We prove that the theoretical analysis provides very good estimate of the amount of communications induced by a dynamic strategy and can be used in order to efficiently determine thresholds used in dynamic scheduler, thus enabling to choose among them for a given problem and architecture.

6.12. Determining the optimal redistribution

The classical redistribution problem aims at optimally scheduling communications when reshuffling from an initial data distribution to a target data distribution. This target data distribution is usually chosen to optimise some objective for the algorithmic kernel under study (good computational balance or low communication volume or cost), and therefore to provide high efficiency for that kernel. However, the choice of a distribution minimizing the target objective is not unique. This leads to generalizing the redistribution problem as follows: find a re-mapping of data items onto processors such that the data redistribution cost is minimal, and the operation remains as efficient. This work studies the complexity of this generalized problem. We compute optimal solutions and evaluate, through simulations, their gain over classical redistribution. We also show the NP-hardness of the problem to find the optimal data partition and processor permutation (defined by new subsets) that minimize the cost of redistribution followed by a simple computational kernel. Finally, experimental validation of the new redistribution algorithms are conducted on a multicore cluster, for both a 1D-stencil kernel and a more compute-intensive dense linear algebra routine.

6.13. On the hierarchically structured bin packing problem

We study the hierarchically structured bin packing problem [14]. In this problem, the items to be packed into bins are at the leaves of a tree. The objective of the packing is to minimize the total number of bins into which the descendants of an internal node are packed, summed over all internal nodes. We investigate an existing algorithm and make a correction to the analysis of its approximation ratio. Further results regarding the structure of an optimal solution and a strengthened inapproximability result are given.

6.14. Heuristics for the bipartite matching problem

We propose two heuristics for the bipartite matching problem that are amenable to shared-memory parallelization [18]. The first heuristic is very intriguing from parallelization perspective. It has no significant algorithmic synchronization overhead and no conflict resolution is needed across threads. We show that this heuristic has an approximation ratio of around 0.632. The second heuristic is designed to obtain a larger matching by employing the well-known Karp-Sipser heuristic on a judiciously chosen subgraph of the original graph. We show that the Karp-Sipser heuristic always finds a maximum cardinality matching in the chosen subgraph. Although the Karp-Sipser heuristic is hard to parallelize for general graphs, we exploit the structure of the selected subgraphs to propose a specialized implementation which demonstrates a very good scalability. Based on our experiments and theoretical evidence, we conjecture that this second heuristic obtains matchings with cardinality of at least 0.866 of the maximum cardinality. We discuss parallel implementations of the proposed heuristics on shared memory systems. Experimental results, for demonstrating speed-ups and verifying the theoretical results in practice, are provided.

6.15. Fill-in reduction in sparse matrix factorizations using hypergraphs

We discuss the use of hypergraph partitioning based methods in fill-reducing orderings of sparse matrices for Cholesky, LU and QR factorizations [33]. For the Cholesky factorization, we investigate a recent result on pattern-wise decomposition of sparse matrices, generalize the result, and develop algorithmic tools to obtain more effective ordering methods. The generalized results help us formulate the fill-reducing ordering

problem for LU factorization as we do for the Cholesky case, without ever symmetrizing the given matrix A as $|A| + |A^T|$ or $|A^T||A|$. For the QR factorization, we adopt a recently proposed technique to use hypergraph models in a fairly standard manner. The method again does not form the possibly much denser matrix $|A^T||A|$. We also discuss alternatives for LU and QR factorization cases where the symmetrized matrix can be used. We provide comparisons with the most common alternatives in all three cases.

6.16. On partitioning two dimensional finite difference meshes for distributed memory parallel computers

We investigate the problem of partitioning finite difference meshes in two dimensions among the processors of a parallel computer [20]. The objective is to achieve a perfect load balance while minimizing the communication cost. There are well-known graph, hypergraph, and geometry-based partitioning algorithms for this problem. The known geometric algorithms have linear running time and obtain the best results for very special mesh sizes and processor numbers. We propose another geometric algorithm. The proposed algorithm is linear; is applicable to much more cases than some well-known alternatives; obtains better results than the graph partitioning algorithms; obtains better results than the hypergraph partitioning algorithms almost always. Our algorithm also obtains better results than a known asymptotically-optimal algorithm for some small number of processors. We also catalog related theoretical results.

6.17. A symmetry preserving algorithm for matrix scaling

We present an iterative algorithm which asymptotically scales the ∞-norm of each row and each column of a matrix to one [12]. This scaling algorithm preserves symmetry of the original matrix and shows fast linear convergence with an asymptotic rate of 1/2. We discuss extensions of the algorithm to the one-norm, and by inference to other norms. For the 1-norm case, we show again that convergence is linear, with the rate dependent on the spectrum of the scaled matrix. We demonstrate experimentally that the scaling algorithm improves the conditioning of the matrix and that it helps direct solvers by reducing the need for pivoting. In particular, for symmetric matrices the theoretical and experimental results highlight the potential of the proposed algorithm over existing alternatives.

6.18. Direct solvers for sparse linear systems

In the context of the MUMPS sparse direct solver (see Section 5.1), we worked in 2014 on: block-low-rank solvers and shared-memory parallelism [4], [13], hybrid (shared-distributed) parallelism and efficient collective communications in asynchronous environments [2], and scheduling strategies to decrease the memory-usage of multifrontal solvers. Quite significant performance gains have been obtained on up to 2000 cores of a Bullx DLC system (CALMIP mesocentre), some of the corresponding developments will be made available in the next release of our solver. We also worked on setting up a consortium of industrial users to fund engineers working on MUMPS (see Section 7.1). These activities were done in collaboration with INP Toulouse and with CERFACS, CNRS, ENS Lyon, Univ. Bordeaux, EDF, LSTC (Livermore, California) and EMGS (Norway).

SOCRATE Project-Team

6. New Results

6.1. Highlights of the Year

6.1.1. FIT/CortexLab Inauguration

FIT(Future Internet of Things) is a french Equipex (Équipement d'excellence) which aims to develop an experimental facility, a federated and competitive infrastructure with international visibility and a broad panel of customers. FIT is be composed of four main parts: a Network Operations Center (NOC), a set of Embedded Communicating Object (ECO) test-beds, a set of wireless OneLab test-beds, and a cognitive radio test-bed (CorteXlab) deployed by the Socrate team in the Citi lab. In 2014 the construction of the room was finished see Figure 5 . SDR nodes have installed in the room, 42 industrial PCs (Aplus Nuvo-3000E/P), 22 NI radio boards (USRP) and 18 Nutaq boards (PicoSDR, 2x2 and 4X4) can be programmed from internet now.

A very successfully inauguration took place on the 28th October 2014 [0], with the noticable venue of Vincent Poor, Dean of School of Engineering and Applied Science of Princeton University.
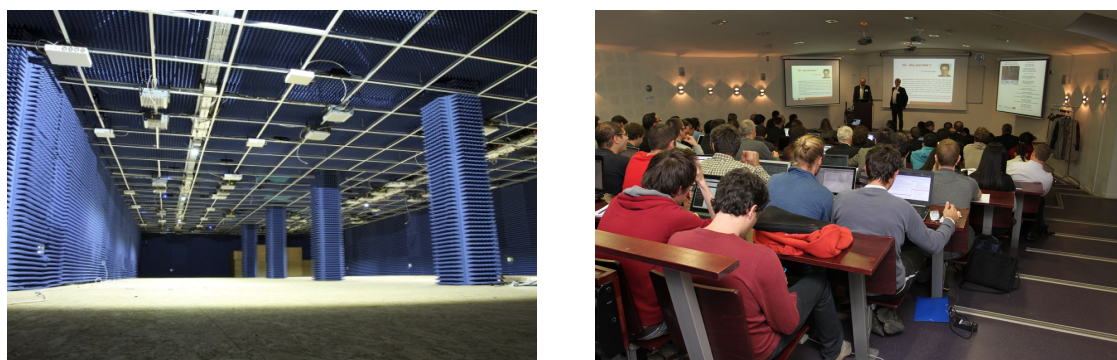
Figure 5. Photo of the FIT/CortexLab experimentation room installed and a snaptshot of the inauguration meeting

6.2. Flexible Radio Front-End

The innovative Wake-Up radio architecture proposed by the Socrate team, based on a classical WiFi standard with a specific OFDM pattern, has been deeply studied in theory and simulations [1], [25], [24]. Great enhancements on the sensitivity study, the choice of identifiers and the comparison of the energy consumption relative to classical systems have led to the development of a first prototype (ongoing work).

6.2.1. Wake-Up Radios

The innovative Wake-Up radio architecture proposed by the Socrate team, based on a classical WiFi standard with a specific OFDM pattern, has been deeply studied in theory and simulations [HUTU-JWCN][KHOUMERI-ECUMICT][HUTU-RWS]. Great enhancements on the sensitivity study, the choice of identifiers and the comparison of the energy consumption relative to classical systems have led to the development of a first prototype (ongoing work).

---

[0]http://www.inria.fr/centre/grenoble/actualites/inauguration-reussie-de-la-plateforme-cortexlab-equipex-fit

6.2.2. Full-Duplex systems

In the development of wideband OFDM Full-Duplex systems, [33] proposes an analysis of the impact of the thermal noise on the quality of the self-interference cancellation in such systems. A method is proposed to reduce the impact on the bit-error-rate by increasing the level of certain parts of the preamble in each frame. [35] add to the analog RF cancellation proposed previously a stage of digital cancellation enabling to increase more the performance of Full-Duplex terminals.

Furthermore, [34] extend the study to a dualband Full-Duplex systems, enabling the very promising combination of Full-Duplex and carrier aggregation. The proposed structure being sensitive to IQ impairments, a digital mitigation algorithm is also designed.

6.2.3. SDR for SRD

In collaboration with Orange labs, [32] analyses the requirements of an SDR gateway for urban networks collecting SRD (short range devices) information. This study is particularly focused on the ADC resolution, showing that the required resolution in realistic scenarios is too high, therefore emphasizing the need to develop specific hardware techniques.

6.2.4. Experimental Facilities

For the development of the CorteXlab testbed, lots of radio hardware and propagation constraints had to be taken into account [15], [14]. Moreover,[36] had proposed a first implementation of Full-Duplex on USRPs which is expected to be deployed on this tesbed.

Another testbed dedicated to the measurement of the energy consumption of radio devices was also designed and implemented.

6.3. Agile Radio Resource Sharing

This axis addresses the challenges relative to the network perspective of software radio. While the two other axes work on the design of the software radio nodes, we focus herein on their coexistence in a multi-user communications perspective. We are first interested in theoretical limits of some reference scenarios where trade-offs between spectral efficiency, energy efficiency, stability and/or fairness are analyzed. Our research activities are further driven by applicative frameworks. We focused on radio access networks with new results on energy efficiency-spectral efficiency trade-off in LTE networks and multi-band CSMA strategies in Wifi networks. We also studied pure random access and success probabilities for the challenging ultra-narrow band (UNB) technology of SigFox. Lot of efforts has been put on body area networks [8] with deep studies on positioning strategies and distributed decisions and information gathering. As mentioned above, our research follows three objectives:

- Establishing theoretical limits of cooperative wireless networks in the network information theory framework.
- Designing MAC procedures, coding and signal processing techniques for optimal transmissions (e.g. interference alignment).
- Developing distributed mechanisms for distributed decision at layer 1 and 2, using game theory, consensus and graph modeling.

6.3.1. Theoretical limits from information theory

The group strengthened his activities from a formal perspective in the framework of network information theory as initiated with the recruitment of Samir Perlaza and the sabbatical of Jean-Marie Gorce at Princeton University in the group of Prof. H. Vincent Poor. The first scenario is devoted to cellular networks with a random distribution on base stations. The main contribution concerns the broadcast channel (BC) generalized to a continuum of users. The second scenario concerns the interference channel (IC) and the main contribution is relative to the characterization of the Nash stable region for the interference channel with noisy feedback.

6.3.1.1. Broadcast channel with a continuum of users in a typical cell

The theoretical Energy efficiency-Spectral efficiency Pareto optimal front in a typical cell has been evaluated by associating stochastic geometry (Poisson point processes, PPP) and information theory.[21], the broadcast channel is extended to a continuum of users. We derived the theoretical uniform achievable rate with superposition coding principles. We show the potential gain of superposition coding techniques compared to the conventional time sharing. These results are however limited to Gaussian channels and the extension to the vector Gaussian channel is still under investigation. The PPP modeling for multi-cells has been also introduced as well as the price of interference management.

6.3.1.2. Interference Channel with feedback

The decentralized interference channel (DIC) with noisy feedback has been analyzed. In [31], all the rate-pairs that are achievable at a Nash equilibrium (NE) in the two-user linear deterministic symmetric decentralized interference channel (LD-S-DIC) with noisy feedback are identified. A second result provides closed form expressions for the PoA, which allows the full characterization of the reduction of the sum rate due to the anarchic behavior of all transmitter-receiver pairs. The price of anarchy (PoA) and the price of stability (PoS) of the game in which transmit-receiver pairs seek an optimal individual transmission rate are fully characterized in [9]. In particular, it is shown that in all interference regimes, there always exists at least one Pareto optimal Nash equilibrium (NE).

6.3.2. Coding, signal processing and MAC procedures for optimal transmissions

6.3.2.1. Implementation

While theoretical studies provide interesting insights about potential gain and limits of cognitive networks, the achievable efficiency may depend on practical issues related to quantization, synchronization and real-time processing limits. We developed the CortexLab facility offering a reproducible environment for fostering the validation of cooperative communication schemes. The first demo has been presented at the Infocom conference [28] and also at the Melbourne Greentouch meeting. We also contributed to the implementation and analysis of a cognitive transceiver for opportunistic networks [ref Maso JWCN]. The work first focused on a previously introduced dynamic spectrum access (DSA) - cognitive radio (CR) solution for primary-secondary coexistence in opportunistic orthogonal frequency division multiplexing (OFDM) networks, called cognitive interference alignment (CIA). The implementation is based on software-defined radio (SDR) and uses GNU Radio and the universal software radio peripheral (USRP) as the implementation toolkit. The proposed flexible transceiver architecture allows efficient on-the-fly reconfigurations of the physical layer into OFDM, CIA or a combination of both.

6.3.2.2. Interference alignment

In the framework of Greentouch, we studied interference alignment as a mean for improving the EE-SE tradeoff in cellular networks [43]. We combined theoretical studies with stochastic geometry and simulations to show the potential interest. We are also developing a demo with CorteXlab enhancing the IA capability from a real perspective.

6.3.2.3. Multiband MAC

In collaboration with CEA-Leti, we studied MAC strategies for multiband systems. The main idea is based on exploiting the multiband system as a slotted Aloha channel for the RTS/CTS initiation but keeping the total band as a whole for data transmission. We proved that this strategy outperforms classical approaches [39], [40], [30].

6.3.2.4. MAC for localization

In the context of the ANR Cormoran project, we account for radiolocation experiments aiming at both indoor navigation and mobility detection applications for Wireless Body Area Networks (WBAN) [7]. We also studied the relation between the MAC protocol and ranging techniques for localization. The impact of mobility on the distance estimation between 2 nodes of a Wireless Body Area Network (WBAN) by comparing the Two-Way Ranging (2WR) and Three-Way Ranging (3WR) protocols has been proposed in [23]. We also investigated the impact of mobility on the Motion Capture applications [22].

6.3.2.5. random access in Ultra-narrow band networks

Ultra narrow band (UNB) transmission is a very promising technology for low-throughput wireless sensor networks. This technology has already been deployed and has proved to be ultra-efficient for point-to-point communications in terms of power-efficiency, and coverage area. We studied the scalability of UNB for a multi-point to point network. In particular, we proposed a new multiple access scheme: random frequency division multiple access (R-FDMA) and studied the impact of the induced interference on the system performance in terms of bit error rate and outage probability [20]. We also analyzed the system performance in terms of bit error rate and outage probability [37].

6.3.3. Distributed decision mechanisms

Distributed decisions appear in many situations in the wireless world. Resource allocation, power management or relaying techniques are all expecting distributed decisions. To avoid strong coordination, distributed mechanisms inspired e.g. by game theory or consensus algorithms are appealing. Some of the results obtained below also rely on information theory but with a more important focus on algorithms and decision processes when several pairs of wireless transceivers are willing to simultaneously transmit in the same environment.

6.3.3.1. Cognitive radio networks

The problem of joint channel selection and power control is analyzed in the context of multiple-channel clustered ad-hoc networks in[ref Rose [3], i.e., decentralized networks in which radio devices are arranged into groups (clusters) and each cluster is managed by a central controller (CC). The problem is modeled by a game in normal form in which the corresponding utility functions are designed for making some of the Nash equilibria (NE) to coincide with the solutions to a global network optimization problem. A second scenario has been considered where multiple source-destination pairs communicate with each other via an energy harvesting relay [5]. The focus was put on the relay's strategies to distribute the harvested energy among the multiple users and their impact on the system performance. Specifically, a non-cooperative strategy that uses the energy harvested from the i-th source as the relay transmission power to the i-th destination is considered first. An auction based power allocation scheme is also proposed to achieve a better tradeoff between system performance and complexity.

6.3.3.2. Distributed decisions and consensus in MANETs

In the large research area of wireless body area networks, cooperative applications involving several users is attracting strong interests. This cooperation may target a simple information exchange or even some cooperative decision such as swarm coordination. We considered in [26]such a swarm of users moving in a common direction and we are interested in the mechanisms allowing to propagate and share some common information. We extend and improve a previous algorithm derived as a max-consensus approach. We describe a complete experimental setup deployed during a real bike race with 200 runners.

6.4. Software Radio Programming Model

6.4.1. Data Flow Programming

Software defined radio (SDR) technology has evolved rapidly and is now reaching market maturity. However, no standard has emerged for programming the new type of machine that will manage the access to the radio channel. Mickaël Dardaillon, Kevin Marquet, Tanguy Risset have been working in collaboration with the CEA LETI on compiling waveform for heterogeneous Multi-processor SoCs . This research leaded to a prototype compiler for the Magali MP-SoC developed in Mickael Dardaillon' PhD thesis (passed in November 2014) which was the first attempt to compile the SPDF format to a real architecture [18], [16], [17]. This study highlighted in particular the fact that SPDF was a good computation model for waveform description langage, easier to compile than dynamic dataflow format.

6.4.2. *Non-volatile memory management for ultra low power systems*

To enable non-trivial computation on very resource-constrained platforms powered by energy harvested from RF communications, an embedded OS has to save and restore program state to and from non-volatile memory. By doing so, the application program does not lose all progress when power is lost, which happens very often in environmentally-powered systems. This can be achieved [13] thanks to an incremental checkpointing scheme which aims at minimizing the amount of data written to non-volatile memory, while keeping the execution overhead as low as possible.

6.4.3. *FPGA-based Implementation of physical Layers for SDR*

A VHDL implementation of the three available options of the IEEE 802.15.4 physical layer was developed [29] in the context of FIT/CorteXlab. This parametrized design was validated on a Nutaq platform which combines Xilinx Virtex-6 FPGA and tunable Radio420x RF transceiver. This work participates to the building of an open source hardware SDR library similar to GNU radio but targeted to FPGA-based platforms.

6.4.4. *Towards filters and functions computing just right*

A FIR filter is specified by its coefficients (real numbers) and its input and output formats. The implementation of a FIR should be as accurate as its output format allows, but no more. This very simple specification enables the automatic construction of FIR filter implementations that are provably accurate at a minimal hardware cost [19]. The corresponding FIR generator is available in FloPoCo.

The fixed-point Atan2 function is very useful to recover the phase of a complex signal. A careful study of three implementation techniques (including a novel one based on two-variable quadratic approximation) shows that, on current FPGAs, the good old CORDIC technique is more efficient than multiplier-based techniques [46].

TYREX Project-Team

6. New Results

6.1. Automated Refactoring for Size Reduction of CSS Style Sheets

Cascading Style Sheets (CSS) is a standard language for stylizing and formatting web documents [17]. Its role in web user experience becomes increasingly important. However, CSS files tend to be designed from a result-driven point of view, without much attention devoted to the CSS file structure as long as it produces the desired results. Furthermore, the rendering intended in the browser is often checked and debugged with a document instance. Style sheets normally apply to a set of documents, therefore modifications added while focusing on a particular instance might affect other documents of the set.

We present a first prototype and a new CSS semantic analyzer and optimizer that is capable of automatically detecting and removing redundant property declarations and rules. We build on earlier work on tree logics to locate redundancies due to the semantics of selectors and properties. Existing purely syntactic CSS optimizers can be used in conjunction with our tool, for performing complementary (and orthogonal) size reduction, toward the common goal of providing smaller and cleaner CSS files. We have been able to detect large numbers of unnecessary property declarations in complex web pages; and we have also found mistakes in the style sheets of some of the most popular web sites. The number of safe modifications can easily grow as more components of CSS are supported and more features are implemented, such as property inheritance, translation of pseudo-classes into query languages, analysis of media queries, merging of equivalent selectors or containment involving grouped selectors.

6.2. Equipping IDEs with XML-Path Reasoning Capabilities

One of the challenges in Web development is to achieve a good level of quality in terms of code size and runtime performance for popular domain-specific languages such as XQuery, XSLT, and XML Schema. We developed an IDE augmented with static detection of inconsistent XPath expressions that assists the programmer with simplifying development and debugging of any application involving XPath expressions [12]. The tool is based on newly developed formal verification techniques based on expressive modal logics, which are now efficient enough to be used in the process of software development. We applied this to a full XQuery compiler for which we introduced an analysis for identifying and eliminating dead code automatically.

6.3. XQuery and Static Typing: Tackling the Problem of Backward Axes

XQuery is a functional language dedicated to XML data querying and manipulation. As opposed to other W3C-standardized languages for XML (e.g. XSLT), it has been intended to feature strong static typing. Currently, however, some expressions of the language cannot be statically typed with any precision. This is due to a discrepancy between the semantics of the language and its type algebra: namely, the values of the language are (possibly inner) tree nodes, which may have siblings and ancestors in the data. The types on the other hand are regular tree types, as usual in the XML world: they describe sets of trees. The type associated to a node then corresponds to the subtree whose root is that node and contains no information about the rest of the data. This makes navigation expressions using "backward axes," which return e.g. the siblings of a node, impossible to type.

We show how to handle this discrepancy by improving the type system. We describe a logic-based language of extended types able to represent inner tree nodes and show how it can dramatically increase the precision of typing for navigation expressions. We describe how inclusion between these extended types and the classical regular tree types can be decided, allowing a hybrid system combining both type languages. The result is a net increase in precision of typing [20].

6.4. A Core Calculus for XQuery 3.0: Combining Navigational and Pattern Matching Approaches

XML processing languages can be classified according to whether they extract XML data by paths or patterns. The strengths of one category correspond to the weaknesses of the other. In this work, we propose to bridge the gap between these two classes by considering two languages, one in each class: XQuery (for path-based extraction) and CDuce (for pattern-based extraction). To this end, we extend CDuce so as it can be seen as a succinct core λ-calculus that captures XQuery 3.0. The extensions we consider essentially allow CDuce to implement XPath-like navigational expressions by pattern matching and precisely type them. The elaboration of XQuery 3.0 into the extended CDuce provides a formal semantics and a sound static type system for XQuery 3.0 programs [18].

6.5. Session Types as Generic Process Types

Behavioural type systems ensure more than the usual safety guarantees of static analysis [15]. They are based on the idea of "types-as-processes", providing dedicated type algebras for particular properties, ranging from protocol compatibility to race-freedom, lock-freedom, or even responsiveness. Two successful, although rather different, approaches, are session types and process types. The former allows to specify and verify (distributed) communication protocols using specific type (proof) systems; the latter allows to infer from a system specification a process abstraction on which it is simpler to verify properties, using a generic type (proof) system. What is the relationship between these approaches? Can the generic one subsume the specific one? At what price? And can the former be used as a compiler for the latter? This work is a step towards answers to such questions. Concretely, we have defined a stepwise encoding of a pi-calculus with sessions and session types (the system of Gay and Hole) into a pi-calculus with process types (the Generic Type System of Igarashi and Kobayashi). We encode session type environments, polarities (which distinguish session channels end-points), and labelled sums. We show forward and reverse operational correspondences for the encodings, as well as typing correspondences. To faithfully encode session subtyping in process types subtyping, one needs to add to the target language record constructors and new subtyping rules. This work shows how the programming convenience of session types as protocol abstractions can be combined with the simplicity and power of the pi-calculus, taking advantage in particular of the framework provided by the Generic Type System.

6.6. Personal Shopping and Navigator System for Visually Impaired People

We have developed a a personal assistant and navigator system for visually impaired people [14]. This system has been built using a set of domain specific languages based on XML such as OpenStreetMap extended for Augmented Reality. It demonstrate how partially sighted people could be aided by the technology in performing an ordinary activity, like going to a mall and moving inside it to find a specific product. We propose an Android application that integrates Pedestrian Dead Reckoning and Computer Vision algorithms, using an off-the-shelf Smartphone connected to a Smart-watch. The detection, recognition and pose estimation of specific objects or features in the scene derive an estimate of user location with sub-meter accuracy when combined with a hardware-sensor pedometer. The proposed prototype interfaces with a user by means of Augmented Reality, exploring a variety of sensorial modalities other than just visual overlay, namely audio and haptic modalities, to create a seamless immersive user experience. The interface and interaction of the preliminary platform have been studied through specific evaluation methods. The feedback gathered will be taken into consideration to further improve the proposed system.

<p align="center" style="color:red">URBANET Team</p>

6. New Results

6.1. Highlights of the Year

Two scientific results can be distinguished in UrbaNet activity this year. First of all, the work did in collaboration with Orange Labs during the PhD thesis of O. Erdene-Ochir (defended in 2013) led to a patent [38] related to routing in wireless sensor networks under resiliency constraints.

A second important result is represented by the book chapter "Wireless Access Networks for Smart Cities" [31], a common contribution of all the permanent members of the team. We hope that this chapter will become the reference on wireless networking within the new and dynamic smart cities community.

6.2. Characterizing and measuring urban networks

Participants: R. Domga Komguem, M. Fiore, D. Naboulsi, P. Raveneau, R. Stanica, F. Valois

6.2.1. *Collection and Analysis of Mobile Phone Data*

Cellular communications are undergoing significant evolutions in order to accommodate the load generated by increasingly pervasive smart mobile devices. At the same time, recent generations of mobile phones, embedding a wide variety of sensors, have fostered the development of open sensing applications, such as network quality or weather forecast applications.

In this sense, we contributed with a novel privacy-preserving mobile data collection platform [21], leveraging the dynamic deployment of crowdsourcing tasks across a pouplation of mobile phones.

Using such data, or other datasets coming from network operators, we can propose dynamic access network mechanisms that adapt to customers' demands. To that end, one must be able to process large amount of mobile traffic data and outline the network utilization in an automated manner. In [28], we propose a framework to analyze broad sets of Call Detail Records (CDRs) so as to define categories of mobile call profiles and classify network usages accordingly. We evaluated our framework on a CDR dataset including more than 300 million calls recorded in an urban area over 5 months. We showed how our approach allows to classify similar network usage profiles and to tell apart normal and outlying call behaviors.

6.2.2. *Generation and Analysis of Vehicular Mobility Datasets*

The surge in vehicular network research has led, over the last few years, to the proposal of countless network solutions specifically designed for vehicular environments. A vast majority of such solutions has been evaluated by means of simulation, since experimental and analytical approaches are often impractical and intractable, respectively. The reliability of the simulative evaluation is thus paramount to the performance analysis of vehicular networks, and the first distinctive feature that has to be properly accounted for is the mobility of vehicles, i.e., network nodes. Notwithstanding the improvements that vehicular mobility modeling has undergone over the last decade, no vehicular mobility dataset was publicly available that captures both the macroscopic and microscopic dynamics of road traffic over a large urban region.

In [12], we present a realistic synthetic dataset, covering 24 hours of car traffic in a 400-km2 region around the city of Ko¨ln, in Germany. We describe the generation process and outline how the dataset improves the traces currently employed for the simulative evaluation of vehicular networks. We also show the potential impact that such a comprehensive mobility dataset has on the network protocol performance analysis, demonstrating how incomplete representations of vehicular mobility may result in over-optimistic network connectivity and protocol performance.

Moreover, using a similar methodology we contribute to the ongoing effort to define such mobility scenarios by introducing a second set of traces for vehicular network simulation, this time focusing on a highway environment. Our traces are derived from high-resolution real-world traffic counts, and describe the road traffic on two highways around Madrid, Spain, at several hours of different working days. We provide a thorough discussion of the real-world data underlying our study, and of the synthetic trace generation process [20] [35] [29]. Finally, we assess the potential impact of our dataset on networking studies, by characterizing the connectivity of vehicular networks built on the different traces. Our results underscore the dramatic impact that relatively small communication range variations have on the network. Also, they unveil previously unknown temporal dynamics of the topology of highway vehicular networks, and identify their causes.

6.2.3. *Characterizing Novel Wireless Networks for Urban Intelligent Transportation Solutions*

Vehicular networks are not the only contribution communication technologies can bring in the field of Intelligent Transportation Systems. Two other examples have been studied this year in the team.

The first example is related to traffic light control in an urban environment [17]. A traffic light controller takes as input an estimation of the number of vehicles entering the intersection and produces as output a light plan, with the objective to reduce the traffic jam. The quality of the input traffic estimation is a key consideration on the performance of the traffic light controller. The advent of Wireless Sensor Networks, with their relatively low deployment and operation price, led to the development of several sensor-based architectures for intersection monitoring. We show in this work that the solutions proposed in the literature are unrealistic in terms of communication possibilities and that they do not allow a measure of the vehicular queue length at a lane level. Based on extensive experimental results, we propose an energy efficient, low cost and lightweight multi-hop wireless sensor network architecture to measure with a good accuracy the vehicle queue length, in order to have a more precise vision of traffic at the intersection.

On a second example, these last years have witnessed the rise of the smart cities and several mechanisms to render the cities more sustainable and more energy-efficient. Among all different aspects, a noteworthy one is urban bike development. Besides the growing enthusiast provoked by bicycles and the benefit for health they bring, there still exists some reluctance in using bikes because of safety, road state, weather, etc. To counter-balance these feelings, there is a need to better understand bicycle users habits, path, road utilization rate in order to improve the bicycle path quality. In this perspective, in [25], we propose to deploy a set of mobile sensors on bicycles to gather this different data and to exploit them to make the bike easier and make people want to ride bicycles more often. Such a network will also be useful for several entities like city authorities for road maintenance and deployment, doctors and environment authorities, etc. Based on such a framework, we propose a first basis model that help to dimension the network infrastructure and the kind of data to be real time gathered from bikes. More specifically, we present a theoretical model that computes the quantity of data a bike will be able to send along a travel and the quantity of data a base station should be able to absorb. We have based our study on real data to provide first numerical results and be able to draw some preliminary conclusions and open new research directions.

6.3. Technology specific solutions

Participants: I. Augé-Blum, W. Bechkit, J. Cui, A. Mouradian, T. Lin, H. Rivano, R. Stanica, F. Valois

6.3.1. *Medium Access Control in Wireless Sensor Networks*

Protocols developed during the last years for Wireless Sensor Networks (WSNs) are mainly focused on energy efficiency and autonomous mechanisms (e.g. self-organization, self-configuration, etc.). Nevertheless, with new WSN applications, new QoS requirements appear, such as time constraints. Real-time applications require the packets to be delivered before a known time bound which depends on the application requirements. We particularly focus on applications which consist in alarms sent to the sink node. We propose Real-Time X-layer Protocol (RTXP) [8], a real-time communication protocol. RTXP is a MAC and routing real-time communication protocol that is not centralized, but instead relies only on local information. To the best of our knowledge, it is the first real-time protocol for WSNs using an opportunistic routing scheme in order

to increase the packet delivery ratio. In the paper above, we describe the protocol mechanisms. We give theoretical bounds on the end-to-end delay and the capacity of the protocol. Intensive simulation results confirm the theoretical predictions and allow to compare RTXP with a real-time scheduled solution. RTXP is also simulated under harsh radio channel, in which case the radio link introduces probabilistic behavior. Nevertheless, we show that RTXP performs better than a non-deterministic solution. It thus advocates for the usefulness of designing real-time (deterministic) protocols even for highly unreliable networks such as WSNs.

Continuing on the idea of WSN applications with strict temporal constraints, these critical applications require correct behavior, reliability, and, of course, the respect of time constraints. Otherwise, if they fail, consequences on human life and the environment could be catastrophic. For this reason, we argue that the WSN protocols used in these applications must be formally verified. Unfortunately the radio link is unreliable, it is thus difficult to give hard guarantees on the temporal behavior of the protocols (on wired systems the link error probability is very low, so they are considered reliable). Indeed, in WSN a message may experience a very high number of retransmissions. The temporal guarantee has thus to be given with a probability that it is achieved. This probability must meet the requirements of the application. Network protocols have been successfully verified on a given network topology without taking into account unreliable links. Nevertheless, the probabilistic nature of radio links may change the topology (links which appear and disappear). Thus, instead of a single topology we have a set of possible topologies, each topology having a probability to exist. In this paper, we propose a method that produces the set of topologies, checks the property on every topology, and gives the probability that the property is verified. This technique is independent from the verification technique, i.e. each topology can be verified using any formal method which can give a "yes" or "no" answer to the question: "Does the model of the protocol respect the property?". In [27], we apply this method on the previously proposed f-MAC protocol, a real-time medium access protocol for WSNs. We use UPPAAL model checker as verification tool, and we perform simulations to observe the difference between average and worst case behaviors.

One WSN application gaining a lot of importance in the team in the last few years targets Intelligent Transportation Systems (ITS), as also explained in the previous section. In this ITS field, parking sensor networks are rapidly deploying around the world and are also regarded as one of the first implemented urban services in smart cities. To provide the best network performance in this context, the MAC protocol shall be adaptive enough in order to satisfy the traffic intensity and variation of parking sensors. In this sense, in [24] and [36], we compare the performance of two off-the-shelf medium access control protocols on two different kinds of traffic models, and then evaluate their application-end information delay and energy consumption while varying traffic parameters and network density. From the simulation results, we highlight some limits induced by network density and occurrence frequency of event-driven applications. When it comes to real-time urban services, a protocol selection shall be taken into account - even dynamically - with a special attention to the energy-delay trade-off. In a follow-up study [23], we use real world data, more precisely the heavy-tailed parking and vacant time models from the SmartSantander platform, and then we apply the traffic model in the simulation with four different kinds of MAC protocols, that is, contention-based, schedule-based and two hybrid versions of these. The result shows that the packet inter-arrival time is no longer heavy-tailed while collecting a group of parking sensors, and then choosing an appropriate MAC protocol highly depends on the network configuration. Also, the information delay is bounded by traffic and MAC parameters which are important criteria while the timely message is required.

6.3.2. *Routing in Wireless Sensor Networks*

Routing represents another major challenging issue in WSN, because of the application diversity and energy efficiency constraints. Gradient broadcast routing is a robust scheme for data gathering in WSNs. At each hop, the sender broadcasts the packet to its neighbors and one or more nodes among its neighbors closer to the sink forward it. As long as a node has at least one neighbor with a smaller hop-count, it can route packets. Nevertheless, nodes can disappear because of energy depletion, hardware failure, etc. In this case, it cannot be ensured that a packet reaches the sink. Usually this issue is addressed by updating the gradient with a periodical flooding. Nevertheless, it consumes an important amount of energy, moreover, parts of the network may not need to be updated. In [26], we propose GRABUP (GRAdient Broadcast UPdate), a traffic-based gradient

maintenance algorithm which updates the gradient thanks to the data packets. We simulate the proposition and compare it with the classic gradient broadcast routing.

Another specific application that we target is smart metering, which heavily rely on the communication network for efficient data gathering, thus eliminating manual meter reading. Smart electronic devices are deployed in open, unattended and possibly hostile environment such as consumer's home and office areas, making them particularly vulnerable to physical attacks. Resilience is needed in this case to mitigate such inherent vulnerabilities and risks related to security and reliability. In [18], a general overview of the resilience including definition, metric and resilient techniques relevant for smart metering is presented. A quantitative metric, visual and meaningful, based on the graphical representation is adopted to compare routing protocols in the sense of resilience against active insider attacks. Five well-known routing protocols from the main categories have been studied through simulations and their resilience is evaluated according to the given metric. Resilient techniques introduced to these protocols have enhanced significantly the resilience against attacks providing route diversification.

6.3.3. *Other Research Issues Related to Wireless Sensor Networks*

Important features of WSNs, such as low battery consumption, changing topology awareness, open environment, non reliable radio links, raise other research issues than classical MAC and routing problems. For example, in [32], we investigate the benefits of Network Coding in WSN, especially with respect to resiliency. We have seen in our previous work that resiliency could be described as a multi dimensional metric, taking parameters such as Average Delivery Ratio, Delay Efficiency, Energy Efficiency, Average Throughput and Delivery Fairness into account. Resiliency can then be graphically represented as a kiviat diagram created by the previous weighted parameters. In order to introduce these metrics, previous works have been leaded on the Random Gradient Based Routing, which proved good resiliency in malicious environment. We look for seeing the improvements in term of resiliency, when adding network coding in the Random Gradient Based Routing with malicious nodes.

Another challenge is represented by the deployment of sensor nodes, which can take into account the impact of multiple parameters. For example, temperature variations have a significant effect on low power WSNs as wireless communication links drastically deteriorate when temperature increases. A reliable deployment should take temperature into account to avoid network connectivity problems resulting from poor wireless links when temperature increases. A good deployment needs also to adapt its operation and save resources when temperature decreases and wireless links improve. Taking into account the probabilistic nature of the wireless communication channel, we develop [4] a mathematical model that provides the most energy efficient deployment in function of temperature without compromising the correct operation of the network by preserving both connectivity and coverage. We use our model to design three temperature-aware algorithms that seek to save energy (i) by putting some nodes in hibernate mode as in the SO (Stop-Operate) algorithm, or (ii) by using transmission power control as in PC (Power-Control), or (iii) by doing both techniques as in SOPC (Stop-Operate Power-Control). All proposed algorithms are fully distributed and solely rely on temperature readings without any information exchange between neighbors, which makes them low overhead and robust. Our results identify the optimal operation of each algorithm and show that a significant amount of energy can be saved by taking temperature into account.

Finally, the notion of Shared Risk Link Groups (SRLG) captures survivability issues when a set of links of a network may fail simultaneously, such as a WSN where link conditions are extremely dynamic. The theory of survivable network design relies on basic combinatorial objects that are rather easy to compute in the classical graph models: shortest paths, minimum cuts, or pairs of disjoint paths. In the SRLG context, the optimization criterion for these objects is no longer the number of edges they use, but the number of SRLGs involved. Unfortunately, computing these combinatorial objects is NP-hard and hard to approximate with this objective in general. Nevertheless some objects can be computed in polynomial time when the SRLGs satisfy certain structural properties of locality which correspond to practical ones, namely the star property (all links affected by a given SRLG are incident to a unique node, for example a battery depleted sensor) and the span property (the links affected by a given SRLG form a connected component of the network). The star property is defined in a multi-colored model where a link can be affected by several SRLGs while the span property is defined

only in a mono-colored model where a link can be affected by at most one SRLG. In [33], we extend these notions to characterize new cases in which these optimization problems can be solved in polynomial time or are fixed parameter tractable. We also investigate on the computational impact of the transformation from the multi-colored model to the mono-colored one. Experimental results are presented to validate the proposed algorithms and principles.

6.3.4. Data Aggregation and Gathering

In the data gathering problem, a particular network node, the base station or the sink, aims at receiving messages from some other network nodes. In [5], we model this network as a graph, and we consider that, at each step, a node can send one message to one of its neighbors (such an action is called a call). However, a node cannot send and receive a message during the same step. Moreover, the communication is subject to interference constraints, more precisely, two calls interfere in a step, if one sender is at distance below a certain threshold from the other receiver. Given a graph with a base station and a set of nodes having some messages, the goal of the gathering problem is to compute a schedule of calls for the base station to receive all messages as fast as possible, i.e., minimizing the number of steps (called makespan). The gathering problem is equivalent in this case to the personalized broadcasting problem where the base station has to send messages to some nodes in the graph, with same transmission constraints. We focus on the gathering and personalized broadcasting problem in grids (regular networks, with nodes deployed in a grid-like shape, e.g. parking or intersection monitoring WSNs). Moreover, we consider the non-buffering model: when a node receives a message at some step, it must transmit it during the next step. In this setting, though the problem of determining the complexity of computing the optimal makespan in a grid is still open, we present linear (in the number of messages) algorithms that compute optimal schedules for data gathering.

Data aggregation is a particular solution for the data gathering problem, which reduces the amount of data sent to the base station. In [16], we show that data aggregation can effectively reduce the energy consumption and improve the network capacity. Moreover, we present the state-of-the-art aggregation functions, including compressing-based and forecasting-based method; compressing-based aggregation focuses on compressing the data packets accompanied with transmitting based on spatial correlation, while forecasting aggregation tends to use mathematical models to fit the time series and predict the new value due to highly temporal correlation. We detail these two methods and characterize them respectively. We propose comparison between A-ARMA and Compressing Sensing, which are noteworthy examples of forecasting aggregation and compressing aggregation respectively.

6.3.5. Safety Vehicular Ad Hoc Networks

Vehicular ad hoc networks can play an important role in enhancing transportation efficiency and improving road safety. Therefore, direct vehicle-to-vehicle communications are considered as one of the main building blocks of a future Intelligent Transportation System. The success and availability of IEEE 802.11 radios made this technology the most probable choice for the medium access control layer in vehicular networks. However, IEEE 802.11 was originally designed in a wireless local area network context and it is not optimized for a dynamic, ad hoc vehicular scenario. In [11], we investigate the compatibility of the IEEE 802.11 medium access control protocol with the requirements of safety vehicular applications. As the protocols in this family are well-known for their scalability problems, we are especially interested in high density scenarios, quite frequent on today's roads. Using an analytical framework, we study the performance of the back-off mechanism and the role of the contention window on the control channel of a vehicular network. Based on these findings, we propose a reverse back-off mechanism, specifically designed with road safety applications in mind. Extensive simulations are carried out to prove the efficiency of the proposed enhancement scheme and to better understand the characteristics of vehicular communications.

One of the major roles of vehicular communication is the dissemination of information on the road in order to increase the awareness of the drivers. The facilities layer is a recently standardized component in the vehicular communication architecture, with an important role to play in the process of information dissemination. In [22], we propose facilities layer-based mechanisms for information propagation and we show they outperform classical network layer solutions. We also demonstrate that previous studies that do

not consider the cohabitation of different types of safety messages on the vehicular control channel highly under-estimate the dissemination delay, which can lead to unrealistic assumptions in the design of safety applications.

6.4. Capillary solutions

Participants: M. Fiore, G. Gaillard, D. Naboulsi, H. Rivano, R. Stanica, F. Valois

6.4.1. Connected Vehicles

Bandwidth availability in the cellular backhaul is challenged by ever-increasing demand by mobile users. Vehicular users, in particular, are likely to retrieve large quantities of data, choking the cellular infrastructure along major thoroughfares and in urban areas. It is envisioned that alternative roadside network connectivity can play an important role in offloading the cellular infrastructure. We investigate [7] the effectiveness of vehicular networks in this task, considering that roadside units can exploit mobility prediction to decide which data they should fetch from the Internet and to schedule transmissions to vehicles. Rather than adopting a specific prediction scheme, we propose a fog-of-war model that allows us to express and account for different degrees of prediction accuracy in a simple, yet effective, manner. We show that our fog-of-war model can closely reproduce the prediction accuracy of Markovian techniques. We then provide a probabilistic graph-based representation of the system that includes the prediction information and lets us optimize content prefetching and transmission scheduling. Analytical and simulation results show that our approach to content downloading through vehicular networks can achieve a 70% offload of the cellular network.

Vehicles also produce large quantities of Floating Car Data (FCD), which consist of information generated by moving vehicles and uploaded to Internet-based control centers for processing and analysis. As upcoming mobile services based on or built for networked vehicles largely rely on uplink transfers of small-sized but high-frequency messages, FCD traffic is expected to become increasingly common in the next few years. Presently, FCD are managed through a traditional cellular network paradigm: however, the scalability of such a model is unclear in the face of massive FCD upload, involving large fractions of the vehicles over short time intervals. In [13], we explore the use of vehicle-to-vehicle (V2V) communication to partially relieve the cellular infrastructure from FCD traffic. Specifically, we study the performance boundaries of such a FCD offloading approach in presence of best- and worst-case data aggregation possibilities at vehicles. We show the gain that can be obtained by offloading FCD via vehicular communication, and propose a simple distributed heuristic that has nearly optimal performance under any FCD aggregation model.

We also advocate the use of a data shuttle service model to offload bulk transfers of delay-tolerant data from the Internet onto standard vehicles equipped with data storage capabilities [14]. We first propose an embedding algorithm that computes an offloading overlay on top of the road infrastructure. The goal is to simplify the representation of the road infrastructure as raw maps are too complex to handle. In this overlay, each logical link maps multiple stretches of road from the underlying road infrastructure. We formulate then the data transfer assignment problem as a novel linear programming model that determines the most appropriate logical paths in the offloading overlay for a data transfer request. We evaluate our proposal using actual road traffic counts in France. Numerical results show that we can satisfy weekly aggregate requests in the petabyte range while achieving cumulative bandwidth above 10 Gbps with a market share of 20% and only one terabyte of storage per vehicle.

6.4.2. Energy Consumption in Communication Networks

Providing high data rates with minimum energy consumption is a crucial challenge for next generation wireless networks. There are few papers in the literature which combine these two issues. The work we propose in [10] focuses on multi-hop wireless mesh networks using a MAC layer based on S-TDMA (Spatial Time Division Multiple Access). We develop an optimization framework based on linear programming to study the relationship between throughput and energy consumption. Our contributions are twofold. First, we formulate and solve, using column generation, a new MILP to compute offline energy-throughput tradeoff curve. We use a physical interference model where the nodes can perform continuous power control and can use a discrete set

of data rates. Second, we highlight network engineering insights. We show, via numerical results, that power control and multirate functionalities allow optimal throughput to be reached, with lower energy consumption, using a mix of single hop and multihop routes.

Another strategy with regard to energy consumption is switching off some network nodes that are not carrying any data or control traffic. In [37], we tackle the problem of on-grid energy saving in cellular networks based on switch-on/off techniques for base stations and the usage of renewable energy. We aim to evaluate how much power can be saved in the network and dimension the renew able energy system according to the consumptions in real-world networks.

6.4.3. *Service Level Agreements*

The era of the Internet of Things (IoT) brings complexity and deployment costs in smart cities, particularly in WSNs. Utilities such as gas or water providers are keen on delegating the management of the communications to specialized firms, namely WSN Operators, that will share the WSN resource among their various clients. For this reason, in [34] we provide a guideline to write Service Level Agreements (SLAs) for IoT operation, borrowing a well studied concept from the web services domain. We extend the SLA definition with specific items that integrate the WSN constraints, and we facilitate the construction of complex metrics that express the performance of the WSN.

Furthermore. WSN operators will need a robust and reliable technology in order to guarantee QoS constraints in a wireless environment, as in the industrial world. IEEE 802.15.4e Time Slotted Channel Hopping (TSCH) is one good candidate. Moreover, the IETF experience in IP networks management is an important input for monitoring and QoS control over WSNs. In [19], we give formal guidelines for the implementation of a SLA architecture for operated WSNs. We distinguish the various formal algorithms that are necessary to operate a WSN according to SLAs, and determines which functional entities are necessarily technology-dependent. Detailed examples of such entities are developed in an IPv6 over IEEE 802.15.4e TSCH context, such as advocated in the IETF 6TiSCH Working Group.

E-MOTION Project-Team

5. New Results

5.1. Highlights of the Year

- C. Laugier, E. Mazer and K. Mekhnacha have been finalists for the Eurobotics Technology Award 2014. Title "Bayesian perception & Decision: from theory to industrial applications". March 2014.
- A. Nègre, L. Rummelhard, M. Perrollaz and C. Laugier had applied for a petenent "Procédé d'analyise d'une scene dybnamique, module d'analyse et programme d'ordinateur associés".

5.2. A new formulation of the Bayesian Occupancy Filter: a hybrid sampling based framework

Participants: Lukas Rummelhard, Amaury Nègre, Christian Laugier.

The Bayesian Occupancy Filter (BOF) is a discretized grid structure based bayesian algorithm, in which the environment is subdivised in cells to which random variables are linked. These random variables represent the state of occupancy and the motion field of the scene, without any notion of object detection and tracking, making the updating part of the filter an evaluation of the distribution of these variables, according to the new data acquisition. In the classic representation of the BOF, the motion field of each cell is represented as a neighborhood grid, the probability of the cell moving from the current one to another of the neighborhood being stocked in an histogram. If this representation is convenient for the update, since the potential antecedents of any cell is exactly determined by the structure, and so the propagation model is easily parallelizable, it also raises determinant issues :

- the structure requires the process rate to be constant, and a priori known.
- in the case of a moving grid, such as an application of car perception, many aliasing problems can appear, not only in the occupation grid, but in the motion fields of cells. A linear interpolation in 4-dimension field to fill each value of the histograms can quickly become unreasonable.
- to be able to match the slowest moves in the scene and the tiniest objects, the resolution of the grid and the motion histogram must be the high. On the other hand, since the system must be able to evaluate the speed of highly dynamic objects (typically, a moving car), the maximum encoded speed is to be high as well. This results in a necessary huge resolution grid, which prevent the system from being used with satisfying results on an embedded device. This huge grid is also mostly empty (most of the motion field histogram for a occupied cell will be empty). On top of that, the perception system being used to represent the direct environment of a moving car, the encoded velocity is a relative velocity, which implies, if we consider the maximal speed of a car to be V_max, to maintain a motion field able to represent speeds from $-2 * V_{max}$ to $2 * V_{max}$. The necessity of such a sized structure is a huge limitation of practical use of the method.

Considering those limitations, a new way to represent the motion field has been developed. To do so, a new formulation of the BOF has been elaborated. This new version allow to introduce in the filter itself a distinction between static and dynamic parts, and so adapt the computation power.

The Hybrid Sampling Bayesian Occupancy Filter (HSBOF) [21] is an evolution of the BOF, in which are introduced additionnal concepts and variables, such as probabilistic classification of the environment between static and dynamic areas, and adaptative motion model structure. The main idea of this new representation is to mix two forms of sampling of the surrounding :

- a uniform sampling, represented as a dense regular grid, for the static objects and the empty areas. In this part, only the occupancy is stored, as the motion model of the static part of the scene is inherent. In practice, the section of the environment includes the vast majority of the scene.

- a non uniform sampling, based on particles drawn in dynamic regions, allowing to focus the computational power on the estimation of their motion. The number of particles used to represent the motion of a particular cell is calculated according to various criterions, such as the confidence in the dynamism of the cell, in its estimated motion, the global needs in the scene, etc. Dynamic regions are resampled at every time step, the amount of particles associated to the different parts of the scene is dynamicly calculated.

The motion field in a cell is then represented as a set of samples from the distribution for values which are not null, and a weight given to the static hypothesis. The use of a set of samples to represent the motion field leads to a important decrease of the needed memory space, so do the classification between dynamic objects and static objects or free areas. In the updating process, the antecedent of a cell can be either from the static configuration or from the dynamic configuration, which are both way easier to project in the new reference frame of the moving grid: the static part requires a 2-dimension interpolation to be expressed in the new reference frame, the dynamic part a immediate particle association and a simple rotation of the velocity vectors.

This new version HSBOF is now used in the core of our systems in place of the previous version of the BOF. It presents important improvements in the quality of the estimations, while drastically reducing the memory and computation costs (easily by a 100 factor in term of memory).

5.2.1. *Probabilistic grid-based collision risk prediction*

Participants: Lukas Rummelhard, Amaury Nègre, Mathias Perrollaz, Christian Laugier.

We developed a new grid-based approach for collision risk prediction [23], based on the Hybrid-Sampling Bayesian Occupancy Filter framework. The idea is to compute an estimation of the Time To Contact (TTC) for each cell of the grid, instead of reasoning on objects. This strategy avoids to solve the difficult problem of multi-objects detection and tracking and provides a probabilistic estimation of the risk associated to each TTC value.

Using motion sensors embedded in the mobile robot (Inertial Measurement Unit, GPS, Wheel speed and steering sensor, visual odometry, etc.), the displacement of the grid between two updates is estimated. The full description of occupancy and dynamics of the scene given by the HSBOF is then used to assess collision risks in the future and even localize them in the grid. The risk evaluation consists in a short-term prediction of the scene configuration (figure 4 and of the robot position. This way a collision likelihood can be computed over time. Using those likelihoods, computed by cell and particle, an estimation of the risk over a period, and a localization of this risk in the grid are performed.

5.3. A new experimental platform for the Technological Research Institute (NanoElec)

Participants: Mathias Perrollaz, Nicolas Turro, Jean-François Cuniberto.

Within the framework of the PERFECT projet (founded by the IRT NanoElec), e-Motion has developed a new experimental platform, based on a Renault Zoe electrical vehicle (Fig. 7). This developpement takes advantage of the experience developed for creating the previous experimental plateform (a Lexus LS600H), and go further by integrating more sensors and more functionalities.

The vehicle is eqquipped with:

- 4 IBEO LUX laser scanners. Each of them scans 4 layers with a field of view of 85 degrees.
- one Velodyne HD64L 3D laser scnanner, capable of scanning 64 layers over 360 degrees.
- one trinocular stereo camera, Point Grey Bumblebee XB3, placed behind the windshield.
- 2 Ueye RGB cameras, looking forward and backward the vehicle.
- one XSens IMU/GPS sensor, used for positionning and ego-motion estimation.
- one ITRI 802.11p on-board unit, allowing V2X communication.
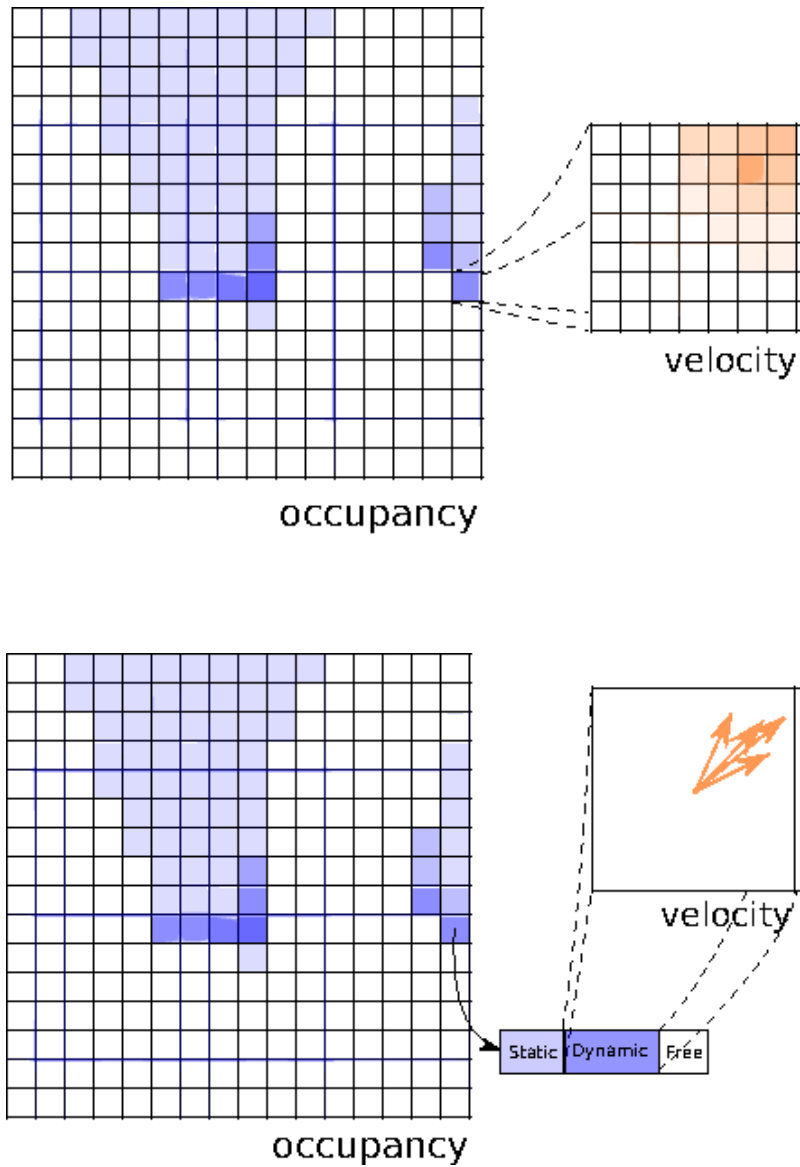
Figure 1.

Data representations in BOF and HSBOF formulation :
(a) Classic BOF representation : a 2 dimension grid, to each cell are assigned an occupancy value and a velocity histogram,
(b) Proposed representation : a 2 dimension grid, to each cell are assigned an occupancy value, a static coefficient $P(V = 0)$ and a set of particles drawn along $P(V = v | V \neq 0)$
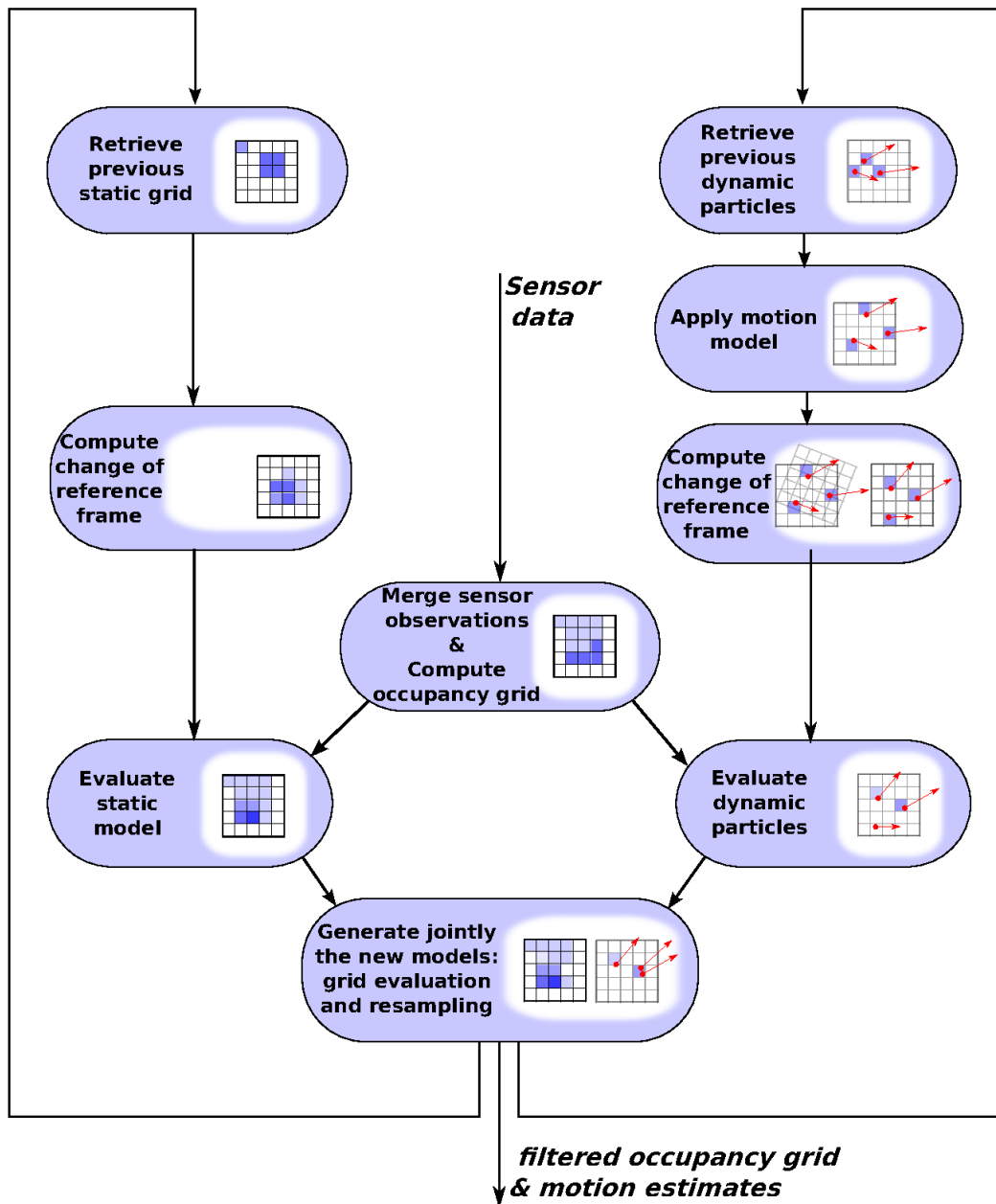
Figure 2. HSBOF algorithm summary. From sensor data instantaneous occupancy grids are successively computed. Those observations are integrated in a Bayesian filter in which coexist and jointly adapt two models, a static grid and a dynamic set of moving particles. The result is obtained by their combination, which provides a filtered occupancy grid as well as inferred motion distributions for cells.
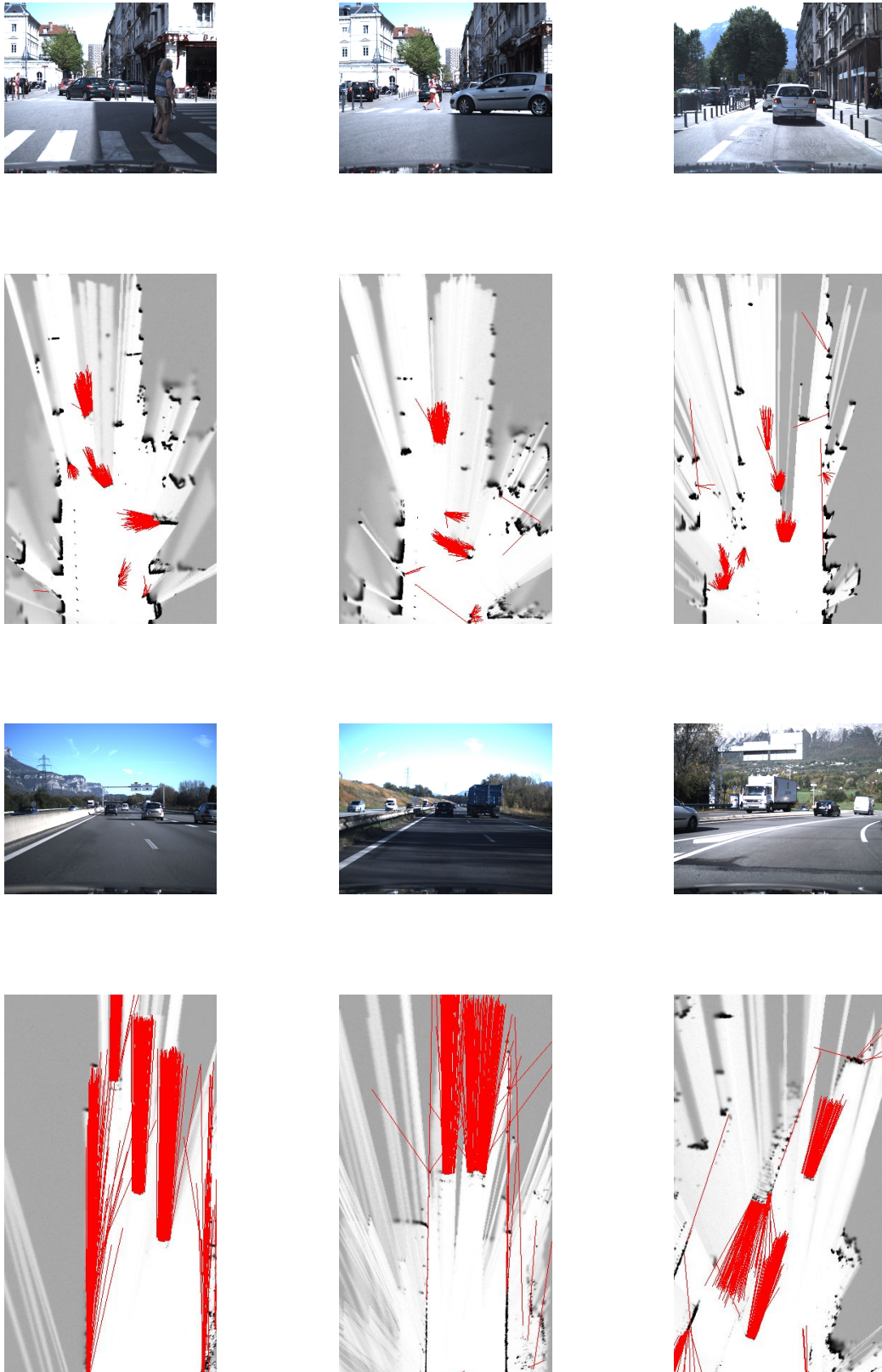
Figure 3. Resulting occupancy grid and velocity field on different urban and highway situations. White cells represent the free space, grey one the unknown space (hidden). Black cells represent the occupied space and red lines represent the average velocity vector for cell with a high dynamic probability.
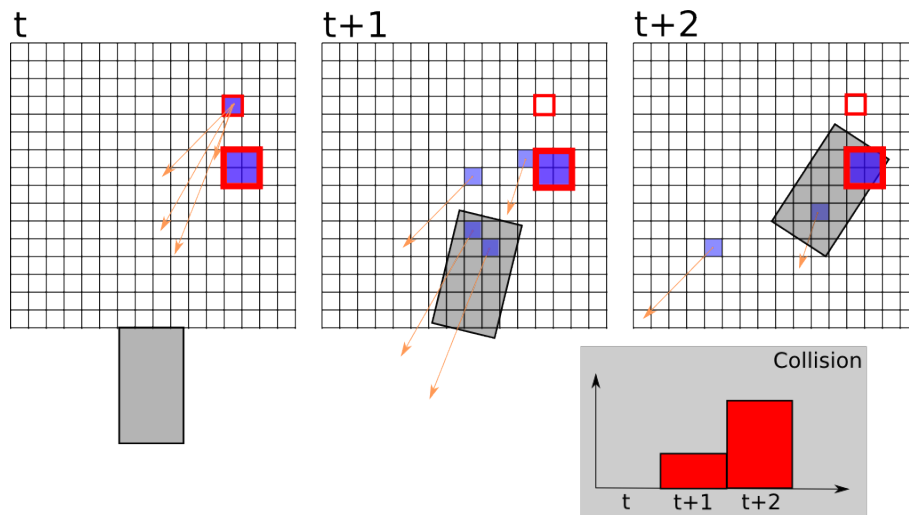
Figure 4. Collision risk estimation over time for a specific cell. The cell position is predicted according to its velocity, along with the mobile robot. This risk profile is computed for every cell, and then used to integrate over time the global collision risk.

(a)
(b)

Figure 5. (a) Fake pedestrian used for experiments. (b) The mannequin is attached to a system with a runner, in order to allow lateral displacements.

(a)

(b)
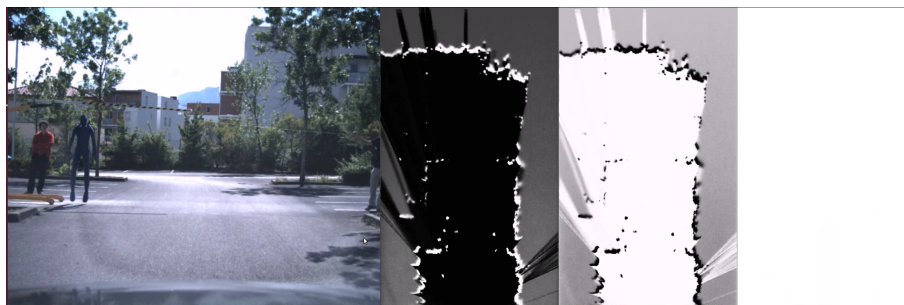
(c)
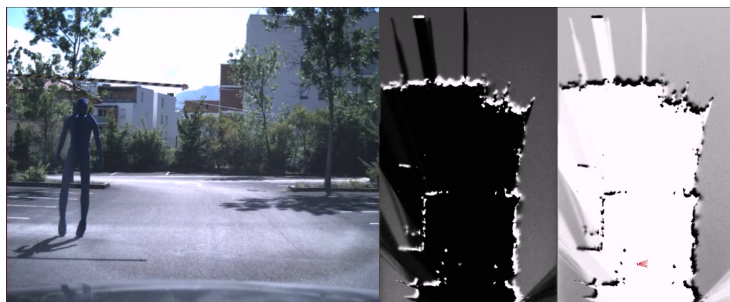
(d)

Figure 6. Results of the system. Each image is a visual capture from the embedded camera, the estimated occupancy grid in front of the car (white for occupied, grey for unknown, black for empty), the estimated motion field (if a case is seen as dynamic, a red motion vector showing the average velocity in the cell is drawn on the map) and finally the estimated risk map for 0.5s. The first sequence (a) (b) presents the appearance of an occluded pedestrian, the second (c) (d) a moving pedestrian heading towards the road.

All the synchronization, display, play/record, and developpments capabilities are relying on the ROS middleware. The vehicle is fully operationnal at the end of 2014.

The vehicle is designed for experimenting in both ADAS (Advanced Driver Assistance Systems) and autonomous driving applications. In parallel, V2X communications are installed on the IRT "smart city" environment, so that the vehicle can evolve on this site and interact with it.

Figure 7. The Zoe experimental platform.

5.4. Visual localization with Open Street Map

Participants: Jean-Alix David, Amaury Nègre.

Given the lack of precision of GPS for localization, it is necessary to implement new ways to improve localization. Here we introduce a new method using a geographic map and a camera to do so. The main point of this method is to combine sensor readings and known data about the environment. We detect lines on the road with the camera, and then compare the extracted lines to the ones stored in the map using ICP (Iterative Closest Point) algorithm.

The used map is OpenStreetMap, it allow to have information on the roads and lanes for example, but there is no information about white marking. So we generated semi-autonomously the lines given roads and number of lanes. Moreover we manually corrected the lines for crossroads using satellite image (see Figure 8).

The line extraction is done using ridge detection on a top-down view of the camera image. Moreover we use GPU acceleration to improve performances during image processing (see Figure 9).

The OSM generated data and the lines extracted from the camera will then be matched and the transformation between the camera and the absolute map will be compute by using an Iterative Closest Point algorithm. In order to improve the precision, a bayesian filtering approach will also be used to merge the previous results with GPS and Inertial Measurement Unit data.

(a) Open Street Map data   (b) Generated lines   (c) Reprojected lines on aerial image after rectification
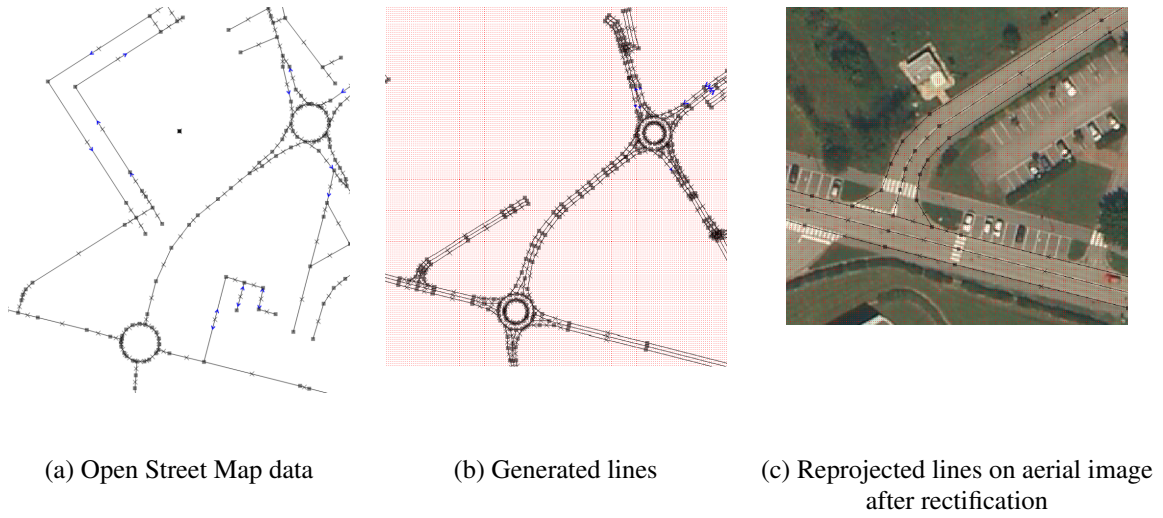
Figure 8. Semi-automatic lroad line generation from Open Street Map.

(a)   (b)   (c)
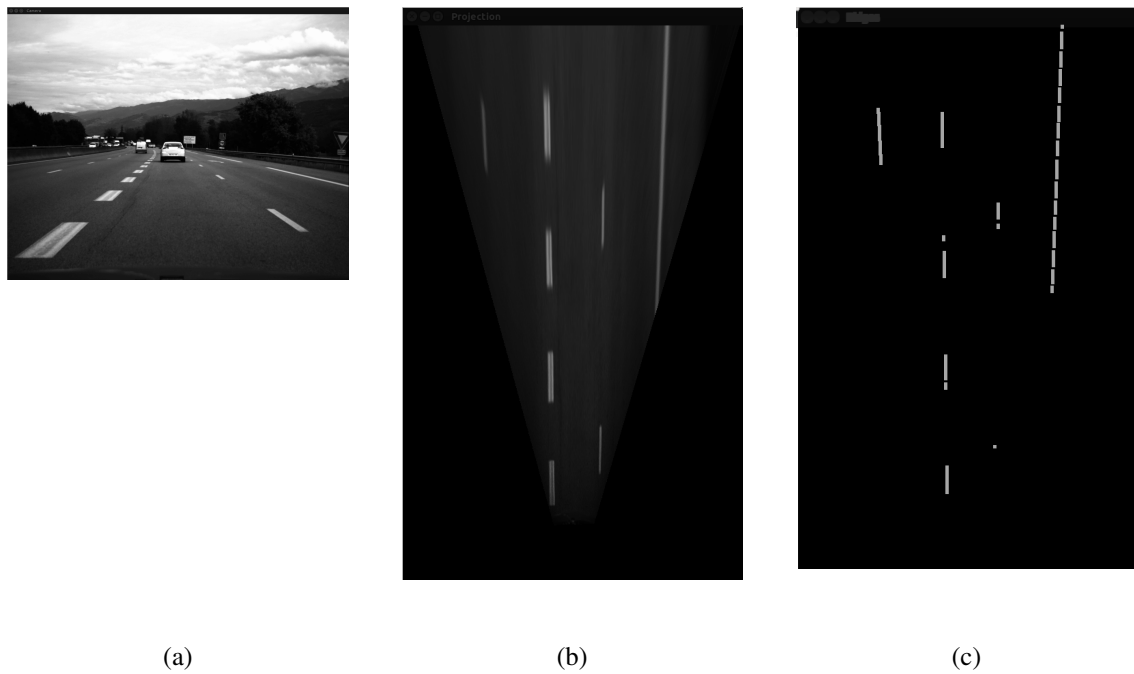
Figure 9. Line detection in camera images: projection in the ground plane (b) ridge extraction (c).

5.5. Human Centered Navigation in the physical world

5.5.1. *Social Mapping*

Participants: Panagiotis Papadakis, Anne Spalanzani, Christian Laugier.

With robots technology shifting towards entering human populated environments, the need for augmented perceptual robotic skills emerges that complement to human presence. In this integration, perception and adaptation to the implicit human social conventions plays a fundamental role. Toward this goal, we introduce in 2013 a novel methodology to detect and analyse complex spatial interactions of multiple people and encode them in the form of a social map, whose structure is obtained by computing a latent space representation of human proxemic behaviour [32]. In 2014, Panagiotis left to Lagadic-Sophia and we carried on this work by integrating a planning algorithm to validate the perception part on a real robot. This work was published at IROS 2014 [22].

5.5.2. *Goal oriented risk based navigation in social and dynamic environment*

Participants: Anne Spalanzani, Procopio Silveira-Stein, Gregoire Vignon, Christian Laugier.

Since 2008 we have proposed a new concept to integrate a probabilistic collision risk function linking planning and navigation methods with the perception and the prediction of the dynamic environments [31]. The likelihood of the obstacles' future trajectory and the probability of occupation are used to compute the risk of collision. A social filter was added to give the robot the ability to move in a social way (see Figure 10). In 2014, we obtained an Inria ADT(ADT PN2) to optimize and share the RiskRRT algorithm. This work is under development. We published in [15] a survey on human-aware navigation.
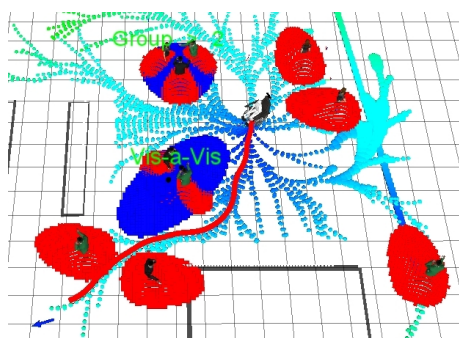
Figure 10. Illustration of the RiskRRT in a social environnment

5.5.3. *Navigation Taking Advantage of Moving Agents*

Participants: Procopio Silveira-Stein, Anne Spalanzani, Christian Laugier.

In this work, we proposes a different form of robotic navigation in dynamic environments, where the robot takes advantage of the motion of pedestrians, in order to improve its own navigation capabilities. The main idea is that, instead of treating persons as dynamic obstacles that should be avoided, they should be treated as special agents with an expert knowledge of navigating in dynamic scenarios. To benefit from the motion of pedestrians, this work proposes that the robot selects and follows them, so it can move along optimal paths, deviate from undetected obstacles, improve navigation in densely populated areas and increase its acceptance by other humans. In 2014, we focused on real experiments (see Figure 11 using the wheelchair and results were published in [16], [25], [26].

Figure 11. Switching navigation method between leader following and independent navi- gation. In image 1 the robot is engaged in leader following, while in the remaining it uses RiskRRT for the navigation.

5.5.4. *Autonomous Wheelchair for Elders Assistance*

Participants: Arturo Escobedo-Cabello, Gregoire Vignon, Anne Spalanzani, Christian Laugier.

The aging of world's population is bringing the need to provide robotic platforms capable to assist elder people to move [33]. It is necessary that such transportation is reliable, safe and comfortable. People with motor disabilities and elders are expected to benefit from new developments in the field of autonomous navigation robotics.

Autonomously driven wheelchairs are a real need for those patients who lack the strength or skills to drive a normal electric wheelchair. The services provided by this kind of robots can also be used to provide a service of comfort, assisting the user to perform difficult tasks as traversing a door, driving in a narrow corridor etc. In 2014, we combined user intention estimation, a navigation using social convention to perform comfortable trajectories (see Figure 12 . Results were published in the IROS conference [19]. Arturo Escobedo defended his phD in october 2014.

Figure 12. The Robotic wheelchair assists its user to acheive his task of navigation/.

5.6. Human modeling for situation understanding

5.6.1. *Situation understanding and risk assessment for intelligent vehicles*

Participants: Dizan Vasquez, Stéphanie Lefèvre, Suryansh Kumar, Yufeng Yu.

The work on this period has been aimed at establishing a solid theoretical and technological base for our research on situation understanding. A step in this direction was the elaboration of an in-depth survey of the current state of the art on the field, prepared together with the university of Berkeley [12]. In the framework of the same collaboration, we have been working on the introduction of human models in current Advanced Driving Assistance Systems (ADAS). This has led to the development of a novel Lane Keeping Assistance System (LKAS) which is able to learn the driver's driving patterns and use them to predict lane departures as well as to generate controls that mimic the driver's style and are, thus, deemed to be more acceptable. The approach has been evaluated against commercial LKAS using real field data, and the results show that the proposed approach is both more efficient and less intrusive than current approaches. This is, for the best of our knowledge the first use of human models withing LKAS and these results illustrate the strong potential that these models may have in ADAS.

Concerning autonomous navigation, we have focused on human-like motion planning for motion prediction. The main hypothesis is that people behave like planners whose motion optimizes some an unknown cost function. Under this assumption, the main challenge becomes to model that cost function and to learn its parameters from demonstrated behavior. This is called, depending on the community, either *Inverse Reinforcement Learning* (IRL) or *Inverse Optimal Control* (IOC). Now, a problem with IRL is that it requires examples of both desirable and undesirable behavior, which are difficult to obtain with a real platform. Additionally, there is no consistent benchmarking methodology to evaluate different approaches. This has motivated our work in a benchmark comprised of: (a) an evaluation methodology; (b) a simulated experimental platform (Fig. 13) based on the Torcs simulator ; and (c) real data gathered with our instrumented Lexus vehicle. The first prototype of this benchmark, developed together with students from Beijing University and IIIT Hyderabad, has been presented this year in a vehicular technologies conference .

Figure 13. Experimental platforms: left) Our Torcs-based racing simulator; right) sensor-equipped Lexus vehicle.

5.6.2. *Socially compliant robot navigation in human environments*
Participant: Dizan Vasquez.

The models we have applied to intelligent vehicles are also adapted in general to situations where mobile robots share their environment with humans. This has lead us to apply this techniques to the assistive robotics fields, given that it is one of e-motion's major applications axes. Our first effort in this sense has been to design and develop a robust experimental platform with baseline modules for motion planning, perception and social awareness .

In parallel we started working, in collaboration with the University of Freiburg, on a benchmarking platform for social compliant motion planning, close in spirit to the one proposed for intelligent vehicles. The platform (Fig. 14) is described in , it includes several motion planning and feature extraction algorithms as well as a pedestrian simulator based on Helbing's social force model .

5.7. Sensor Fusion for state parameters identification
Participants: Agostino Martinelli, Chiara Troiani.

5.7.1. *General theoretical results*

During this year we have focused our research on two distinct domains:
- the visual-inertial structure from motion problem;
- the derivation of analytical solutions for the probability distribution of a Brownian motion that satisfies the unicycle constraint.
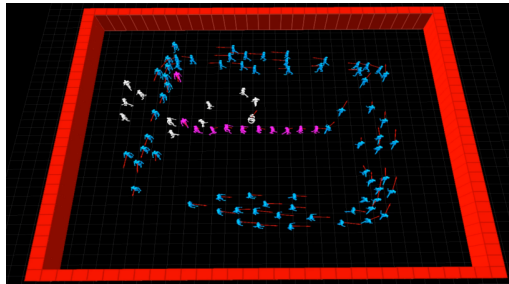
Figure 14. A screenshot of our pedestrian simulator.

The research carried out on the first domain is the follow up of our previous activity. We continued to investigate the observability properties of the visual inertial structure from motion and in particular we have analyzed the case when some of the inertial sensors are missing. This analysis has never been provided before and we started this investigation at the end of last year. During this year we confirmed the validity of our preliminary analysis and we also extended them. The preliminary results were obtained by referring to the case when at least five pont features are available and showed that the observability properties of visual inertial structure from motion do not change by removing all the three gyroscopes and one accelerometer. By removing a further accelerometer, if the camera is not extrinsically calibrated, the system loses part of its observability properties. On the other hand, if the camera is extrinsically calibrated, the system maintains the same observability properties as in the standard case. These results have been published on the journal Foundations and Trends in Robotics and have also been presented at the last ICRA conference[20].

We recently extended these results by considering the extreme case of a single point feature (i.e., not five). This analysis required to approach an open problem in control theory, called the Unknown Input Observability (UIO). In [20] we proposed a possible method to solve this UIO problem. However, we had to improve this method to deal with this extreme case (i.e., the case of one single point feature). Preliminary results on the extension of this method have been published as a research report [30] and we also plan to present them at the next American Control Conference. By applying this method to our problem, we obtained new interesting results. The new investigation allowed us to conclude that, even in the case of a single point feature, the information provided by a sensor suit composed by a monocular camera and two inertial sensors (along two independent axes and where at least one is an accelerometer) is the same as in the case of a complete inertial measurement unit (i.e., when the inertial sensors consist of three orthogonal accelerometers and three orthogonal gyroscopes). Our first objective is to validate these new results.

Regarding the second domain mentioned above, we have derived a complete analytical solution for the probability distribution of the configuration of a non-holonomic mobile robot that moves in two spatial dimensions by satisfying the unicycle kinematic constraints. The proposed solution differs from previous solutions since it is obtained by deriving the analytical expression of any-order moment of the probability distribution. To the best of our knowledge, an analytical expression for any-order moment that holds even in the case of arbitrary linear and angular speed, has never been derived before. To compute these moments, a direct integration of the Langevin equation has been carried out and each moment was expressed as a multiple integral of the deterministic motion (i.e., the known motion that would result in absence of noise).

For the special case when the ratio between the linear and angular speed is constant, the multiple integrals can be easily solved and expressed as the real or the imaginary part of suitable analytic functions. As an application of the derived analytical results, we also investigated the diffusivity of the considered Brownian motion for constant and for arbitrary time-dependent linear and angular speed. These results have been published on

the journal of statistical mechanics [13] and also as a research report [29] where we added more specific considerations about the impact of the derived results on mobile robotics.

5.7.2. Applications with a Micro Aerial Vehicle

We continued our previous activity about the estimation of the relative motion between two consecutive camera views in order to introduce very efficient algorithms to remove the outliers of the feature-matching process. Thanks to their inherent efficiency, the proposed algorithms are very suitable for computationally-limited robots.

In particular, during this year, we extended the previous results by removing the assumption of planar motions. In this case, to obtain useful results, we had to include one more point feature (i.e., the proposed algorithms only use two feature correspondences and gyroscopic data from IMU measurerments to compute the motion hypothesis). By exploiting this 2-point motion parametrization, we proposed two algorithms to remove wrong data associations in the feature matching process for case of a 6DoF motion. We showed that in the case of a monocular camera mounted on a quadrotor vehicle, motion priors from IMU can be used to discard wrong estimations in the framework of a 2 -point-RANSAC based approach. The proposed methods have been evaluated on both synthetic and real data and presented at the last ICRA conference [27].

5.8. Compiling Probabilistic Programs Onto Reconfigurable Logic Using Stochastic Arithmetic

Participants: Emmanuel Mazer, Marvin Faix.

It is of great interest to perform light weight probabilistic inferences for applications such as sensor fusion. Our goal is to design systems to perform these inferences without using a Von Newman machine nor standard floating point arithmetic. By addressing the core of how computations are made, we can explore the tradeoffs between system precision with power consumption and computation time, enabling artificial systems with limited resources, such as mobile and embedded systems, to better operate under uncertainty. Figure 15 illustrates the tool-chain, which starts from the specification of the Bayesian Program in Bayesian programming language , and evaluates it on a reconfigurable device.

This study is part of BAMBI (Bottom-up Approaches to Machines dedicated to Bayesian Inference, www. bambi-fet.eu) : a European collaborative research project relying on the theory of Bayesian inference to understand the natural cognition and aiming at designing bio-inspired computing devices.

A Bayesian machine has probability distributions as inputs and returns a probability distribution as output. It is defined by a joint probability distribution on a set of discrete and finite variables: $P(M \wedge D \wedge L)$. Where M, D and L are themselves conjunctions of variables, for example $D = D_1 \wedge ... \wedge D_k$. We define the soft evidences on the variables D_k as the probability distribution $\widetilde{P}(D_k)$. These soft evidences will be the inputs of the Bayesian machine.

So, given the soft evidences $\widetilde{P}(D_k)$ and the joint distribution $P(M \wedge D \wedge L)$, the machine will fulfil the specification if it computes:

$$P'(M) = \frac{1}{Z} \sum_{D_1} \widetilde{P}(D_1) ... \sum_{D_k} \widetilde{P}(D_k) \sum_{L} P(M \wedge D \wedge L) \tag{2}$$

with

$$Z = \sum_{M} (\sum_{D_1} \widetilde{P}(D_1) ... \sum_{D_k} \widetilde{P}(D_k) \sum_{L} P(M \wedge D \wedge L)$$

In other words the machine computes a soft inference based on the joint distribution $P(M \wedge D \wedge L)$.

A modified version of the probabilistic language ProBT is used to specify the machine: the joint distribution, the output and the inputs are specified with this language [0]. The next program is an example of a simple specification using the Python bindings of ProBT.
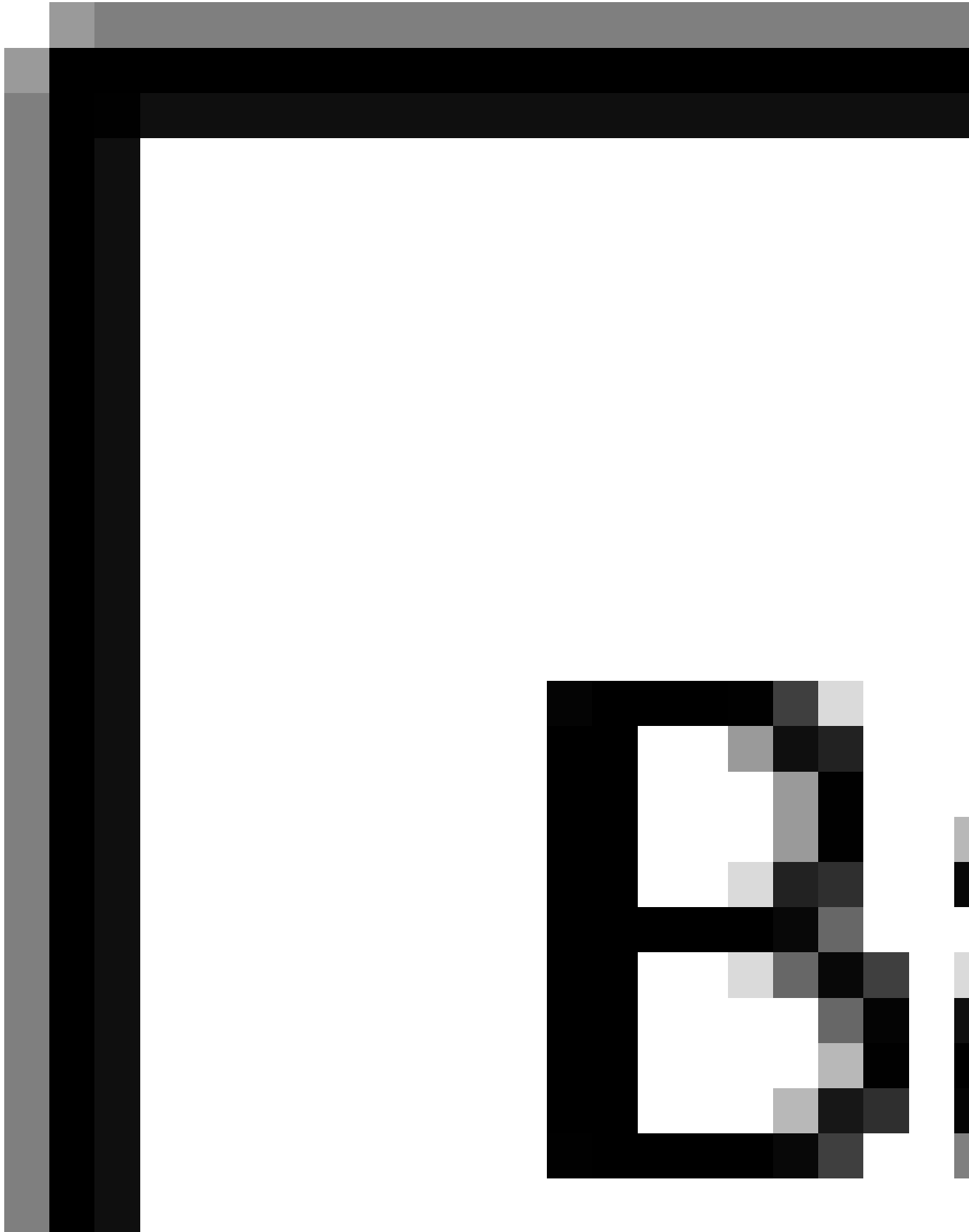
```
#import the ProBT bindings
from pypl import *
#define the variables
dim3 = plIntegerType(0,2)
D1 = plSymbol(D1,dim3)
D2 = plSymbol(D2,dim3)
M= plSymbol(M,dim3)
#define the distribution on M
PM= plProbTable(M,[0.8,0.1,0.1])
#define a conditional distribution on D1
PD1_k_M = plDistributionTable(D1,M)
PD1_k_M.push(plProbTable(D1,[0.5,0.2,0.3]),0)
PD1_k_M.push(plProbTable(D1,[0.5,0.3,0.2]),1)
PD1_k_M.push(plProbTable(D1,[0.4,0.3,0.3]),2)
#define a conditional distribution on D2
PD2_k_M = plDistributionTable(D2,M)
PD2_k_M.push(plProbTable(D2,[0.2,0.6,0.2]),0)
PD2_k_M.push(plProbTable(D2,[0.6,0.3,0.1]),1)
PD2_k_M.push(plProbTable(D2,[0.3,0.6,0.1]),2)
#define the joint distribution
model=plJointDistribution(PM*PD1_k_M*PD2_k_M)
#define the soft evidence variables
model.set_soft_evidence_variables(D1^D2)
#define the output
question=model.ask(M)
```
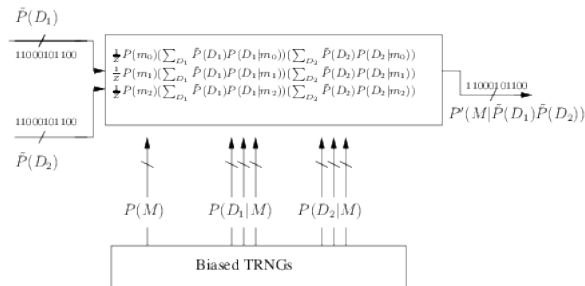

Figure 16. The probabilistic machine corresponding to the given program..

Figure 16 presents the high-level representation of the architecture for the Bayesian Machine. It comprises the main stochastic machine along with the True Random Generators (TRNG), responsible for the generation of the stochastic bit streams for the constants considered in the problem.

---

[0]A free version of ProBT is available at http://www.probayes.com/fr/Bayesian-Programming-Book/ and the version with soft evidence will be placed on www.bambi-fet.eu before the NIPS conference

The proposed tool-chain is working and accepts any ProBT program with discrete variables as entry. The tool-chain generates a VHDL file which is the description of the stochastic circuit and can be implemented on a FPGA. A Cyclone IV FPGA, from Altera has been targeted as supporting device. A machine has been synthesised to demonstrate the applicability and scalability of the proposed tool-chain. ProBT is also used to compute the exact result using standard arithmetic. This allows to evaluate the results given by FPGA with the synthesised VHDL program.

Figure 17 (right) shows the RTL generated by the synthesis tool, where it is possible to identify the connections between the components, corresponding to the circuit in Figure 17 (left). This circuit was implemented using 6 Logic Elements. The circuit was tested with bit streams integrated over 2^{31} to do the conversion from stochastic to binary.

The proposed tool-chain is working and accepts any ProBT program with discrete variables as entry. The tool-chain generates a VHDL file which is the description of the stochastic circuit and can be implemented on a FPGA. A Cyclone IV FPGA, from Altera has been targeted as supporting device. A machine has been synthesised to demonstrate the applicability and scalability of the proposed tool-chain.

ProBT is also used to compute the exact result using standard arithmetic. This allows to evaluate the results given by FPGA with the synthesised VHDL program.

Figure 17 (right) shows the RTL generated by the synthesis tool, where it is possible to identify the connections between the components, corresponding to the circuit in Figure 17 (left). This circuit was implemented using 6 Logic Elements. The circuit was tested with bit streams integrated over 2^{31} to do the conversion from stochastic to binary.

Figure 17. Stochastic circuit computing $\sum_{D1} \widetilde{P}(D_1)P(D_1|m)$ and the corresponding RTL.

We are now focusing on solving the time dilution problem by introducing memory in the architecture. Then we will make an attempt to build a filter with similar ideas by re-fitting the output into the initial joint distribution.

EXMO Project-Team

6. New Results

6.1. Highlights of the Year

- Our work on link key extraction and evaluation (§6.3.4) has been published at ECAI 2014.
- Jérôme Euzenat has been elected fellow of the European Coordination Committee for Artificial Intelligence (ECCAI).

6.2. Ontology matching and alignments

We pursue our work on ontology matching and alignment support [8] [10] with contributions to evaluation and alignment semantics.

6.2.1. *Evaluation*

Participant: Jérôme Euzenat.

Since 2004, we run the Ontology Alignment Evaluation Initiative (OAEI) which organises evaluation campaigns for assessing the degree of achievement of actual ontology matching algorithms [2].

This year, we ran the OAEI 2014 evaluation campaign [15]. We used again our generator for generating new version of benchmarks. The Alignment API was used for manipulating alignments and evaluating results.

A novelty of this year was that data interlinking evaluation was using the SEALS platform and a new query-based evaluation track was created.

The participating systems and evaluation results were presented in the 9th Ontology Matching workshop, held in Riva de Garda, Italy. More information on OAEI can be found at http://oaei.ontologymatching.org/.

6.2.2. *Algebras of alignment relations*

Participants: Armen Inants [Correspondent], Jérôme Euzenat.

Qualitative calculus is the central concept in qualitative binary constraint satisfaction problems. All formalisms developed so far are homogeneous – they assume a single universe. We had previously shown the advantages of using a homogeneous qualitative calculus for expressing ontology alignment relations between concepts. We tackle the problem of combining two or more calculi over disjoint universes into a single calculus. The problem is important, because in the ontology matching domain we deal with various kinds of ontological entities: concepts, individuals, properties. We define a new formalism called a heterogeneous qualitative calculus, based on an algebraic construct called Schröder category. A Schröder category is to binary relations over heterogeneous universes what a relation algebra is to homogeneous ones. We establish the connection between homogeneous and heterogeneous qualitative calculi by defining two mutually inverse transition operators. We provide an algorithm for combining two homogeneous calculi with different universes into a single calculus.

This work has vocation to support developments of the Alignment API towards relation algebras. It is part of the PhD of Armen Inants.

6.3. Data interlinking

The web of data uses semantic web technologies to publish data on the web in such a way that they can be interpreted and connected together. It is thus important to be able to establish links between these data, both for the web of data and for the semantic web that it contributes to feed. We consider this problem from different perspectives.

6.3.1. Interlinking cross-lingual RDF data sets

Participants: Tatiana Lesnikova [Correspondent], Jérôme David, Jérôme Euzenat.

RDF data sets are being published with labels that may be expressed in different languages. Even systems based on graph structure, ultimately rely on anchors based on language fragments. In this context, data interlinking requires specific approaches in order to tackle cross-lingualism. We proposed a general framework for interlinking RDF data in different languages and implemented two approaches: one approach is based on machine translation, the other one is based taking advantage of multilingual refernces, such as BabelNet. We evaluated variation of theses two settings on English (DBPedia) and Chinese (XLore) datasets. Both approaches demonstrated promising results [20]. We will conduct more experiments including other language pairs and larger corpus.

This work is part of the PhD of Tatiana Lesnikova developed in the LINDICLE project (§7.1.2).

6.3.2. Interactive learning of interlinking patterns

Participants: Zhengjie Fan [Correspondent], Jérôme Euzenat.

We proposed an interlinking method which, from class correspondences between data source ontologies, uses k-means or k-medoids clustering to produce property correspondences. It then generates a first interlinking pattern which is a combination of a link key and similarity measures. Such patterns can be transformed into a SILK script for generating an initial link set. A sample of these links are assessed by users as either correct or incorrect. These are taken as positive and negative example by an extension of the disjunctive version space method to find an interlinking pattern, that can justify correct links and incorrect links. Experiments show that, with only 1% of sample links, this method reaches a F-measure over 96%. The F-measure quickly converges, being improved by nearly 10% than other comparable approaches [19].

This work is part of the PhD of Zhengjie Fan [4], co-supervised with François Scharffe (LIRMM), and developed in the DATALIFT project (§7.1.1).

6.3.3. An iterative import-by-query approach to data interlinking

Participants: Manuel Atencia Arcas [Correspondent], Mustafa Al-Bakri, Steffen Lalande, Marie-Christine Rousset.

We modelled the problem of data interlinking as a reasoning problem on possibly decentralised data. We described an import-by-query algorithm that alternates steps of sub-query rewriting and of tailored querying of data sources. It only imports data as specific as possible for inferring or contradicting target sameAs assertions. Experiments conducted on a real-world dataset have demonstrated in practice the feasibility and usefulness of this approach for data interlinking and disambiguation purposes.

This work is part of the PhD thesis of Mustafa Al-Bakri, co-supervised by Manuel Atencia and Marie-Christine Rousset, developed in the QUALINCA project.

6.3.4. Link key extraction

Participants: Jérôme David [Correspondent], Manuel Atencia Arcas, Jérôme Euzenat.

Ontologies do not necessarily come with key descriptions, and never with link key assertions. Keys can be extracted from data by assuming that keys holding for specific data sets, may hold universally. We have extended such a classical key extraction technique for extracting weak link keys. We designed an algorithm to generate first a small set of candidate link keys and described this approach in the framework of formal context analysis [13]. Depending on whether some of the, valid or invalid, links are known, we defined supervised and non supervised measures for selecting the appropriate link keys. The supervised measures approximate precision and recall on a sample, while the non supervised measures are the ratio of pairs of entities a link key covers (coverage), and the ratio of entities from the same data set it identifies (discrimination). We have experimented these techniques, showing the accuracy and robustness of both approaches [12].

This work has been developed partly in the LINDICLE project (§7.1.2).

6.4. Dynamic aspects of networks of ontologies

Huge quantities of data described by ontologies and linked together are made available. These are generated in an independent manner by autonomous providers such as individuals or companies. They are heterogeneous and their joint exploitation requires connecting them, ending up as a mesh of reticulated knowledge.

However, data and knowledge have to evolve facing changes in what they represent, changes in the context in which they are used and connections to new data and knowledge sources. As their production and exchange are growing larger and more connected, their evolution is not anymore compatible with manual curation and maintenance. We work towards their continuous evolution as it is critical to their sustainability.

Two different approaches are currently explored.

6.4.1. *Evolution of ontology networks and linked data*

Participants: Adam Sanchez Ayte [Correspondent], Jérôme David, Jérôme Euzenat.

As link keys are obtained by statistical analysis of datasets (§6.3.4), they are both data-dependent and computation-intensive. Therefore, their recalculation should be avoided if possible. We are developing methods to analyse if changes performed in the data, necessarily require link key recomputation.

To reach this goal, we are developing an approach considering datasets as logical theories. In this context, changes that affect a link key are meta-logical operations. We adopt the framework of belief revision to define postulates that evolution operators must satisfy.

This work is part of the PhD thesis of Adam Sanchez Ayte developed in the LINDICLE project (§7.1.2).

6.4.2. *Cultural alignment repair*

Participant: Jérôme Euzenat [Correspondent].

Alignments between ontologies may be established through agents holding such ontologies attempting at communicating and taking appropriate action when communication fails. This approach, that we call cultural repair, has the advantage of not assuming that everything should be set correctly before trying to communicate and of being able to overcome failures. We tested this approach on alignment repair, i.e., the improvement of incorrect alignments. For that purpose, we performed a series of experiments in which agents react to mistakes in alignments. Agents only know about their ontologies and alignments with others and they act in a fully decentralised way. We showed that cultural repair is able to converge towards successful communication through improving the objective correctness of alignments. The obtained results are on par with a baseline of state-of-the-art alignment repair algorithms [7] [17]

The benchmarks, results and software are available at http://lazylav.gforge.inria.fr.

IMAGINE Project-Team

6. New Results

6.1. Highlights of the Year

- Vector Graphics Complexes, an new structure for 2D illustration developed in collaboration with UBC, resulted into a publication at ACM SIGGRAPH [4]. This superset of multi-layers graphics and of planar maps, enable intuitive design and deformation of 2D illustrations thanks to the separation of geometry from topology.

- Our work on elastic implicit skinning, a collaboration with U. Toulouse, Victoria University, and Inria Bordeaux was accepted at ACM SIGGRAPH Asia [16]. Thanks to robust iso-surface tracking, this method captures dynamic skin siding effects and can be used with extreme bending angles.

6.2. User-centered Models for Shapes

- **Scientist in charge**: Stefanie Hahmann.
- **Other permanent researchers**: Marie-Paule Cani, Jean-Claude Léon.

Our goal, is to develop responsive shape models, i.e. 3D models that respond in the expected way under any user action, by maintaining specific application-dependent constraints (such as a volumetric objects keeping their volume when bent, or cloth-like surfaces remaining developable during deformation, etc). We are extending this approach to composite objects made of distributions and/or combination of sub-shapes of various dimensions.

6.2.1. Implicit modeling

Participants: Antoine Bégault, Marie-Paule Cani, Michael Gleicher, Cédric Zanni.

Figure 4. Illustration from [17] showing some results of our N-ary implicit blends.

Our insight towards 3D shapes that respond in an intuitive way during both design and animation is to develop representations that clearly separate changes of structure - namely, the morphology of the shape - from changes of posture (its current 3D isometric embedding). Using skeletons is an excellent way to do so for 3D solids: the structure of a shape is represented by the topology of the skeleton, the length of its components and the shape thickness around it, while the shape posture is defined by the embedding of the skeleton in 3D space. Implicit surfaces (iso-surfaces of scalar fields) are the best mathematical model so far for generating 3D shapes from skeletons. However, a number of long standing problems - blending at distance that makes topology unpredictable, bulges at junctions, blurring of details - reduced the interest for this representation for many years. We addressed several of these issues in the last few years. Our most recent contribution is a method for enabling topology control in the case n-ary implicit blends [17]. Shapes are modeled using scale-invariant integral primitives (SCALIS) along skeletons, and blend with a plus. We use field warping to avoid unwanted blending and provide a unique control (based for instance on the angle) on the way skeleton-based primitives are allowed to blend. See Figure 4

6.2.2. *Towards responsive assemblies*

Participants: Stefanie Hahmann, Jean-Claude Léon, Aarohi Singh Johal.
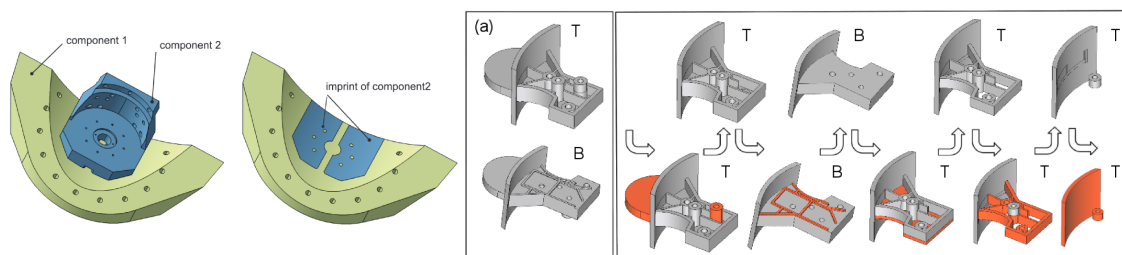
Figure 5. Illustration from [8] (Left) showing the automatic extraction of geometric interfaces, and from [2] (Right) showing the automatic generation of a construction graph.

We chose to focus on man-made objects to tackle the topic of shape assemblies, since CAD models of virtual industrial prototypes provide an excellent, real-size test-bed for our methods. Moreover, this is perfectly fitting the demand from industrial partners such as Airbus group and EDF.

Assemblies representing products are most often reduced to a collection of independent CAD models representing each component. The designation of each component and information about its function are often missing. As a result, geometric interfaces between components are unknown. These interfaces are particularly useful for structural mechanics to be able to quickly generate a Finite Element model of the assembly. This is especially critical when the latter gets very complex. [8] addresses the problem of automatically generating a class of geometric interfaces for very complex assemblies. GPU-based algorithms have proved suitable to obtain reliable results on CAD models.

Precisely determining interfaces between components is also a first requirement to enrich geometric models with functional information, since a subset of functions derives from interfaces between components. Based on both geometric interfaces and on a new concept of conventional interfaces, we proposed a series of approaches [13], [3] that make use of qualitative and ontology-based reasoning to connect CAD components and their geometric interfaces to functions or to functional designations of components: this results into an intrinsic identifier of a component in an assembly that connects it to its function.

To efficiently process assemblies of components, shape analysis [40] is particularly useful to generate the dimensionally reduced models needed for structural mechanics. [2] shows that analyzing a B-Rep CAD model to derive a construction graph, i.e. a set of construction trees, can be a robust basis to generate dimensionally reduced models.

Lastly, we extended shape analysis methods to detect some sets of symmetries [9]. Recovering this knowledge and embedding it into a model is the first step towards functionality-preserving deformations of complex man-made prototypes.

6.2.3. Parametric shapes

Participants: Stefanie Hahmann, Léo Allemand-Giorgis, Tibor Stanko.

Figure 6. Illustration from [38](Left) showing our results on monotonic interpolation, and from [1](Right) with our G1 interpolation surfaces for quad meshes.

We are developing new smooth parametric surface models defined on irregular quad meshes. They are in fact a powerful alternative to subdivision surface and singularly parameterized tensor product surfaces since they combine the advantages of both, the arbitrary topology of quad meshes and the smoothness of the tensor product patches. In collaboration with G.-P. Bonneau (Maverick team) several parametric triangular surface models for arbitrary topologies have been developed in the past. A new surface spline model has been published [1] and presented at GMP'14. It solves the problem of defining a G^1-continuous surface interpolating the vertices of an arbitrary quad mesh with low degree polynomial tensor product patches. It further aims to produce shapes of very high visual quality while reducing the number of control points, see Figure 6(right).

Another contribution concerns the modeling and smoothing of shapes using the Morse-Smale complex. The Morse-Smale complex is a topological structure defined on scalar functions which extracts critical points of the function and the links between them. By encoding a hierarchy between critical points, less important critical points can be deleted in order to simplify the structure. Our goal is to reconstruct a new shape, which corresponds to the simplified structure while approximating the initial data and preserving the most salient features. We first developed a method for interpolating monotone increasing 2D scalar data with a monotone piecewise cubic C^1-continuous surface. Monotonicity is a sufficient condition for a function to be free of critical points inside its domain. We overcome the restrictive standard axial monotonicity for tensor-product surfaces and introduce sufficient conditions and two algorithms for a more relaxed monotonicity constraint [38], see a piecewise monotonic shape in Figure 6(left). Then, some preliminary results on shape reconstruction from Morse-Smale complexes have been presented as a Posterand at a national conference [35].

In collaboration with Hans Hagen and Anne Berres from University of Kaiserslautern, we investigated conditions under which shape deformations preserve surface curvatures. The work has been published as a chapter in a scientific book [39].

6.3. Models for Motion Synthesis

- **Scientist in charge**: François Faure.
- **Other permanent researchers**: Marie-Paule Cani, Damien Rohmer, Rémi Ronfard.

Animating objects in real-time is mandatory to enable user interaction during motion design. Physically-based models, an excellent paradigm for generating motions that a human user would expect, tend to lack efficiency for complex shapes due to their use of low-level geometry (such as fine meshes). Our goal is therefore two-folds: first, develop efficient physically-based models and collision processing methods for arbitrary passive objects, by decoupling deformations from the possibly complex, geometric representation; second, study the combination of animation models with geometric responsive shapes, enabling the animation of complex constrained shapes in real-time. The last goal is to start developing coarse to fine animation models for virtual creatures, towards easier authoring of character animation for our work on narrative design.

6.3.1. Real-time physically-based models

Participants: Armelle Bauer, Ali Hamadi Dicko, François Faure, Matthieu Nesme.

Figure 7. Illustration from [14](Left) showing our velocity based adaptive simulation, and from [21](Right) on interactive visualization of muscle activity.

Following the success of frame-based elastic models (Siggraph 2011), a real-time animation framework provided in SOFA and currently used in many of our applications with external partners, we further improved this year the efficiency of this approach: we developed an adaptive version of frame-based elastic models, where frames get seamlessly attached to other ones during deformations when appropriate, in order to reduce computations [14], [33].

Frame-based models were successfully used to model limb movements in anatomical modeling [21]. The efficiency of this mehod enables us to advance towards the concept of a *Living book of anatomy*, where users move their own body and observe it through a tablet to get some visual illustration of anatomy in motion (see Figure 7).

6.3.2. Specific models for virtual creatures

Participants: Marie-Paule Cani, Michael Gleicher.

In collaboration with Loic Barthe and Rodolphe Vaillant from IRIT (U. Toulouse), Brian Wyvill (U. Victoria) and Gael Guennebaud (Manao, Inria), we developed a new automatic method for character skinning: Based on the approximation of character limbs with Hermite RBF implicit volumes, we adjust the mesh vertices representing the skin by projecting them back, at each animation step, to their iso-surface of interest. Since the vertices start from their previous position at the last animation step, there is no need of specifying skinning weights and using another skinning method as pre-computation, as in our previous implicit skinning method. Our solution avoids the well known blending artifacts of linear blend skinning and of dual quaternions, accommodates extreme blending angles and captures elastic effect in skin deformation [16].
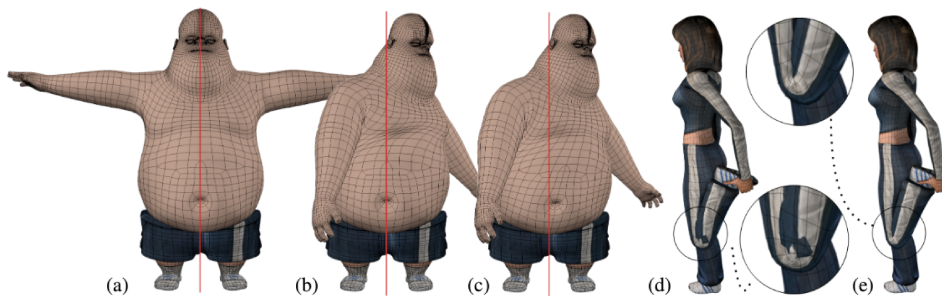
Figure 8. Illustration from [16] showing our results on dynamic skin deformations.

This year, we also studied the way character eyes and gazes are to be animated. This extensive study resulted into a state if the art report published at the Eurographics conference [32]

6.4. Knowledge-based Models for Narrative Design

- **Scientist in charge**: Rémi Ronfard.
- **Other permanent researchers**: Marie-Paule Cani, François Faure, Jean-Claude Léon, Olivier Palombi.

Our long term goal is to develop high-level models helping users to express and convey their own narrative content (from fiction stories to more practical educational or demonstrative scenarios). Before being able to specify the narration, a first step is to define models able to express some a priori knowledge on the background scene and on the object(s) or character(s) of interest. Our first goal is to develop 3D ontologies able to express such knowledge. The second goal is to define a representation for narration, to be used in future storyboarding frameworks and virtual direction tools. Our last goal is to develop high-level models for virtual cinematography such as rule-based cameras able to automatically follow the ongoing action and semi-automatic editing tools enabling to easily convey the narration via a movie.

6.4.1. Knowledge representation through 3D ontologies
Participants: Armelle Bauer, Jean-Claude Léon, Olivier Palombi.
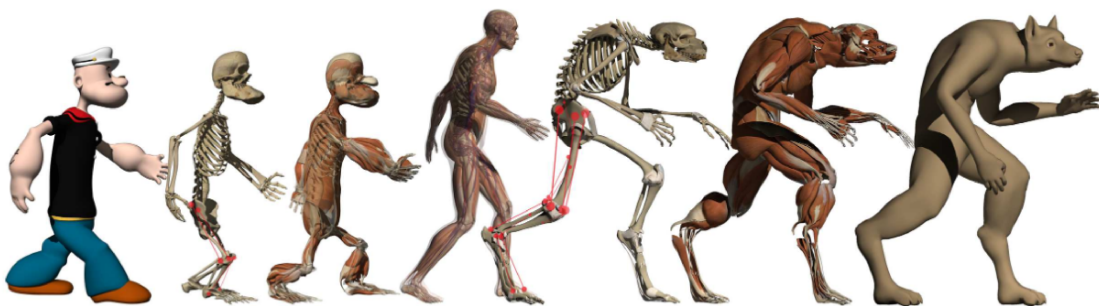
Figure 9. Illustration from [22] showing our results on ontology based anatomy generation.

We chose to develop 3D ontologies for being able to express combined knowledge on geometry, motion and function for assemblies or hierarchies of 3D objects. This is done in collaboration with a specialized group from the LIG laboratory in Grenoble. We decided to first focus these ontologies developments on two topics on which group members have a strong expertise: the anatomical domain (an interesting application test-bed for educational scenarios) and the industrial prototyping domain (where assembly scenarios can be defined).

We developed an anatomical knowledge database called My Corporis Fabrica (MyCF). We first linked functional entities defined in MyCF to the involved anatomical structures, using the musculoskeletal system as a test-bed. Based on this new formal description of the functional anatomy of limbs, we presented a novel pipeline for the construction of biomechanical simulations by combining generic anatomical knowledge with specific data which can handle complex reasoning and querying in MyCF. This resulted into a publication in the Journal of Biomedical Semantics [11]. We also used MyCF within our previous framework of anatomical transfer to set up an assistant tool for modeling and simulating anatomical structure such as bones, muscles, viscera and fat tissues easily while ensuring a correct anatomical consistency [22].

Secondly, in analysing the similarities and differences between existing ontology based describtion of products and virtual humans, we developed a common framework for combining 3D models and functional describtion to both models [15], [34].

6.4.2. Virtual direction tools

Participants: Adela Barbulescu, Rémi Ronfard.
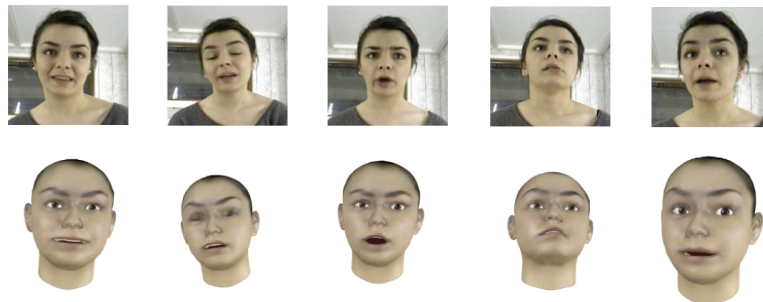
Figure 10. Illustration from [20] showing our results on expressive virtual actors.

We are developing a new approach to transfer speech signals and 3D facial expressions to virtual actors of a different identity. The converted sequences should be perceived as belonging to the target actors. This is the goal of Adela Barbulescu's thesis, co-advised by Gérard Bailly from GIPSA-lab. Our work started with conversion of speaking styles through speech signals only. This year, we started extending this approach to visual prosody and advanced on communicating social attitudes through head gestures [20].

6.4.3. Virtual cinematography

Participants: Quentin Galvane, Vineet Ghandi, Christophe Lino, Rémi Ronfard.

Our goal is to model automatic cameras for covering 3D scenes, as well as to develop semi-automatic film editing techniques to help conveying narration. This work was first conducted on video data, enabling us to test our ideas without the need for complex 3D movies: we designed an automatic method for the identification of actors in a video, and are using it for the automatic re-framing and editing high-resolution videos shots of theater rehearsals [25].

Figure 11. Illustration from [23] showing our results on narrative-driven camera control of computer games.

In parallel, we started extending this methodology to 3D animation, in collaboration with the Mimetic group in Rennes and with Geneva University: this year, we proposed a new method for replaying first person video games with automatic camera control based on the narration [23]. We also advanced towards semi-automatic film editing: A paper was just accepted to AAI 2015. To stress the difficulty of validating film editing methods, we devoted a specific work to validation methodologies [27].

We also addressed other issues related to cinematography and narratives: We designed a pre-visualization system for 3D cinematography to be used in the Action3DS project [30]: the method makes use of 3D modeling to show what the spectators watching a 3D movie are going to see, in order to ease 3D camera control by the film director. Lastly, we worked on computer generation of narrative discourses with the university of Geneva [31].

This year, Remi Ronfard and Vineet Gandhi wrote a patent application "Dispositif de génération de rushes cinématographiques par traitement vidéo", demande de brevet français no. 1460957, déposée le 13 novembre 2014.

6.5. Creating and Interacting with Virtual Prototypes

- **Scientist in charge**: Jean-Claude Léon.
- **Other permanent researchers**: Marie-Paule Cani, Olivier Palombi, Damien Rohmer, Rémi Ronfard.

The challenge is to develop more effective ways to put the user in the loop during content authoring. We generally rely on sketching techniques for quickly drafting new content, and on sculpting methods (in the sense of gesture-driven, continuous distortion) for further 3D content refinement and editing. The objective is to extend these expressive modeling techniques to general content, from complex shapes and assemblies to animated content. As a complement, we are exploring the use of various 2D or 3D input devices to ease interactive 3D content creation.

6.5.1. *Sketch-based modeling and editing of 3D shapes*
Participants: Marie-Paule Cani, Arnaud Emilien, Even Entem, Stefanie Hahmann, Rémi Ronfard.

While a lot of work has been done on sketch-based modeling of solid shapes, only a few methods do tackle surface models. Terrain surfaces are particular challenging: their fractal-like distribution of details makes them easy to identify, but these cannot be fully drawn by a user. In our work, users only need to draw the main silhouettes they would like to see from a first person viewpoint (enabling, for instance, an art director to set the background scene behind his actors). We generate a plausible, complex terrain that matches the sketch by deforming an existing terrain model. This is done by analyzing the complex silhouettes with cups and T-junctions in the input sketch and matching them with perceptually close features of the input terrain. The rest
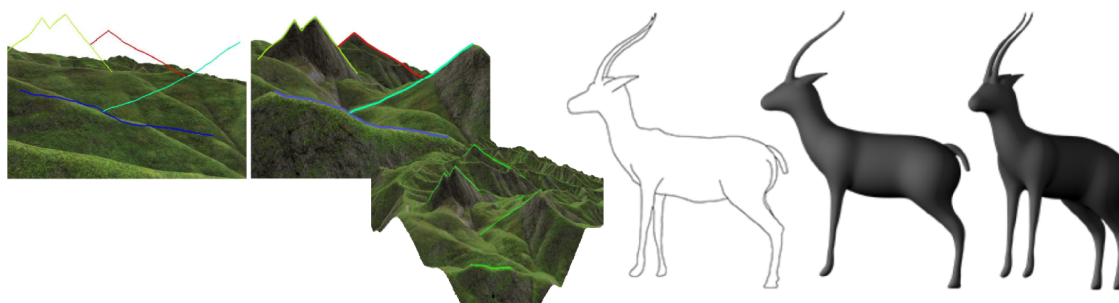
Figure 12. Illustration from [29](Left) showing our results on terrain editing, and from [6](Right) on 3D animal modeling from sketch.

of the terrain is seamlessly deformed while keeping its visual complexity and style. This work was presented at Graphics Interface 2014 [29] and extended to enable the combination of silhouettes from multiple viewpoints in the Computer and Graphics journal [12].

In collaboration with UBC, we introduced the vector graphics complex (VGC), a simple data structure that supports non-manifold topological modeling for 2D vector graphics illustrations. The representation faithfully captures the intended semantics of a wide variety of illustrations, and is a proper superset of scalable vector graphics and planar map representations. VGC nearly separates the geometry of vector graphics objects from their topology, making it easy to deform objects in interesting and intuitive ways, a premise for enabling their animation. This work was published at SIGGRAPH 2014 [4]. We also developed a method for generating 3D animals from a single sketch. This method takes a complex sketch with cups and T-junctions as input (see Figure 12 , and makes use of symmetry hypotheses to analyze it into regions corresponding to the main body and to front and back limbs. The different regions are then automatically reconstructed and blended together using on our implicit modeling methodologies (SCALIS surfaces).

Lastly, we designed a sketch-based interface for authoring illustrative animations. The method makes use of hierarchical motion brushes, a new concept for specifying complex hierarchical motion with a few strokes [28].

6.5.2. *Sketching and Sculpting Motion*

Participants: Marie-Paule Cani, Arnaud Emilien, Kevin Jordao.

We extended sculpting methods, which had been restricted so far to homogeneous geometric models of a single dimension, to the handling of complex structured shapes and to the interactive sculpting of animated environments.

We developed the first method enabling to sculpt animated content in extending our previous elastic mutable model approach. Relying on the crowd patches representation for modeling animated crowds, we extended component mutations to space-time content, enabling a user to stretch, bend or assemble populated streets while ensuring that individual character trajectories remain continuous through space and time, as well as plausible. This work, developed within Kevin Jordao's thesis co-advised with Julien Pettré from the Mimetic project-team in Rennes, was published at Eurographics 2014 [7].

We developed an interactive system for designing complex waterfall scenes: vector elements created by the user (contacts, freefalls and pools) are used to control the procedural creation of complex waterfalls and rivers that match the user intend while ensuring coherent flows and good embedding within the terrain [18].

6.5.3. *Interaction devices and gestural patterns*

Participants: Marie-Paule Cani, Rémi Brouet.
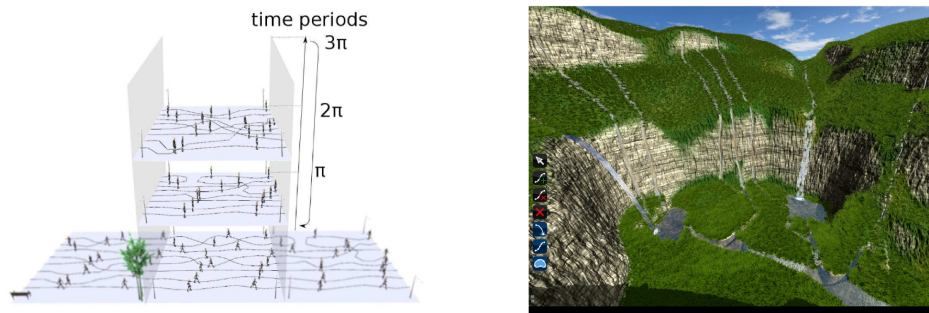
Figure 13. Illustration from [7](Left) on our crowd sculpting interface, and from [18](Right) on interactive design of complex waterfall scenes.

Our work on gestural interaction patterns for 3D design has been developed in collaboration with the HCI team from LIG laboratory towards the exploration of 2D multi-touch tables for the placement and deformation of 3D models.

We are also exploring the use of multi-touch tables for the interactive design and editing of 3D scenes. This is the topic of Rémi Brouet's PhD thesis, co-advised with Renaud Blanch from the human-computer interaction group of LIG laboratory. The main challenge here is to find out how to use 2D interaction media for editing 3D content, hence how to intuitively control the third dimension (depth, non-planar rotations, 3D deformations, etc). Our work on this topic started with a preliminary user study enabling us to analyze all possible hand interaction gestures on table-tops, and to explore the ways users would intuitively try to manipulate 3D environments, either for changing the camera position or for moving objects around. We extracted a general interaction pattern from this study. Our implementation enables both seamless navigation and docking in 3D scenes, without the need for any menu or button to change mode. We are currently extending this work to object editing scenarios, where shapes can be bent or twisted in 3D using 2D interaction and automatic mode selection only.

LEAR Project-Team

6. New Results

6.1. Highlights of the Year

- Cordelia Schmid received the **Longuet-Higgins prize** for fundamental contributions in computer vision that have withstood the test of time, 2014.
- We participated to the **Trecvid 2014 Multimedia Event Detection** challenge. We ranked first on one of the four tracks (Ad-hoc training videos with 10 examples per class).
- We participated to the **THUMOS 2014 challenge**. We obtained top ranked results in the localization track of the Thumos 2014 Action Recognition Challenge. The goal of the challenge is to evaluate large-scale action recognition in natural settings.

6.2. Visual recognition in images

6.2.1. *Multi-fold MIL Training for Weakly Supervised Object Localization*

Participants: Ramazan Cinbis, Cordelia Schmid, Jakob Verbeek.

Object category localization is a challenging problem in computer vision. Standard supervised training requires bounding box annotations of object instances. This time-consuming annotation process is sidestepped in weakly supervised learning. In this case, the supervised information is restricted to binary labels that indicate the absence/presence of object instances in the image, without their locations. In [13], we follow a multiple-instance learning approach that iteratively trains the detector and infers the object locations in the positive training images. Our main contribution is a multi-fold multiple instance learning procedure, which prevents training from prematurely locking onto erroneous object locations. This procedure is particularly important when high-dimensional representations, such as the Fisher vectors, are used. We present a detailed experimental evaluation using the PASCAL VOC 2007 and 2010 datasets. Compared to state-of-the-art weakly supervised detectors, our approach better localizes objects in the training images, which translates into improved detection performance. Figure 1 illustrates the iterative object localization process on several example images.

A journal paper is currently in preparation in which extends [13] by adding experiments with CNN features, and a refinement procedure for the object location inference. These additions improve over related work that has appeared since the publication of the original paper.

6.2.2. *Transformation Pursuit for Image Classification*

Participants: Mattis Paulin, Jerome Revaud, Zaid Harchaoui, Florent Perronnin [XRCE], Cordelia Schmid.

In this work [19], [23], we use data augmentation (see Fig 2 for examples) to improve image classification performances in a large-scale context. A simple approach to learning invariances in image classification consists in augmenting the training set with transformed versions of the original images. However, given a large set of possible transformations, selecting a compact subset is challenging. Indeed, all transformations are not equally informative and adding uninformative transformations increases training time with no gain in accuracy. We propose a principled algorithm – Image Transformation Pursuit (ITP) – for the automatic selection of a compact set of transformations. ITP works in a greedy fashion, by selecting at each iteration the one that yields the highest accuracy gain. ITP also allows to efficiently explore complex transformations, that combine basic transformations. We report results on two public benchmarks: the CUB dataset of bird images and the ImageNet 2010 challenge. Using Fisher Vector representations, we achieve an improvement from 28.2% to 45.2% in top-1 accuracy on CUB, and an improvement from 70.1% to 74.9% in top-5 accuracy on ImageNet. We also show significant improvements for deep convnet features: from 47.3% to 55.4% on CUB and from 77.9% to 81.4% on ImageNet.
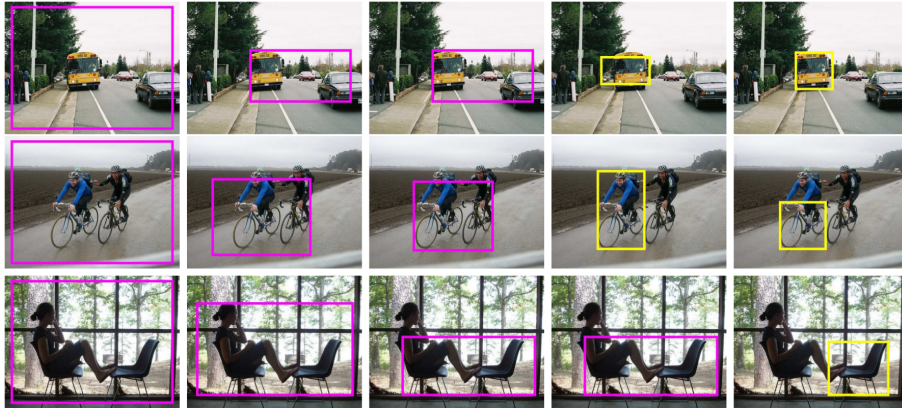
Figure 1. Illustrattion of our iterative object localization process on several example images, from initialization (left) to final localization (right). Yellow bounding boxes indicate that the object location hypothesis is in agreement with the ground-truth, for pink boxes the hypothesis is incorrect.
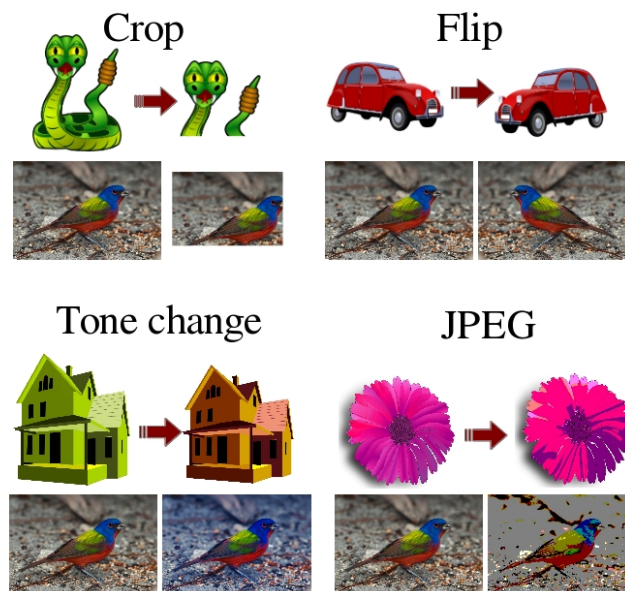

Figure 2. Examples of transformations used in [19], [23].

6.2.3. Convolutional Kernel Networks

Participants: Julien Mairal, Piotr Koniusz, Zaid Harchaoui, Cordelia Schmid.

An important goal in visual recognition is to devise image representations that are invariant to particular transformations. In this paper [16] we address this goal with a new type of convolutional neural network (CNN) whose invariance is encoded by a reproducing kernel. Unlike traditional approaches where neural networks are learned either to represent data or for solving a classification task, our network learns to approximate the kernel feature map on training data. Such an approach enjoys several benefits over classical ones. First, by teaching CNNs to be invariant, we obtain simple network architectures that achieve a similar accuracy to more complex ones, while being easy to train and robust to overfitting. Second, we bridge a gap between the neural network literature and kernels, which are natural tools to model invariance. We evaluate our methodology on visual recognition tasks where CNNs have proven to perform well, e.g., digit recognition with the MNIST dataset, and the more challenging CIFAR-10 and STL-10 datasets, where our accuracy is competitive with the state of the art. Figure 3 illustrates the architecture of our network.

(a) Hierarchy of image feature maps. (b) Zoom between layer $k-1$ and k of the CKN.
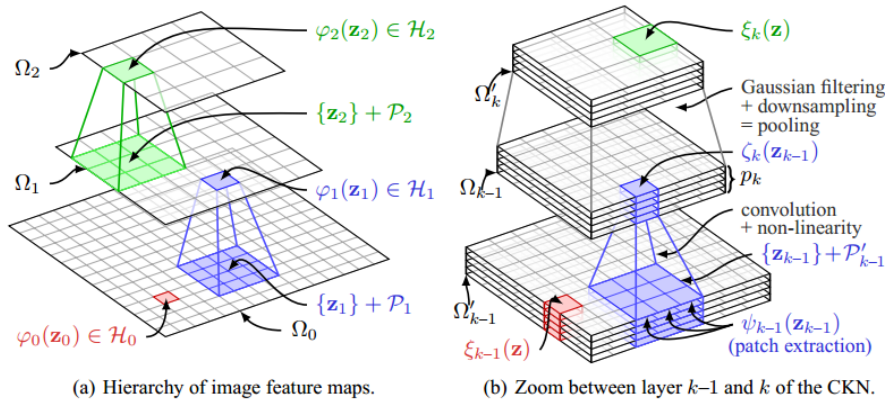
Figure 3. Left: concrete representation of the successive layers for the multilayer convolutional kernel. Right: one layer of the convolutional neural network that approximates the kernel.

6.2.4. Scene Text Recognition and Retrieval for Large Lexicons

Participants: Udit Roy [IIIT Hyderabad, India], Anand Mishra [IIIT Hyderabad, India], Karteek Alahari, C. v. Jawahar [IIIT Hyderabad, India].

In [21], we propose a framework for recognition and retrieval tasks in the context of scene text images. In contrast to many of the recent works, we focus on the case where an image-specific list of words, known as the small lexicon setting, is unavailable. We present a conditional random field model defined on potential character locations and the interactions between them. Observing that the interaction potentials computed in the large lexicon setting are less effective than in the case of a small lexicon, we propose an iterative method, which alternates between finding the most likely solution and refining the interaction potentials. We evaluate our method on public datasets and show that it improves over baseline and state-of-the-art approaches. For example, we obtain nearly 15% improvement in recognition accuracy and precision for our retrieval task over baseline methods on the IIIT-5K word dataset, with a large lexicon containing 0.5 million words.

6.2.5. On Learning to Localize Objects with Minimal Supervision

Participants: Hyun On Song [UC Berkeley], Ross Girschick [UC Berkeley], Stefanie Jegelka [UC Berkeley], Julien Mairal, Zaid Harchaoui, Trevor Darrell [UC Berkeley].

Learning to localize objects with minimal supervision is an important problem in computer vision, since large fully annotated datasets are extremely costly to obtain. In this paper [22], we propose a new method that achieves this goal with only image-level labels of whether the objects are present or not. Our approach combines a discriminative submodular cover problem for automatically discovering a set of positive object windows with a smoothed latent SVM formulation. The latter allows us to leverage efficient quasiNewton optimization techniques. Experimental results are presented in Figure 4 .

Figure 4. Visualization of some common failure cases of constructed positive windows by (Siva et al., 2012) vs our method. Red bounding boxes are constructed positive windows from (Siva et al., 2012). Green bounding boxes are constructed positive windows from our method.

6.2.6. Good Practice in Large-Scale Learning for Image Classification

Participants: Zeynep Akata, Florent Perronnin [XRCE], Zaid Harchaoui, Cordelia Schmid.

In this paper [3], we benchmark several SVM objective functions for large-scale image classification. We consider one-vs-rest, multi-class, ranking, and weighted approximate ranking SVMs. A comparison of online and batch methods for optimizing the objectives shows that online methods perform as well as batch methods in terms of classification accuracy, but with a significant gain in training speed. Using stochastic gradient descent, we can scale the training to millions of images and thousands of classes. Our experimental evaluation shows that ranking-based algorithms do not outperform the one-vs-rest strategy when a large number of training examples are used. Furthermore, the gap in accuracy between the different algorithms shrinks as the dimension of the features increases. We also show that learning through cross-validation the optimal rebalancing of positive and negative examples can result in a significant improvement for the one-vs-rest strategy. Finally, early stopping can be used as an effective regularization strategy when training with online algorithms. Following these "good practices", we were able to improve the state-of-the-art on a large subset of 10K classes and 9M images of ImageNet from 16.7% Top-1 accuracy to 19.1%.

6.3. Learning and statistical models

6.3.1. Fast and Robust Archetypal Analysis for Representation Learning

Participants: Yuansi Chen, Julien Mairal, Zaid Harchaoui.

In [9], we revisit a pioneer unsupervised learning technique called archetypal analysis, which is related to successful data analysis methods such as sparse coding and non-negative matrix factorization. Since it was proposed, archetypal analysis did not gain a lot of popularity even though it produces more interpretable models than other alternatives. Because no efficient implementation has ever been made publicly available, its application to important scientific problems may have been severely limited. Our goal is to bring back into favour archetypal analysis. We propose a fast optimization scheme using an active-set strategy, and provide

an efficient open-source implementation interfaced with Matlab, R, and Python. Then, we demonstrate the usefulness of archetypal analysis for computer vision tasks, such as codebook learning, signal classification, and large image collection visualization.

In Figure 5 , we present some archetypes corresponding to the request "Paris" when downloading 36 600 images uploaded in 2012 and 2013, and sorted by relevance on the Flickr website.

(a) Archetypes representing (expected) landmarks.   (b) Less expected archetypes.   (c) Archetypes representing scene composition.

Figure 5. Classical landmarks appear on the left, which is not surprising since Flickr contains a large number of vacation pictures. In the middle, we display several archetypes that we did not expect, including ones about soccer, graffitis, food, flowers, and social gatherings. Finally, we display on the right some archetypes that do not seem to have some semantic meaning, but they capture some scene composition or texture that are common in the dataset.

6.3.2. Conditional Gradient Algorithms for Norm-Regularized Smooth Convex Optimization

Participants: Zaid Harchaoui, Anatoli Juditsky, Arkadii Nemirovski.

In this paper [6], we consider convex optimization problems arising in machine learning in high-dimensional settings. For several important learning problems, such as e.g. noisy matrix completion, state-of-the-art optimization approaches such as composite minimization algorithms are difficult to apply and do not scale up to large datasets. We study three conditional gradient-type algorithms, *i.e.* first-order optimization algorithms that require a linear minimization oracle but do not require a proximal oracle. These new algorithms are suitable for large-scale problems, and enjoy finite-time convergence guarantees. Promising experimental results are presented on two large-scale real-world datasets. The method is illustrated in Figure 6 .

6.3.3. A Smoothing Approach for Composite Conditional Gradient with Nonsmooth Loss

Participants: Federico Pierucci, Zaid Harchaoui, Jérôme Malick [BIPOP Team, Inria].

In [25], we consider learning problems where the nonsmoothness lies both in the convex empirical risk and in the regularization penalty. Examples of such problems include learning with nonsmooth loss functions and atomic decomposition regularization penalty. Such doubly nonsmooth learning problems prevent the use of recently proposed composite conditional gradient algorithms for training, which are particularly attractive for large-scale applications. Indeed, they rely on the assumption that the empirical risk part of the objective is smooth. We propose a composite conditional gradient algorithm with smoothing to tackle such learning

1. Initialize : $z^{(0)} = [\mathbf{0}, \overline{R}]$, with $\overline{R} > 0$ s.t. $\|x^\star\|_{\mathcal{A}} \le \overline{R}$
2. for $t = 0, \dots, T$ do
   – Compute $\bar{z}^{(t)} = \mathrm{LMO}(\nabla F(z^{(t)})$
   – Update
   $$z^{(t+1)} := \underset{z \in \mathcal{C}^{(t)}}{\arg\min}\ F(z)$$
   $$\mathcal{C}^{(t)} := \mathrm{conv}\{0, z^{(t)}, \overline{R} \cdot \bar{z}^{(t)}\}$$
3. end

$$\overline{z}^{(t)} := \overline{R} \cdot [\nabla f(\boldsymbol{x}^{(t)}), 1]$$

$[\nabla f(\boldsymbol{x}^{(t)}), 1]$

$z^{(t)}$

$\mathbf{0}$
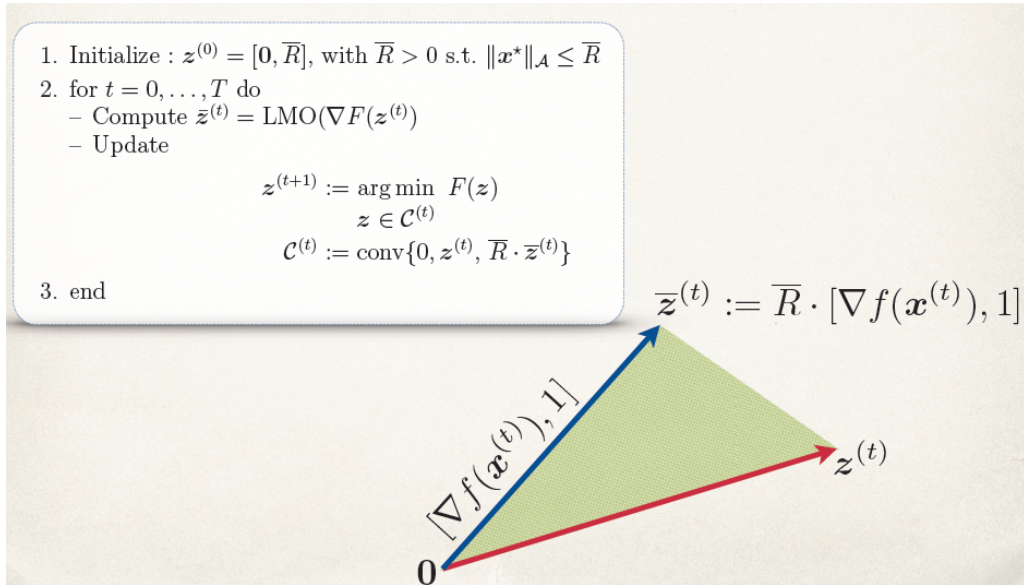
Figure 6. Overview of the composite conditional gradient algorithm which minimizes $F(x) := f(x) + \lambda\|x\|_{\mathcal{A}}$, where f is smooth and $\|\cdot\|_{\mathcal{A}}$ is an atomic-decomposition norm.

problems. We set up a framework allowing to systematically design parametrized smooth surrogates of nonsmooth loss functions. We then propose a smoothed composite conditional gradient algorithm, for which we prove theoretical guarantees on the accuracy. We present promising experimental results on collaborative filtering tasks (see Figure 7).

$\le \frac{\mu}{2}$
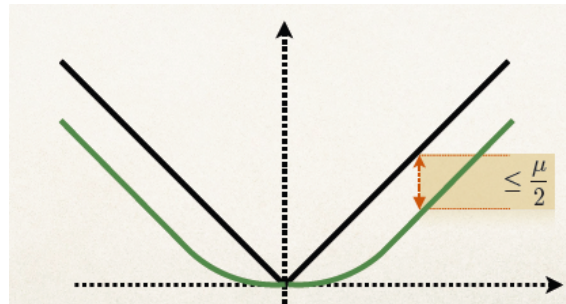
Figure 7. Illustration of the smooth surrogate with parameter μ (green) of the absolute value function (black).

6.3.4. Incremental Majorization-Minimization Optimization with Application to Large-Scale Machine Learning

Participant: Julien Mairal.

In this paper [27], we study optimization methods consisting of iteratively minimizing surrogates of an objective function, as illustrated in Figure 8 . We introduce a new incremental scheme that experimentally matches or outperforms state-of-the-art solvers for large-scale optimization problems typically arising in machine learning.
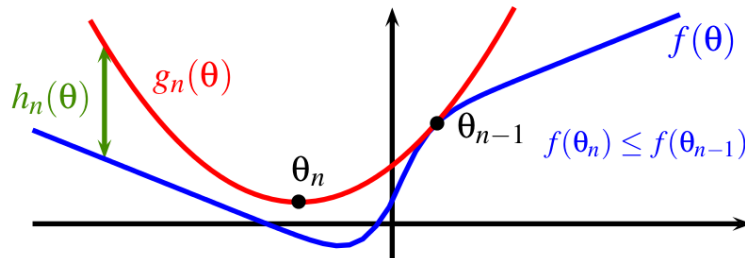
Figure 8. Illustration of the basic majorization-minimization principle. We compute a surrogate g_n of the objective function f around a current estimate θ_{n-1}. The new estimate θ_n is a minimizer of g_n. The approximation error h_n is smooth.

6.3.5. Efficient RNA Isoform Identification and Quantification from RNA-Seq Data with Network Flows

Participants: Elsa Bernard [Institut Curie, Ecoles des Mines-ParisTech], Laurent Jacob [CNRS, LBBE Laboratory], Julien Mairal [correspondant], Jean-Philippe Vert [Institut Curie, Ecoles des Mines-ParisTech].

Several state-of-the-art methods for isoform identification and quantification are based on ℓ_1-regularized regression, such as the Lasso. However, explicitly listing the—possibly exponentially—large set of candidate transcripts is intractable for genes with many exons. For this reason, existing approaches using the ℓ_1-penalty are either restricted to genes with few exons or only run the regression algorithm on a small set of preselected isoforms. In [4], we introduce a new technique called FlipFlop, which can efficiently tackle the sparse estimation problem on the full set of candidate isoforms by using network flow optimization. Our technique removes the need of a preselection step, leading to better isoform identification while keeping a low computational cost. Experiments with synthetic and real RNA-Seq data confirm that our approach is more accurate than alternative methods and one of the fastest available. Figure 9 presents the graph on which the network flow optimization is performed.

6.3.6. Riemannian Sparse Coding for Positive Definite Matrices

Participants: Anoop Cherian, Suvrit Sra [MPI].

Inspired by the great success of sparse coding for vector valued data, our goal in this work [12] is to represent symmetric positive definite (SPD) data matrices as sparse linear combinations of atoms from a dictionary, where each atom itself is an SPD matrix. Since SPD matrices follow a non-Euclidean (in fact a Riemannian) geometry, existing sparse coding techniques for Euclidean data cannot be directly extended. Prior works have approached this problem by defining a sparse coding loss function using either extrinsic similarity measures (such as the log-Euclidean distance) or kernelized variants of statistical measures (such as the Stein divergence, Jeffrey's divergence, etc.). In contrast, we propose to use the intrinsic Riemannian distance on the manifold of SPD matrices. Our main contribution is a novel mathematical model for sparse coding of SPD matrices; we also present a computationally simple algorithm for optimizing our model. Experiments on several computer vision datasets showcase superior classification and retrieval performance compared against state-of-the-art approaches.
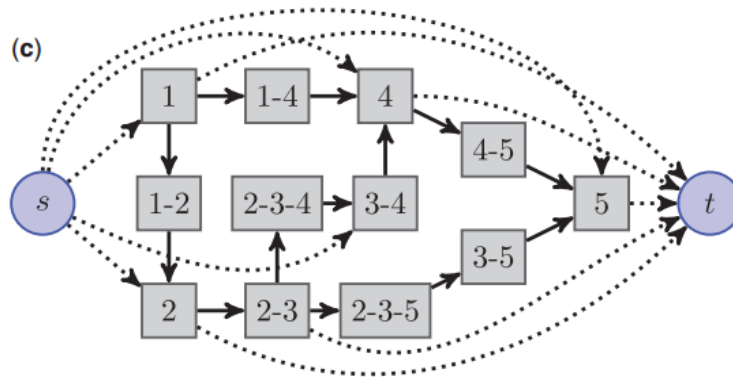
Figure 9. Graph on which we perform network flow optimization. Nodes represent observed reads, and paths on the graph correspond to isoforms.

6.4. Recognition in video

6.4.1. *Occlusion and Motion Reasoning for Long-Term Tracking*

Participants: Yang Hua, Karteek Alahari, Cordelia Schmid.

Object tracking is a reoccurring problem in computer vision. Tracking-by-detection approaches, in particular Struck, have shown to be competitive in recent evaluations. However, such approaches fail in the presence of long-term occlusions as well as severe viewpoint changes of the object. In this paper we propose a principled way to combine occlusion and motion reasoning with a tracking-by-detection approach. Occlusion and motion reasoning is based on state-of-the-art long-term trajectories which are labeled as object or background tracks with an energy-based formulation. The overlap between labeled tracks and detected regions allows to identify occlusions. The motion changes of the object between consecutive frames can be estimated robustly from the geometric relation between object trajectories. If this geometric change is significant, an additional detector is trained. Experimental results show that our tracker obtains state-of-the-art results and handles occlusion and viewpoints changes better than competing tracking methods. This work corresponds to the publication [15] and is illustrated in Figure 10 .

6.4.2. *Category-Specific Video Summarization*

Participants: Danila Potapov, Matthijs Douze, Zaid Harchaoui, Cordelia Schmid.

In large video collections with clusters of typical categories, such as "birthday party" or "flash-mob", category-specific video summarization can produce higher quality video summaries than unsupervised approaches that are blind to the video category. Given a video from a known category, our approach published in [20] first efficiently performs a temporal segmentation into semantically-consistent segments, delimited not only by shot boundaries but also general change points. Then, equipped with an SVM classifier, our approach assigns importance scores to each segment. The resulting video assembles the sequence of segments with the highest scores, as shown in Figure 11 . The obtained video summary is therefore both short and highly informative. Experimental results on videos from the multimedia event detection (MED) dataset of TRECVID'11 show that our approach produces video summaries with higher relevance than the state of the art.

6.4.3. *Efficient Action Localization with Approximately Normalized Fisher Vectors*

Participants: Dan Oneata, Jakob Verbeek, Cordelia Schmid.
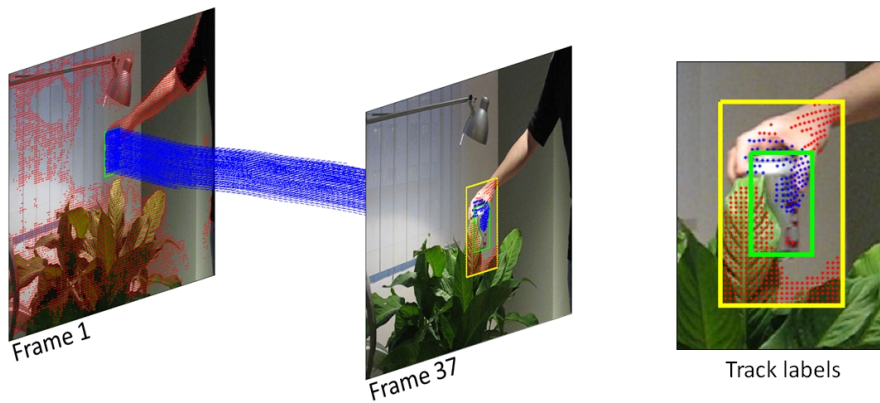
Figure 10. Left: Long-term tracks beginning in frame 1 of the Coke sequence. The yellow box shows the search region used to compute the bounding box most likely to contain the object (green box). We use the tracks to estimate the object state. Right: Close-up of the track labels in frame 37. Here, less than 60% of the tracks within the predicted bounding box are assigned to the object (blue), and the remaining are labelled as background (red). Thus, the object is predicted to be in an occluded state.
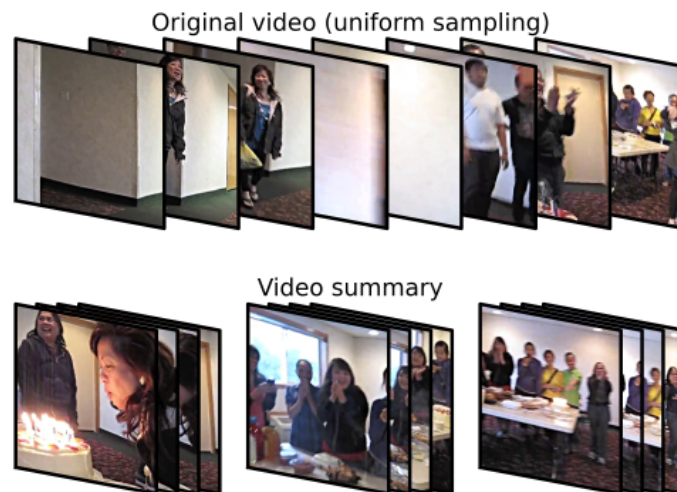
Figure 11. Original video, and its video summary for the category "birthday party".

The Fisher vector (FV) representation is a high-dimensional extension of the popular bag-of-word representation. Transformation of the FV by power and ℓ_2 normalizations has shown to significantly improve its performance, and led to state-of-the-art results for a range of image and video classification and retrieval tasks. These normalizations, however, render the representation non-additive over local descriptors. Combined with its high dimensionality, this makes the FV computationally expensive for the purpose of localization tasks. In [18] we present approximations to both these normalizations (see Figure 12), which yield significant improvements in the memory and computational costs of the FV when used for localization. Second, we show how these approximations can be used to define upper-bounds on the score function that can be efficiently evaluated, which enables the use of branch-and-bound search as an alternative to exhaustive sliding window search. We present experimental evaluation results on classification and temporal localization of actions in videos. These show that the our approximations lead to a speedup of at least one order of magnitude, while maintaining state-of-the-art action recognition and localization performance.
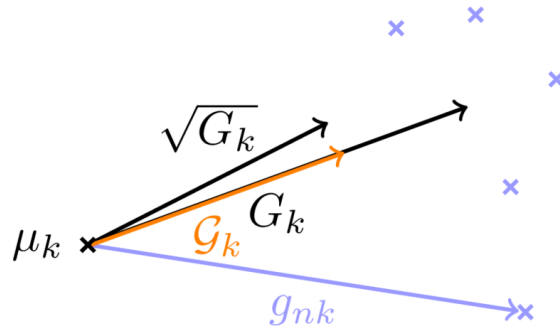
Figure 12. Schematic illustration of the proposed approximation for the square-root normalization. We depict a Fisher vector G_k as an aggregation of individual gradients g_{nk}. Both the exact ($\sqrt{G_k}$) and the approximated ($\mathcal{G}_k$) square-root normalizations scale similarly the Fisher vector G_k; the approximated variant has the property of preserving the orientation of the Fisher vector G_k.

6.4.4. Spatio-Temporal Object Detection Proposals

Participants: Dan Oneata, Jakob Verbeek, Cordelia Schmid, Jerome Revaud.

Spatio-temporal detection of actions and events in video is a challenging problem. Besides the difficulties related to recognition, a major challenge for detection in video is the size of the search space defined by spatio-temporal tubes formed by sequences of bounding boxes along the frames. Recently methods that generate unsupervised detection proposals have proven to be very effective for object detection in still images. These methods open the possibility to use strong but computationally expensive features since only a relatively small number of detection hypotheses need to be assessed. In [17] we make two contributions towards exploiting detection proposals for spatio-temporal detection problems. First, we extend a recent 2D object proposal method, to produce spatio-temporal proposals by a randomized supervoxel merging process (see Figure 13). We introduce spatial, temporal, and spatio-temporal pairwise supervoxel features that are used to guide the merging process. Second, we propose a new efficient supervoxel method. We experimentally evaluate our detection proposals, in combination with our new supervoxel method as well as existing ones. This evaluation shows that our supervoxels lead to more accurate proposals when compared to using existing state-of-the-art supervoxel methods.

6.4.5. EpicFlow: Edge-Preserving Interpolation of Correspondences for Optical Flow

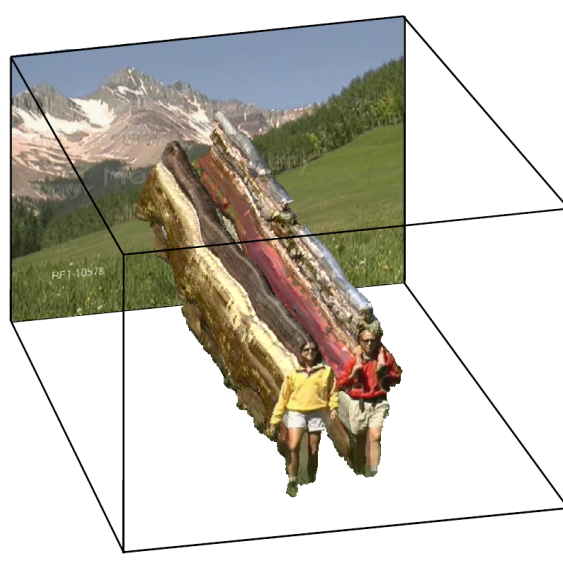**Participants:** Revaud Jerome, Weinzaepfel Philippe, Harchaoui Zaid, Cordelia Schmid.

Figure 13. An example of a spatio-temporal proposal generated by our randomized supervoxel merging process. The video sample belongs to the UCF Sports dataset.

We propose a novel approach [29] for optical flow estimation, targeted at large displacements with significant occlusions. It consists of two steps: i) dense matching by edge-preserving interpolation from a sparse set of matches; ii) variational energy minimization initialized with the dense matches. The sparse-to-dense interpolation relies on an appropriate choice of the distance, namely an edge-aware geodesic distance. This distance is tailored to handle occlusions and motion boundaries (see Figure 14), two common and difficult issues for optical flow computation. We also propose an approximation scheme for the geodesic distance to allow fast computation without loss of performance. Subsequent to the dense interpolation step, standard one-level variational energy minimization is carried out on the dense matches to obtain the final flow estimation. The proposed approach, called Edge-Preserving Interpolation of Correspondences (EpicFlow) is fast and robust to large displacements. It significantly outperforms the state of the art on MPI-Sintel and performs on par on KITTI and Middlebury.

6.4.6. *Weakly Supervised Action Labeling in Videos Under Ordering Constraints.*

Participants: Piotr Bojanowski [Willow team, Inria], Rémi Lajugie [Willow team, Inria], Francis Bach [Sierra team, Inria], Ivan Laptev [Willow team, Inria], Jean Ponce [Willow team, Inria], Cordelia Schmid, Josef Sivic [Willow team, Inria].

Suppose we are given a set of video clips, each one annotated with an ordered list of actions, such as "walk" then "sit" then "answer phone" extracted from, for example, the associated text script. See Fig. 15 for an illustration. In this work [8], we seek to temporally localize the individual actions in each clip as well as to learn a discriminative classifier for each action. We formulate the problem as a weakly supervised temporal assignment with ordering constraints. Each video clip is divided into small time intervals and each time interval of each video clip is assigned one action label, while respecting the order in which the action labels appear in the given annotations. We show that the action label assignment can be determined together with learning a classifier for each action in a discriminative manner. We evaluate the proposed model on a new and challenging dataset of 937 video clips with a total of 787720 frames containing sequences of 16 different actions from 69 Hollywood movies.
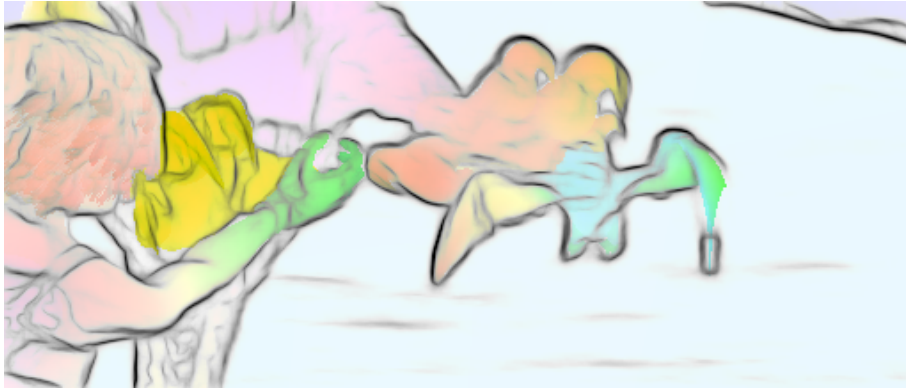
Figure 14. Image edges detected with SED and ground-truth optical flow. Motion discontinuities appear most of the time at image edges
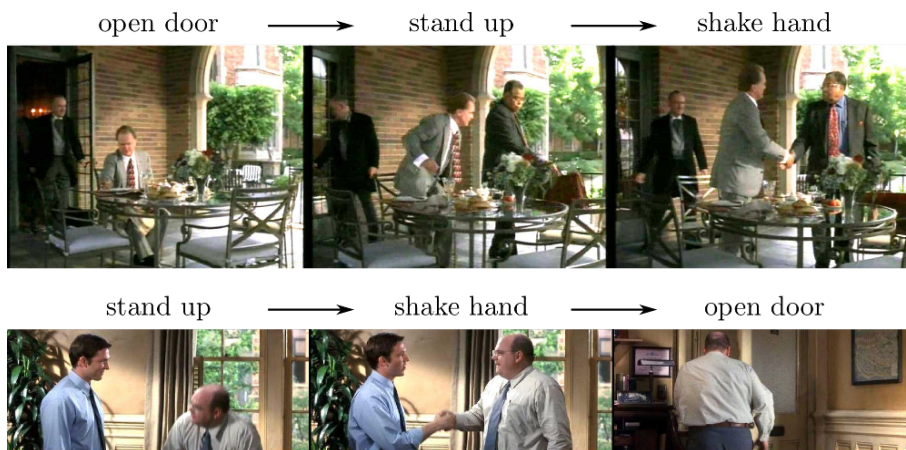

Figure 15. Sample data used as input to our method. Every video clip comes with an ordered list of actions that appears in it. These actions are not temporaly localized, only the order is known. The goal of our paper is to correctly localize these actions according to a discriminative criterion.

6.4.7. *Mixing Body-Part Sequences for Human Pose Estimation*

Participants: Cherian Anoop, Mairal Julien, Alahari Karteek, Schmid Cordelia.

This work [11] presents a method for estimating articulated human poses in videos. We cast this as an optimization problem defined on body parts with spatio-temporal links between them. The resulting formulation is unfortunately intractable and previous approaches only provide approximate solutions. Although such methods perform well on certain body parts, e.g., head, their performance on lower arms, i.e., elbows and wrists, remains poor. We present a new approximate scheme with two steps dedicated to pose estimation. First, our approach takes into account temporal links with subsequent frames for the less-certain parts, namely elbows and wrists. Second, our method decomposes poses into limbs, generates limb sequences across time, and recomposes poses by mixing these body part sequences (See Figure 16   for an illustration). We introduce a new dataset "Poses in the Wild", which is more challenging than the existing ones, with sequences containing background clutter, occlusions, and severe camera motion. We experimentally compare our method with recent approaches on this new dataset as well as on two other benchmark datasets, and show significant improvement.
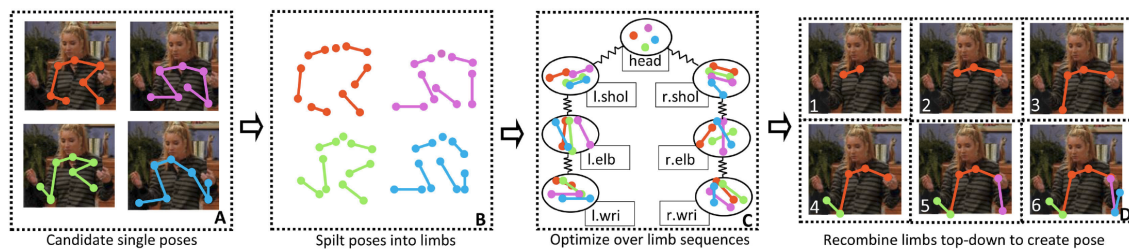
Figure 16. Illustration of our limb recombination scheme. From left to right: Block-A: An image and four candidate poses, where only a part of each pose is well-aligned with the person. Block-B: We divide each candidate pose into limb parts. Block-C: We allow the recombination of limbs from different pose candidates with constraints between two limbs that have a joint in common. Block-D: An example where recombination builds an accurate pose, which is not in the original candidate set.

6.4.8. *The LEAR Submission at Thumos 2014*

Participants: Dan Oneata, Jakob Verbeek, Cordelia Schmid.

In [28] we describe the submission of our team to the THUMOS workshop in conjunction with ECCV 2014. Our system is based on Fisher vector (FV) encoding of dense trajectory features (DTF), which we also used in our 2013 submission. The dataset is based on the UCF101 dataset, which is currently the largest action dataset both in terms of number of categories and clips, with more than 13000 clips drawn from 101 action classes. This year special attention was paid to classification of uncropped videos, where the action of interest appears in videos that contain also non-relevant sections. This year's submission additionally incorporated static-image features (SIFT, Color, and CNN) and audio features (ASR and MFCC) for the classification task. For the detection task, we combined scores from the classification task with FV-DTF features extracted from video slices. We found that these additional visual and audio feature significantly improve the classification results. For localization we found that using the classification scores as a contextual feature besides local motion features leads to significant improvements. In Figure 17   we show the middle frame from the top four ranked videos corresponding to the three hardest classes (as evaluated on the validation data). Our team has ranked second on the classification challenge (out of eleven teams) and first on the detection challenge (out of three teams).
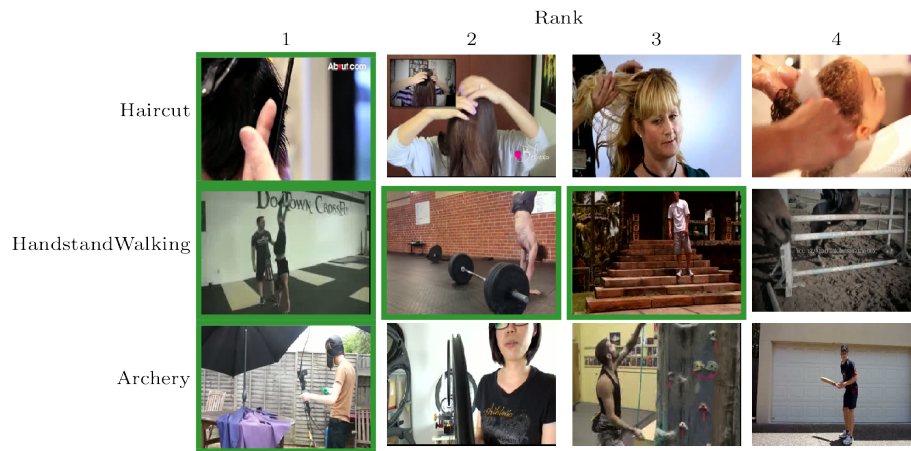
Figure 17. Snapshots from the top four ranked test videos for the three hardest classes; green borders indicate true positives.

6.4.9. The LEAR Submission at TrecVid MED 2014

Participants: Matthijs Douze, Dan Oneata, Mattis Paulin, Clément Leray, Nicolas Chesneau, Danila Potapov, Jakob Verbeek, Karteek Alahari, Zaid Harchaoui, Lori Lamel [Spoken Language Processing group, LIMSI, CNRS], Jean-Luc Gauvain [Spoken Language Processing group, LIMSI, CNRS], Christoph Schmidt [Fraunhofer IAIS, Sankt Augustin], Cordelia Schmid.

In [26] we describe our participation to the 2014 edition of the TrecVid Multimedia Event Detection task. Our system is based on a collection of local visual and audio descriptors, which are aggregated to global descriptors, one for each type of low-level descriptor, using Fisher vectors. Besides these features, we use two features based on convolutional networks: one for the visual channel, and one for the audio channel. Additional high-level features are extracted using ASR and OCR features. Finally, we used mid-level attribute features based on object and action detectors trained on external datasets. In the notebook paper we present an overview of the features and the classification techniques, and experimentally evaluate our system on TrecVid MED 2011 data.

We participated in four tasks, which differ in the amount of training videos for each event (either 10 or 100), and the time that is allowed for the processing. For the 20 pre-specified events several weeks are allowed to extract features, train models, and to score the test videos (which consisted of 8,000 hours of video this year). For the 10 ad-hoc events, we only have five days to do all processing. Across the 11 participating teams, our results ranked first for the 10-example ad-hoc task, and fourth and fifth place for the other tasks.

<p style="text-align:center; color:red; font-weight:bold; font-size:large">MAVERICK Project-Team</p>

5. New Results

5.1. Highlights of the Year

The impacting PhD work [3] of Eric Heitz on *appearance filtering* (see section 5.5.1) has received a very good reception in both academic and industrial world, including several "best paper" prizes in 2013 and 2014, invitation to participate to the Siggraph Course on Photorealistic Rendering [13], and statements of importance and/or integration by reference peoples and CG companies.

BEST PAPER AWARD :

[] **Computer Graphics Forum**. E. HEITZ, E. D'EON.

5.2. Visual perception

5.2.1. *The effects of surface gloss and roughness on color constancy for real 3-D objects*

Participants: Jeoren J. M. Granzier, Romain Vergne [contact], Karl Gegenfurtner.

Color constancy denotes the phenomenon that the appearance of an object remains fairly stable under changes in illumination and background color. Most of what we know about color constancy comes from experiments using flat, matte surfaces placed on a single plane under diffuse illumination simulated on a computer monitor. Here we investigate whether material properties (glossiness and roughness) have an effect on color constancy for real objects. Subjects matched the color and brightness of cylinders (painted red, green, or blue) illuminated by simulated daylight (D65) or by a reddish light with a Munsell color book illuminated by a tungsten lamp. The cylinders were either glossy or matte and either smooth or rough. The object was placed in front of a black background or a colored checkerboard as shown in Figure 6 . We found that color constancy was significantly higher for the glossy objects compared to the matte objects, and higher for the smooth objects compared to the rough objects. This was independent of the background. We conclude that material properties like glossiness and roughness can have significant effects on color constancy [7].

5.3. Visualization

Participants: Léo Allemand-Giorgis, Georges-Pierre Bonneau [contact].

In computer visualization we have worked on two topics: topology for visualization and perception for visualization.

In topology for visualization we have worked on scalar field vizualization methods taking into account the topology of the data. In [14] We have derived theoretical results on monotonic interpolation of scalar data. Our method enables to interpolate given topological data such as minima, maxima and saddle points at the corners of a rectangular domain without adding spurious extrema inside the function domain, as illustrated in Figure 7 .

We have collaborated to a state of the art chapter on Uncertain Visualization [15], in which we described the evaluation of visualization methods based on visual perception.

Furthermore we have worked on two topics related to geometry for visualization. In [6] we introduce a method for interpolating a quad mesh using G1-continuous polynomial surfaces. We plan to use this method in the future for displaying isosurfaces of higher order data. In [11] we have published a method for reconstructing interfaces in highly complex assemblies, as illustrated in Figure 8 . This method has been developed in order to visualize data arising from simulation of complex mechanical assemblies, within the ANR project ROMMA, closed in January 2014.
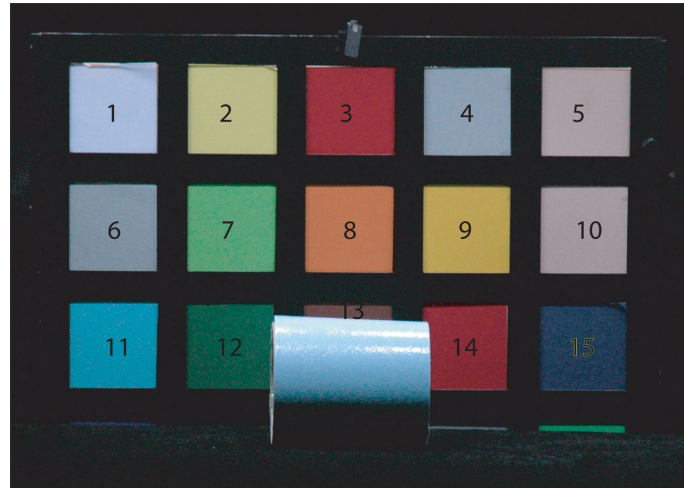
Figure 6. Color perception depends on material properties. This image represents one stimulus used in our experiment to compare the effect of glossiness on color constancy.
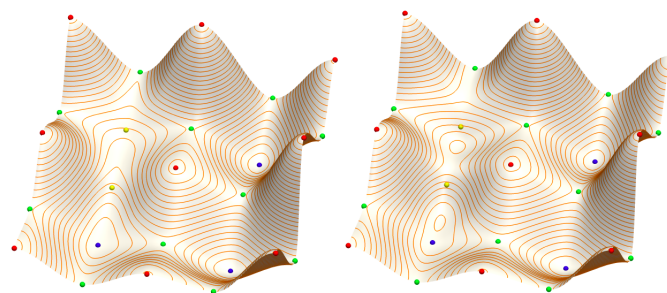

Figure 7. Local maxima (red), minima (blue), saddles (green) and regular (yellow) vertices are interpolated by a C1 piecewise cubic interpolant. Left: no unwanted local extrema exist in the interior of the cubic patches. Right: partial derivatives too large in size are chosen for the yelllow regular vertices implying that additional unwanted local extrema appear inside the cubic polynomial patches.
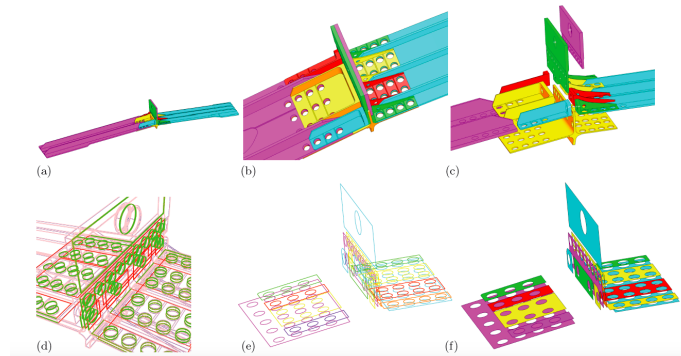
Figure 8. Aircraft part for assembling the wings with the body of an aircraft (model courtesy of EADS). (a,b) two views of the components, (c) exploded view, (d) ray casting results, (e) boundary reconstruction, (f) nal interfaces

5.4. Image creation and editing

5.4.1. Programmable 2D Arrangements for Element Texture Design

Participants: Hugo Loi, Thomas Hurtut, Romain Vergne, Joëlle Thollot [contact].

We introduce a programmable method for designing stationary 2D arrangements for element textures, namely textures made of small geometric elements. These textures are ubiquitous in numerous applications of computer-aided illustration. Previous methods, whether they be example-based or layout-based, lack control and can produce a limited range of possible arrangements. Our approach targets technical artists who will design an arrangement by writing a script. These scripts are using three types of operators: *partitioning operators* for defining the broad-scale organization of the arrangement, *mapping operators* for controlling the local organization of elements, and *merging operators* for mixing different arrangements. These operators are designed so as to guarantee a stationary result meaning that the produced arrangements will always be repetitive. We show (see Figure 10) that this simple set of operators is sufficient to reach a much broader variety of arrangements than previous methods. Editing the script leads to predictable changes in the synthesized arrangement, which allows an easy iterative design of complex structures. Finally, our operator set is extensible and can be adapted to application-dependent needs.

5.4.2. Color transfer guided by summary statistics

Participants: Benoît Arbelot, Romain Vergne [contact], Thomas Hurtut, Joëlle Thollot.

Modifying the colors of an image is an attractive way to edit its ambiance and mood. In practice, manually and directly tuning the color distribution of an image is challenging and tedious. Color transfer methods offer an intuitive alternative by automatically changing an image colors according to a target image. Existing transfer methods mostly rely on global matching processes to reshape and map the color histogram of the source image as close as possible to the target histogram. However, they offer no control over where the colors of the target will be transferred in the source image: they only tend to match colors that have similar intensities and chromaticities. This can lead to unexpected results, especially when some elements do not have the same colors in the two images, but share similar features. In this work, we propose to implicitly segment input images before transferring colors. Instead of relying on colors only, we use a summary of statistics to describe the underlying texture properties of each pixel. This provides a measure of pixel similarity which is then used to guide and ensure the transfer to be done between similar features (see Figure reffig:color for a preliminary result).

Source   Target   Tranfer

Figure 9. By transferring colors between statistically similar pixels, meaningful colors are transferred from the target to the source image.

5.5. Complex scenes

In order to render both efficiently and accurately ultra-detailed large scenes, this approach consists in developing representations and algorithms able to account compactly for the quantitative visual appearance of a regions of space projecting on screen at the size of a pixel.

5.5.1. *Surfacic appearance pre-filtering*

Participants: Eric Heitz, Fabrice Neyret [contact].

Here, we deal with complex surfaces represented by microfacets and material attributes.

Among the various correlations between material ingredients forming the BRDF, we published an extended version of the work on correlation between surface attribute (like color) and visibility [9], and Eric published an comprehensive interpretation of the microfacet model in a journal in the field of physics [8]. He also adapted his microfacet approach to the efficient BRDF sampling for path tracing – published in EGSR/CGF [] –, see Figure 11 , and he was invited to participate to the prestigious Siggraph Course "Physically Based Shading in Theory and Practice" [13]. This work is now implemented in various professional and standard software and thus settled a new standard. Eric defended his PhD on September, 26 2014 [3]

5.5.2. *Volumetric appearance pre-filtering*

Participants: Guillaume Loubet, Fabrice Neyret [contact].

Here, we deal with complex density distributions. The first target is galactic material in the scope of the veRTIGE / Galaxy ANR project, but the long term goal is more general since at long distance complex surfaces or scattered objects can more efficiently be represented as volumetric distributions.

The usual hypothesis in CG is that volumes are homogeneous distribution of matter. But star and (dark) dust distributions are fractal, not homogeneous. This breaks all the existing equations accounting for large scale opacity and lighting of volumes of such material.

first, we developed a new procedural noise able to easily mimic such fractal distributions according to astrophysical models (see fig 12 ,a). Then we studied how to reproduce the same opacity (fig 12 ,b) and reflectivity (fig 12 ,c) for various level of details – this is still ongoing work.

Moreover, volumetric material is often concentrated into bodies, with a boundary delimited by a density jump or gradient. We studied the macroscopic light behavior in such configurations (fig 12 ,d).

5.6. Realistic rendering

Note that Cyril Soler defended his HDR "*Models and Analyses for Image Synthesis*", Université Joseph-Fourier, on June 2014.
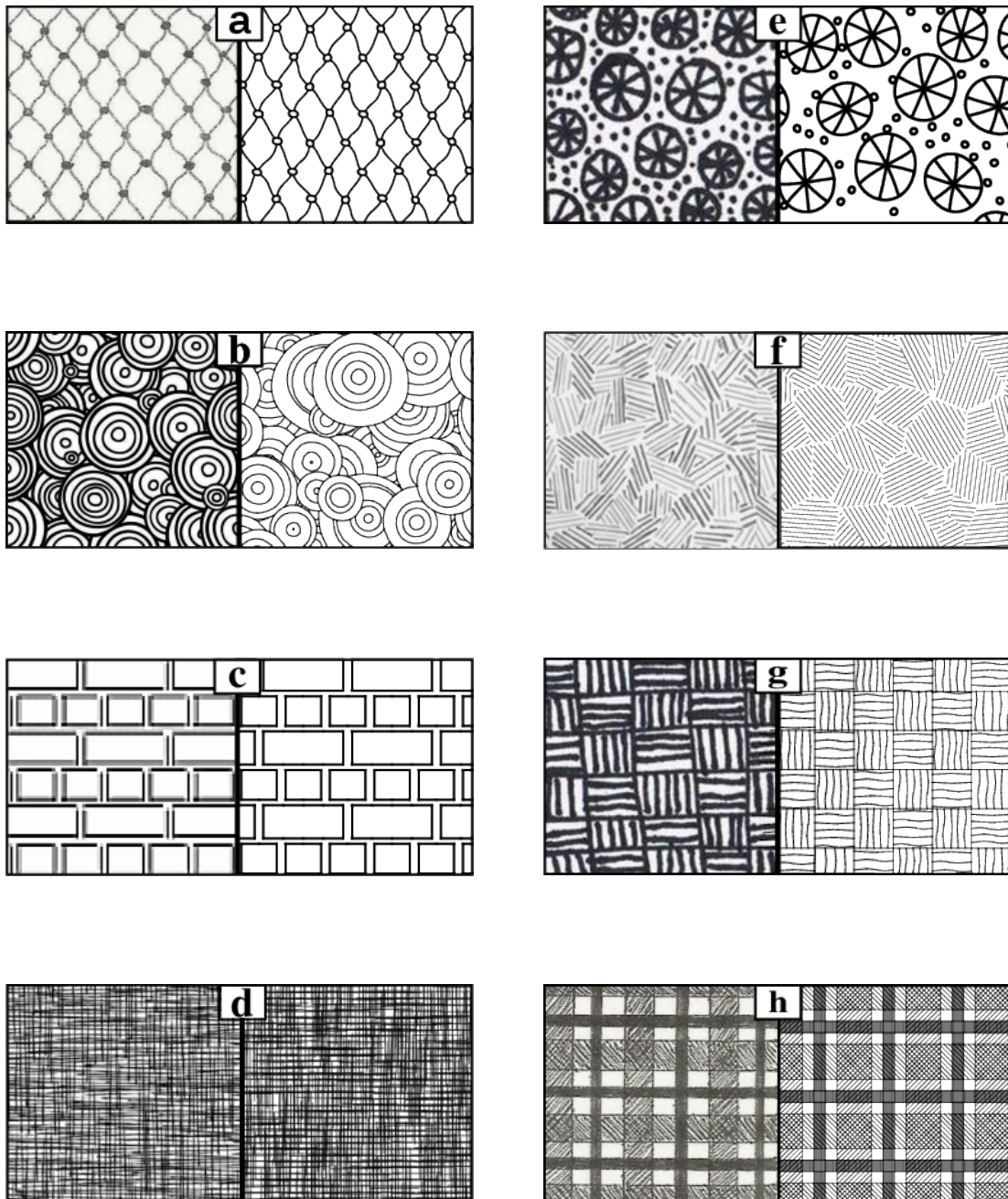
Figure 10. **Element textures commonly used.** *These textures can be found in professional art (d,g,h), casual art (a,e,f), technical productions such as Computer-Assisted Design illustration tools (c), and textile industry (b). For each example, we show a hand-drawn image (left), and our synthesized reproduction of its geometric arrangement (right).* **(a,b,c)** *Classic regular distributions with contact, overlap and no adjacency between elements respectively.* **(d)** *Overlap of two textures creating cross hatching.* **(e)** *Non overlapping combination of two textures.* **(f,g,h) Complex element textures with clusters of elements. — Image credit: (d,g,h) "Rendering in Pen and Ink: The Classic Book On Pen and Ink Techniques for Artists, Illustrators, Architects, and Designers"* [20]; *(a,e) Profusion Art [*profusionart.blogspot.com*]; (f) Hayes' Art Classes [*hayesartclasses.blogspot.com*]; (c) CompugraphX [*www.compugraphx.com*]; (b) 123Stitch [*www.123stitch.com*].*
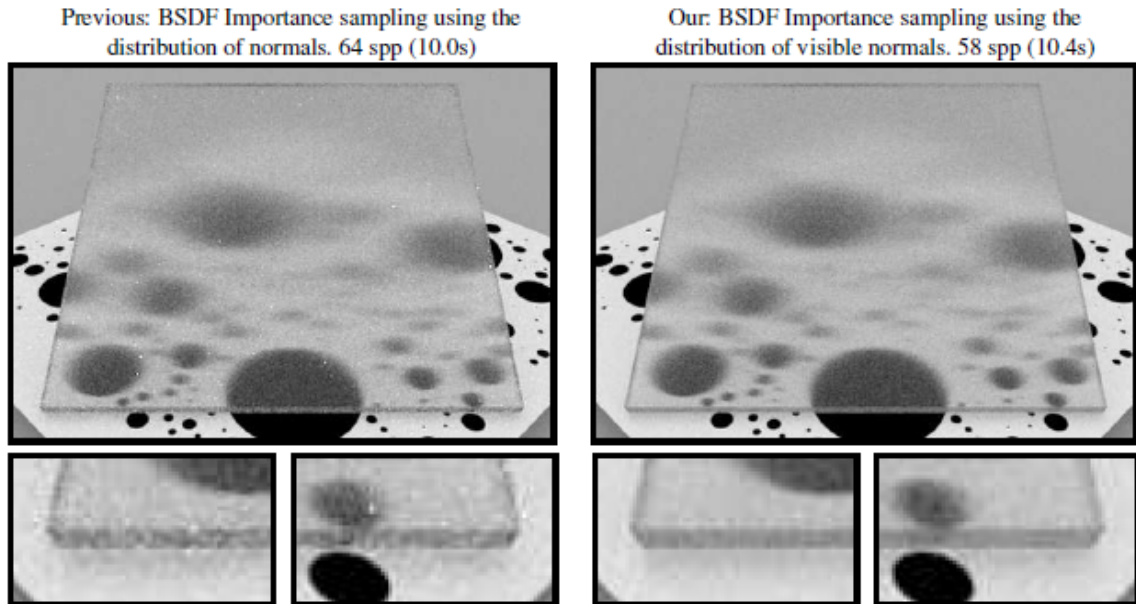
Figure 11. A dielectric glass plate (n = 1.5) with anisotropic GGX roughness (ax = 0.05, ay = 0.4) on all faces (with the Smith masking function). For a similar sample budget and the same render time, our method (right) significantly reduces the variance and converges faster than the common technique used in previous work (left).
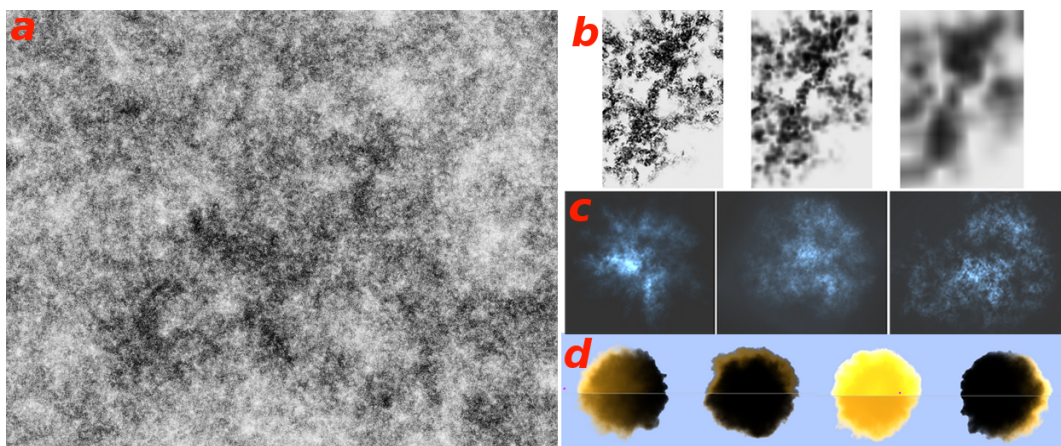
Figure 12. a: Our new fractal procedural noise. b: Multiscale opacity. c: Multiscale reflectance. d: Light reflection at volumetric bodies boundary with gradient (top) or jump (bottom) density distribution, with different light direction (left to right).

5.6.1. *Single Scattering in participating media with refractive boundaries*
Participant: Nicolas Holzschuch [contact].

(a) Our algorithm (169 s) **(b)** Point samples [WZHB09] **(c)** Photon mapping
(27 samples, 170 s) (650 000 photons, 167 s)
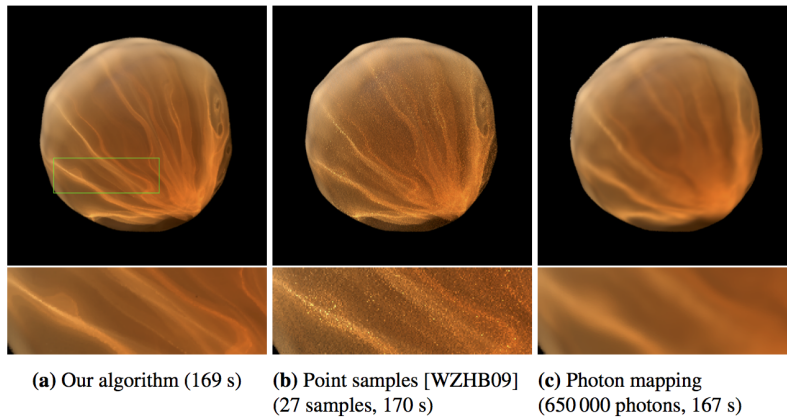
Figure 13. Single scattering: comparison between our algorithm and existing methods (equal computation time) on a translucent sphere illuminated by a point light source from behind.

Volume caustics are high-frequency effects appearing in participating media with low opacity, when refractive interfaces are focusing the light rays (see Figure 13). Refractions make them hard to compute, since screen locality does not correlate with spatial locality in the medium. We have developed a new method for accurate computation of single scattering effects in a participating media enclosed by refractive interfaces. Our algorithm is based on the observation that although radiance along each camera ray is irregular, contributions from individual triangles are smooth. Our method gives more accurate results than existing methods, faster. It uses minimal information and requires no precomputation or additional data structures. This paper was accepted for publication at Computer Graphics Forum [10].

5.6.2. *A Local Frequency Analysis of Light Scattering and Absorption*
Participants: Laurent Belcour, Kavita Bala, Cyril Soler [contact].

We proposed a novel analysis of absorption and scattering of local light fields in the Fourier domain in the neighborhood of light paths. This analysis aims at predicting the changes over the distribution of light energy, so as to allow efficient sampling and integration methods of diffused light in participating media. Our analysis explains that absorbtion increases frequency since it acts as a continuous visibility mask over the local light field, and that scattering lowers frequencies as it operates a low pass convolution filter in the directional domain. In order to combine this analysis with our previous work on covariance tracing—and therefore use it to improve existing algorithms for path tracing in participating media—we derived new sampling metrics all based on a common prediction of the 3D covariance of the fluence in the volume. We demonstrate indeed that the covariance of the fluence can efficiently be computed by combining the 4D covariance matrices of light fields in the neighborhood of light paths, and that it can be used to compute effective metrics (1) for the variance of energy collected along camera rays, (2) for determining the shape and size of reconstruction kernels in screen space, and (3) for drastically improving the convergence of density estimation methods. For the later, we propose an improvement of the method of Progressive Photon Beams. This work has been published in ACM Transactions on Graphics and presented at Siggraph'2014 in Vancouver [5].
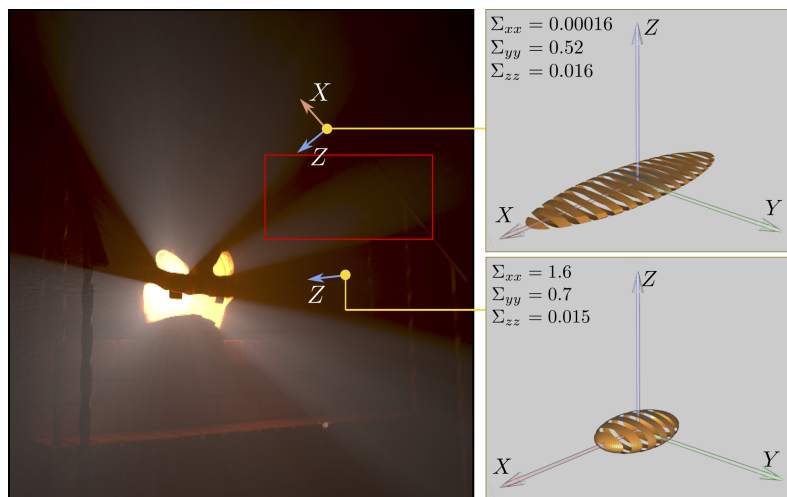
Figure 14. Predictions of the covariance of the Fourier spectrum of the fluence in the volume computed using our Fourier analysis of scattering and absorption.

<p style="text-align:center;">**MORPHEO Project-Team**</p>

6. New Results

6.1. Human Shape and Pose Tracking Using Keyframes

In this work we consider human tracking in multi-view set-ups and investigates a robust strategy that learns online key poses to drive a shape tracking method. The interest arises in realistic dynamic scenes where occlusions or segmentation errors occur. The corrupted observations present missing data and outliers that deteriorate tracking results. We propose to use key poses of the tracked person as multiple reference models. In contrast to many existing approaches that rely on a single reference model, multiple templates represent a larger variability of human poses. They provide therefore better initial hypotheses when tracking with noisy data. Our approach identifies these reference models online as distinctive keyframes during tracking. The most suitable one is then chosen as the reference at each frame. In addition, taking advantage of the proximity between successive frames, an efficient outlier handling technique is proposed to prevent from associating the model to irrelevant outliers. The two strategies are successfully experimented with a surface deformation framework that recovers both the pose and the shape. Evaluations on existing datasets also demonstrate their benefits with respect to the state of the art. This work was presented at CVPR'14 [5].
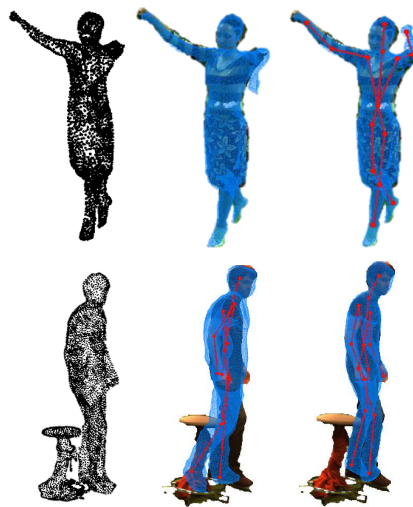

Figure 4. Shape tracking with keframes[5]

6.2. On Mean Pose and Variability of 3D Deformable Models

We present a novel methodology for the analysis of complex object shapes in motion observed by multiple video cameras. In particular, we propose to learn local surface rigidity probabilities (i.e., deformations), and to estimate a mean pose over a temporal sequence. Local deformations can be used for rigidity-based dynamic surface segmentation, while a mean pose can be used as a sequence keyframe or a cluster prototype and has therefore numerous applications, such as motion synthesis or sequential alignment for compression or morphing. We take advantage of recent advances in surface tracking techniques to formulate a generative model of 3D temporal sequences using a probabilistic framework, which conditions shape fitting over all

frames to a simple set of intrinsic surface rigidity properties. Surface tracking and rigidity variable estimation can then be formulated as an Expectation-Maximization inference problem and solved by alternatively minimizing two nested fixed point iterations. We show that this framework provides a new fundamental building block for various applications of shape analysis, and achieves comparable tracking performance to state of the art surface tracking techniques on real datasets, even compared to approaches using strong kinematic priors such as rigid skeletons.
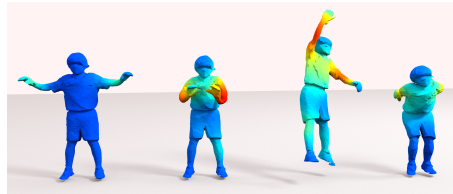
Figure 5. Rigidity probability of the shape tracked with mean pose [4]

6.3. Segmentation multi-vues par coupure de graphes

In this paper, we address the problem of object segmentation in multiple views when two or more viewpoints of the same scene are available. We propose a new approach that propagates segmentation coherence information in space, hence allowing evidence in one image to be shared over the complete set. To this aim the segmentation is cast as a single efficient labeling problem over space and time with graph cuts. In contrast to most existing multi-view segmentation methods that rely on some form of dense reconstruction, ours only requires a sparse 3D sampling to propagate information between viewpoints. The approach is thoroughly evaluated on standard multi-view datasets. The obtained results compete with state of the art methods but they are achieved with significantly fewer viewpoints.
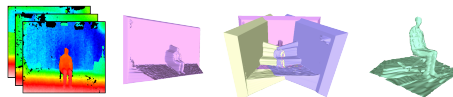

Figure 6. Results of our multi-view segmentation approach over 3 input views, with no user interaction (completely automated). [9]

6.4. Combined Visible and X-Ray 3D Imaging

This work considers 3D imaging of moving objects and introduces a technique that exploits visible and x-ray images to recover dense 3D models. While recent methods such as tomography from cone-beam x-ray can advantageously replace more expensive and higher-dose CT scanners, they still require specific equipment and immobilised patients. We investigate an alternative strategy that combines a single x-ray source and a set of colour cameras to capture rigidly moving samples. The colour cameras allow for coarse marklerless motion tracking, which is further refined with the x-ray information. Once the sample poses are correctly estimated, a dense 3D attenuation model is reconstructed from the set of x-ray frames. Preliminary results on simulated data compared to ground-truth as well as actual in-vivo experiments were presented at the conference MIUA'14 [6].
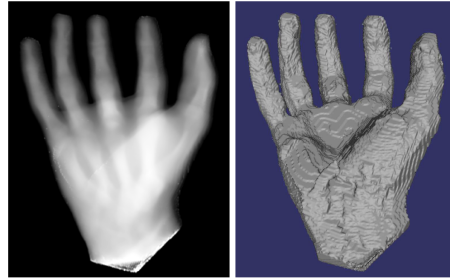
Figure 7. Human hand reconstruction with a combined visible and x-ray system [6]

6.5. Non-Rigid Registration meets Surface Reconstruction

Non rigid registration is an important task in computer vision with many applications in shape and motion modeling. A fundamental step of the registration is the data association between the source and the target sets. Such association proves difficult in practice, due to the discrete nature of the information and its corruption by various types of noise, e.g. outliers and missing data. In this work we investigate the benefit of the implicit representations for the non-rigid registration of 3D point clouds. First, the target points are described with small quadratic patches that are blended through partition of unity weighting. Then, the discrete association between the source and the target can be replaced by a continuous distance field induced by the interface. By combining this distance field with a proper deformation term, the registration energy can be expressed in a linear least square form that is easy and fast to solve. This significantly eases the registration by avoiding direct association between points. Moreover, a hierarchical approach can be easily implemented by employing coarse-to-fine representations. Experimental results were conducted with point clouds from multi-view data sets. The qualitative and quantitative comparisons show the outperformance and robustness of our framework. This work was presented at 3DV'14[7].
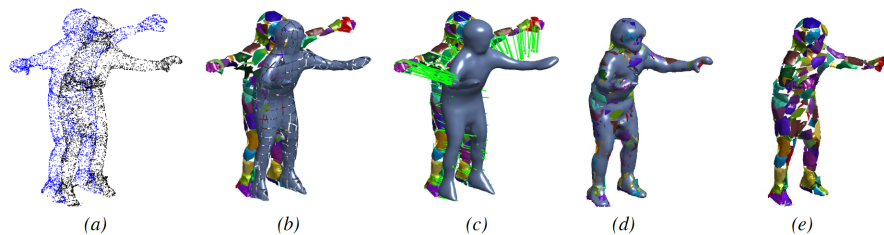

Figure 8. Using implicit interface for registration: (a) initial pose of the source and target sets; (b) source patches and the local quadrics representing the target; (c) the implicit interface induces a gradient field; (d) deformed source patches fitting the interface; a coarse-to-fine interface has been used in (c) - (d); (e) the final deformation of the template [7].

6.6. High Resolution 3D Shape Texture from Multiple Videos

We examine the problem of retrieving high resolution textures of objects observed in multiple videos under small object deformations. In the monocular case, the data redundancy necessary to reconstruct a high-

resolution image stems from temporal accumulation. This has been vastly explored and is known as super-resolution. On the other hand, a handful of methods have considered the texture of a static 3D object observed from several cameras, where the data redundancy is obtained through the different viewpoints. We introduce a unified framework to leverage both possibilities for the estimation of a high resolution texture of an object. This framework uniformly deals with any related geometric variability introduced by the acquisition chain or by the evolution over time. To this goal we use 2D warps for all viewpoints and all temporal frames and a linear projection model from texture to image space. Despite its simplicity, the method is able to successfully handle different views over space and time. As shown experimentally, it demonstrates the interest of temporal information that improves the texture quality. Additionally, we also show that our method outperforms state of the art multi-view super-resolution methods that exist for the static case. This work was presented at CPVR'14 [8].
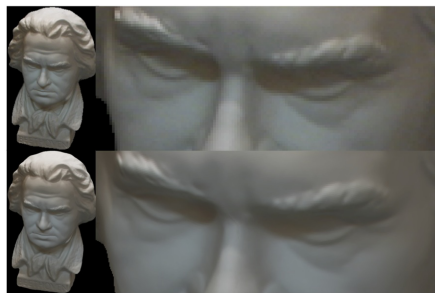

Figure 9. Input view 768 × 576 resolution with up-sampling by factor of three, BEETHOVEN dataset. Super-resolved 2304×1728 output of our algorithm rendered from identical viewpoint [8].

<p style="text-align:center; color:red;">**PERCEPTION Project-Team**</p>

5. New Results

5.1. Highlights of the Year

- In 2014 Antoine Deleforge (team member 2009-2013) received the **Signal, Image and Vision best PhD prize** for his thesis *"Acoustic Space Mapping: A Machine Learning Approach to Sound Source Separation and Localization"*, defended in December 2013 and advised by Radu Horaud. The prize is jointly awarded by GDR ISIS, Club EEA, and GRETSI.

  Website: http://www.inria.fr/centre/grenoble/actualites/apprendre-a-rester-attentif-a-ses-locuteurs

- Radu Horaud was awarded an **ERC Advanced Grant** for his five year project VHIA *"Vision and Hearing in Action"*, grant number 340113, 2014-2019.

  Website: https://team.inria.fr/perception/vhia/.

- The PERCEPTION team was awarded an **ANR BLANC** two year project MIXCAM *"Real-Time Visual Reconstruction by Mixing Multiple Depth and Color Cameras"*, in collaboration with 4D View Solutions, 2014-2016.

  Website: https://team.inria.fr/perception/mixcam-project/

- The PERCEPTION team was awarded an **FP7 STREP** three year project EARS *"Embodied Audition for Robots"*, in collaboration with Friedrich Alexander Universiteit, coordinator (Germany), Ben Gurion University (Israel), Imperial College (UK), Humboldt University Berlin (Germany) and Aldebaran Robotics (France), 2014-2017.

  Website: https://team.inria.fr/perception/ears/

5.2. Acoustic Space Learning on Binaural Manifolds

We addressed the problems of modeling the acoustic space generated by a full-spectrum sound source and of using the learned model for the localization and separation of multiple sources that simultaneously emit sparse-spectrum sounds. We lay theoretical and methodological grounds in order to introduce the binaural manifold paradigm. We perform an in-depth study of the latent low-dimensional structure of the high-dimensional interaural spectral data, based on a corpus recorded with a human-like audiomotor robot head, namely the POPEYE robot shown on Fig 3 (right). A non-linear dimensionality reduction technique is used to show that these data lie on a two-dimensional (2D) smooth manifold parameterized by the motor states of the listener, or equivalently, the sound source directions, e.g., Fig. 4 . We propose a probabilistic piecewise affine mapping model (PPAM) specifically designed to deal with high-dimensional data exhibiting an intrinsic piecewise linear structure. We derive a closed-form expectation-maximization (EM) procedure for estimating the model parameters, followed by Bayes inversion for obtaining the full posterior density function of a sound source direction. We extend this solution to deal with missing data and redundancy in real world spectrograms, and hence for 2D localization of natural sound sources such as speech. We further generalize the model to the challenging case of multiple sound sources and we propose a variational EM framework. The associated algorithm, referred to as variational EM for source separation and localization (VESSL) yields a Bayesian estimation of the 2D locations and time-frequency masks of all the sources. Comparisons of the proposed approach with several existing methods reveal that the combination of acoustic-space learning with Bayesian inference enables our method to outperform state-of-the-art methods [18], [24].

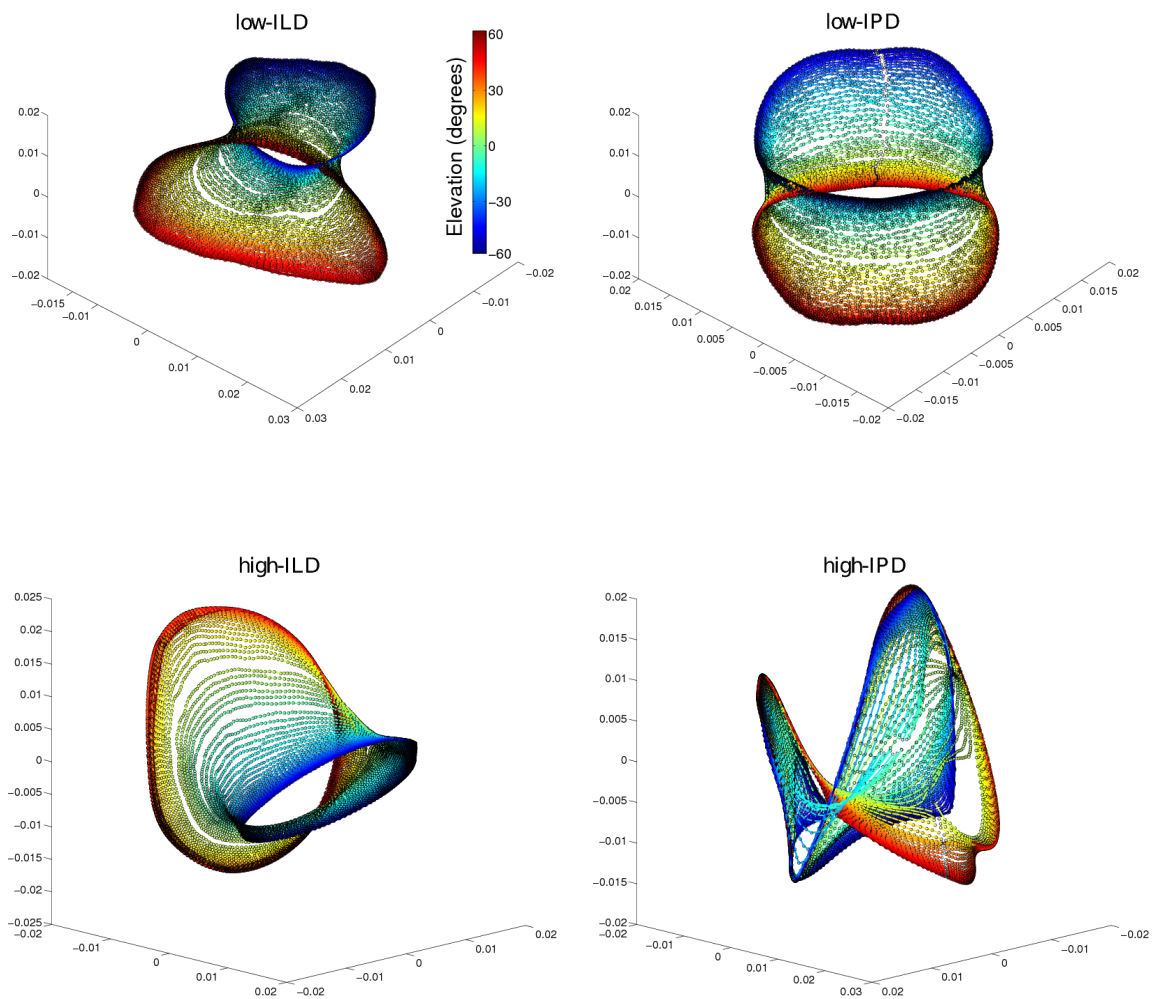Website: https://team.inria.fr/perception/research/acoustic-learning/

Figure 4. This figure illustrates the concept of binaural manifold. A wide-spectrum sound is recorded with a binaural acoustic dummy head and an interaural high-dimensional spectral representation of this sound is mapped onto a low-dimensional (2) space. This reveals the two-dimensional manifold of possible sound-source directions that is embedded in the interaural spectral features. Please consult [18] for more details.

5.3. Geometric Sound Source Localization

We address the problem of sound-source localization from time-delay estimates using arbitrarily-shaped non-coplanar microphone arrays. A novel geometric formulation is proposed, together with a thorough algebraic analysis and a global optimization solver [15]. The proposed model is thoroughly described and evaluated. The geometric analysis, stemming from the direct acoustic propagation model, leads to necessary and sufficient conditions for a set of time delays to correspond to a unique position in the source space. Such sets of time delays are referred to as *feasible sets*. We formally prove that every feasible set corresponds to exactly one position in the source space, whose value can be recovered using a closed-form localization mapping. Therefore we seek for the optimal feasible set of time delays given, as input, the received microphone signals. This time delay estimation problem is naturally cast into a programming task, constrained by the feasibility conditions derived from the geometric analysis. A global branch-and-bound optimization technique is proposed to solve the problem at hand, hence estimating the best set of feasible time delays and, subsequently, localizing the sound source. Extensive experiments with both simulated and real data are reported; we compare our methodology to four state-of-the-art techniques. This comparison shows that the proposed method combined with the branch-and-bound algorithm outperforms existing methods. These in-depth geometric understanding, practical algorithms, and encouraging results, open several opportunities for future work.

Website: https://team.inria.fr/perception/research/geometric-sound-source-localization/

5.4. Joint Registration of Multiple Point Sets

We developed a probabilistic generative model and its associated algorithm to jointly register multiple point sets. The vast majority of state-of-the-art registration techniques select one of the sets as the *model* and perform pairwise alignments between the other sets and this set. The main drawback of this mode of operation is that there is no guarantee that the model-set is free of noise and outliers, which contaminates the estimation of the registration parameters. Unlike previous work, the proposed method treats all the point sets on an equal footing: they are realizations of a Gaussian mixture (GMM) and the registration is cast into a clustering problem [26]. We formally derive an EM algorithm that estimates both the GMM parameters and the rotations and translations that map each individual set onto the *central* model. The mixture means play the role of the registered set of points while the variances provide rich information about the quality of the registration. We thoroughly validate the proposed method with challenging datasets, we compare it with several state-of-the-art methods, and we show its potential for fusing real depth data.

Website: https://team.inria.fr/perception/research/jrmpc/

5.5. High-Dimensional Regression

The problem of approximating high-dimensional data with a low-dimensional representation is addressed. The article makes the following contributions. An inverse regression framework is proposed, which exchanges the roles of input and response, such that the low-dimensional variable becomes the regressor, and which is tractable. A mixture of locally-linear probabilistic mapping model is introduced, that starts with estimating the parameters of the inverse regression, and follows with inferring closed-form solutions for the forward parameters of the high-dimensional regression problem of interest. Moreover, a partially-latent paradigm is introduced, such that the vector-valued response variable is composed of both observed and latent entries, thus being able to deal with data contaminated by experimental artifacts that cannot be explained with noise models. The proposed probabilistic formulation could be viewed as a latent-variable augmentation of regression. Expectation-maximization (EM) procedures are introduced, based on a data augmentation strategy which facilitates the maximum-likelihood search over the model parameters. Two augmentation schemes are proposed and the associated EM inference procedures are described in detail; they may well be viewed as generalizations of a number of EM regression, dimension reduction, and factor analysis algorithms. The proposed framework is validated with both synthetic and real data. Experimental evidence is provided that the method outperforms several existing regression techniques [19], [25].

Website: https://team.inria.fr/perception/research/high-dim-regression/

5.6. Audiovisual Speaker Detection, Localization and Interaction with NAO

In this research we address the problem of audio-visual speaker detection. We introduce an online system working on the humanoid robot NAO. The scene is perceived with two cameras and two microphones. A *multimodal* Gaussian mixture model (GMM) fuses the information extracted from the auditory and visual sensors. The system is implemented based on a platform-independent middleware library and it is able to process the information online (17 visual frames per second). A detailed method description and the system implementation are provided, with special emphasis on the online processing issues that must be addressed, and the proposed solutions. Experimental validation is done over five different scenarios, with no special lighting, nor special acoustic conditions, leading to good results [16].

Website: https://team.inria.fr/perception/research/audiovisual-nao/
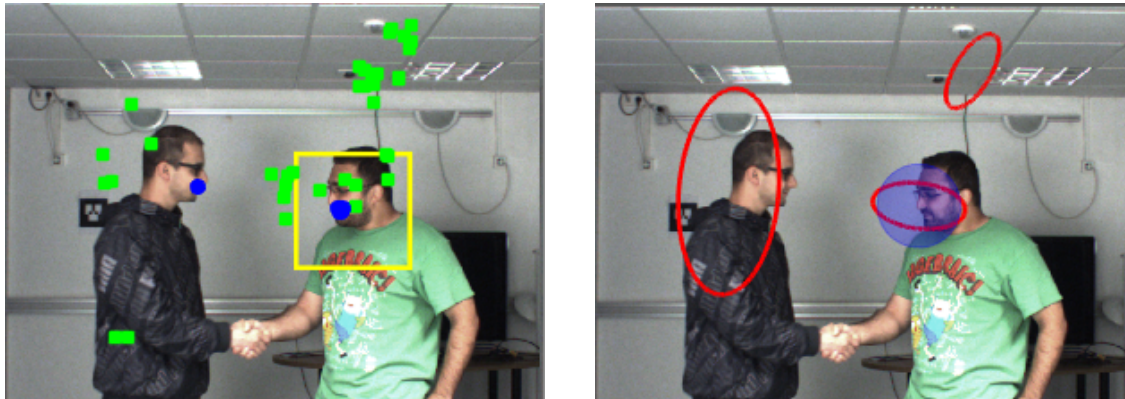
5.7. EM for Weighted-Data Clustering

Figure 5. We developed a novel multimodal clustering method that is based on expectation-maximization (EM) with weighted data. The left image shows auditory features (green), namely sound source positions mapped onto the image plane using [24] and visual features (blue, lip landmarks), as well as the active speaker (yellow square). The right image shows the results of our weighted-data EM algorithm that finds three clusters. Among these clusters, the active audio-visual cluster is marked with a transparent blue circle.

Data clustering has received a lot of attention and many methods, algorithms and software packages are currently available. Among these techniques, parametric finite-mixture models play a central role due to their interesting mathematical properties and to the existence of maximum-likelihood estimators based on expectation-maximization (EM). In this paper we propose a new mixture model that associates a weight with each observed data point. We introduce a Gaussian mixture with weighted data and we derive two EM algorithms [29]: the first one considers the weight of each observed datum to be fixed, while the second one treats each weight as a hidden variable drawn from a gamma distribution. We provide a general-purpose scheme for weight initialization and we thoroughly validate the proposed algorithms by comparing them with several parametric and non-parametric clustering techniques. We demonstrate the utility of our method for clustering heterogeneous data, namely data gathered with different sensorial modalities, e.g., audio and vision.

Website: https://team.inria.fr/perception/research/wdgmm/

5.8. Continuous Action Recognition

Continuous action recognition is more challenging than isolated recognition because classification and segmentation must be simultaneously carried out. We build on the well known dynamic time warping (DTW) framework and devise a novel visual alignment technique, namely dynamic frame warping (DFW), which performs isolated recognition based on per-frame representation of videos, and on aligning a test sequence with a model sequence. Moreover, we propose two extensions which enable to perform recognition concomitant with segmentation, namely one-pass DFW and two-pass DFW. These two methods have their roots in the domain of continuous recognition of speech and, to the best of our knowledge, their extension to continuous visual action recognition has been overlooked. We test and illustrate the proposed techniques with a recently released dataset (RAVEL) [32] and with two public-domain datasets widely used in action recognition (Hollywood-1 and Hollywood-2). We also compare the performances of the proposed isolated and continuous recognition algorithms with several recently published methods [22].

Website: https://team.inria.fr/perception/research/car/

5.9. Skeletal Quads

Recent advances on human motion analysis have made the extraction of human skeleton structure feasible, even from single depth images. This structure has been proven quite informative for discriminating actions in a recognition scenario. In this context, we propose a local skeleton descriptor that encodes the relative position of joint quadruples. Such a coding implies a similarity normalization transform that leads to a compact (6D or 5D) view-invariant skeletal feature, referred to as skeletal quad. In the references below, we use this descriptor in conjunction with Fisher kernel in order to encode gesture or action (sub)sequences. The short length of the descriptor compensates for the large inherent dimensionality associated to Fisher vectors. We investigate the performance in both isolated [28] and continuous [27] recognition scenarios.

Website: https://team.inria.fr/perception/research/skeletalquads/

PRIMA Project-Team

5. New Results

5.1. Highlights of the Year

On March 14, 2014, James Crowley was named Chevalier de l'Ordre national du Mérite.

On August 2014, the paper "Human-Robot Motion: an Attention-Based Navigation Approach" [14] by Thierry Fraichard, Remi Paulin & Patrick Reignier has been nominated for the best paper award at the IEEE Int. Symp. on Robot and Human Interactive Communication (RO-MAN 2014), Edinburgh (UK).

On December 2014, Patrick Reignier was a member of the EDF grand jury for smart energy

BEST PAPER AWARD :

[14] **IEEE Int. Symp. on Robot and Human Interactive Communication (ROMAN)**. T. FRAICHARD, R. PAULIN, P. REIGNIER.

5.2. Attention-Based Navigation

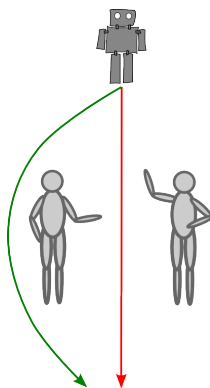**Participants:** Thierry Fraichard, Remi Paulin, Patrick Reignier..

Figure 6. People are not pieces of furniture! Motion in red is definitely shorter but it is not appropriate.

The domain of service-robots is growing fast and has become the focus of many researchers and industrials alike. Their application areas been extremely broad, from logistics to handicap assistance. A large proportion of such robots are expected to share humans' living space and thus must be endowed with navigation capabilities that exceed the standard requirements pertaining to autonomous navigation such as motion safety. In a human populated environment, optimality does not boil down to minimising resources such as time or distance travelled anymore, the robot motion must abide by social/cultural rules and be **appropriate**, *e.g.* Fig. 6 .
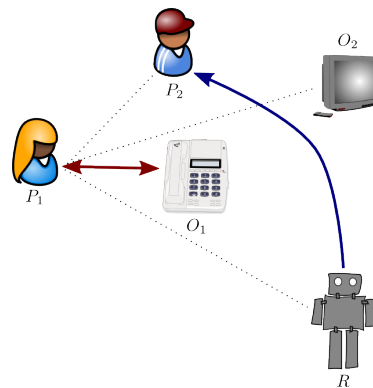
Figure 7. Attention vs activity: although P_1's current activity is being on the phone, part of her attention may be directed towards P_2, the TV set or the robot R. Suppose now that R moves towards P_2 in a way that hides the TV from P_1. Such a behavior would not be appropriate should P_1 be actually paying attention to the TV.

Most of the approaches proposed so far relies upon the definition of so-called **social spaces**, *i.e.* regions in the environment that, for different reasons, the persons consider as psychologically theirs. Such social spaces are primarily characterized using either the position of the person, *e.g.* "Personal space" [42], or the activity it is currently engaged in, *e.g.* "Interaction Space" [47] and "Activity Space" [51]. The most common approach is then to define costmaps on such social spaces: the higher the cost, the less desirable it is for the robot to be at the corresponding position. The costmaps are ultimately used for motion planning and navigation purposes. Such approaches are interesting however their spatial nature (being inside or outside the space) make them less suitable when facing more complicated situations, *e.g.* Fig. 7 . To overcome those limits, we suggest using the psychological concept of **attention**, which plays a central role when humans navigate around each other. Besides lifting the limits of social spaces, this concept brings a new degree of control over the motion of the robot, namely the invasive and distracting character of the robot motion, which have so far proven hard to tackle with the conventional tools such as social spaces. Beside leading appropriate motion, attention-based navigation enable interaction through motion by predicting the quantity of attention the human will give to the robot.

Building upon a computational model of attention that was earlier proposed in [53]. we have developed the novel concept of **attention field**. The attention field is straightforward to define: it is a measure of the amount of attention that a given person would allocate to the robot, should the robot be in a given position/state. It is mapping from the state space of the robot to $\mathbb{R}$. The attention field can serve as an attention predictor that can be used to predict potential attentional situations. This knowledge can in turn be used to decide what the robot should do in the future depending on its current task.

Let us illustrate this on a simple scenario featuring a person, a TV and a robot (denoted by P_1, O_1 and R in Fig. 8 -left). The person is currently watching the TV: this is his current *activity*. This activity relates to his *intention* and is modeled by the yellow vector $\overrightarrow{I}$ in Fig. 8 -left that is directed from the person to the TV. Fig. 8 -right depicts the attention field for the person; it is a mapping from $\mathbb{R}^2$ to $\mathbb{R}$ that gives the amount of attention that the person is paying to the robot when it is at a given position (x, y). Fig. 8 -right should be interpreted as follows: the warmer the color, the higher the amount of attention given by the person to the robot. It integrates both the visual and auditory perception capabilities of P_1.

In 2014, we have furthered the development of the concept of attention field and demonstrated different ways to use its attention prediction capability on various scenarios. The main results obtained have been reported in a conference article that has been nominated for the Best Paper Award [14]. Work is ongoing to quantify
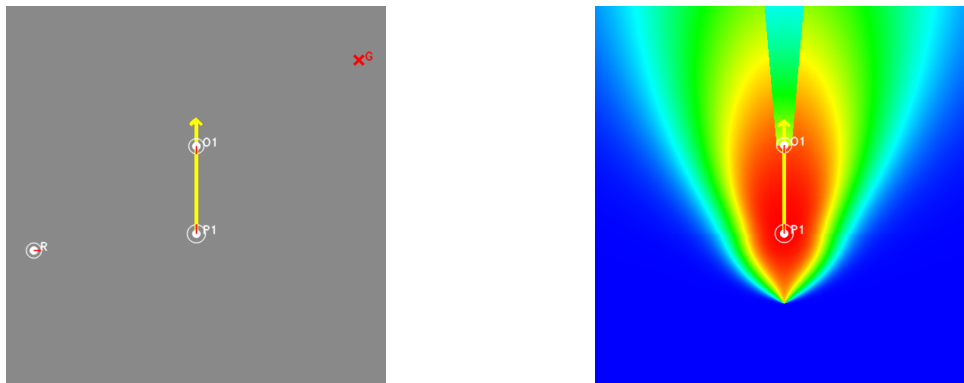
Figure 8. Person-TV-Robot scenario (left), Attention field for the person P_1 (right).

the social "goodness" of the paths provided by our approach, to further the use of the concept of attention on more challenging and dynamic scenarios and to offer an approach to fill the gap between appropriate motion and interaction through motion.

5.3. SPOK: End User Programming for Smart Homes

Participant: Alexandre Demeure.

As part of the CATRENE project AppsGate, we have developed SPOK, an End User Development Environment, that enables inhabitants to control and program their smart Homes via a web interface. The current version of SPOK includes an editor for editing programs using a pseudo-natural language and an interpreter. A multi-syntax editor as well as additional services such as a debugger and a simulator are expected for the second version.

A multi-syntax editor will allow users to build syntactically correct programs using the syntax that is most appropriate to them or by using a combination of them. These syntaxes include pseudo-natural language (i.e. a constrained natural language) and graphical iconic syntax (as exemplified by Scratch [Maloney et al. 2010]). The interaction techniques used to enter programs may be menu-based, free typing, as well as by demonstration in the physical home or by the way of the simulator. The simulator is the dual digital representation of the real home. It is intended to serve also as a debugger for testing and correcting end-user programs.

Whatever syntax used by end-users, programs are translated into syntactic abstract trees whose leaves reference services provided by the Core HMI and/or by the Extended HMI Middleware. The interpreter, executes end-user programs, using the corresponding syntactic abstract trees as input.

In order to support a dynamically extensible grammar as well as to provide end-users with feedforward at the user interface of the editor, the grammar used by the editor is split into 2 parts: the root grammar and the device specific grammars. The root grammar specifies the generic structures of an end-user program: loops, conditions, etc. The device specific grammars are separated from the root grammar to be able to dynamically build the final grammar to be compliant with what is currently installed and detected by the AppsGate server. Each device type brings with it its own events, status and actions. These grammatical elements are injected into the root grammar when generating the parser and for compiling end-user programs.

The language used by end-users to express their programs is a pseudo-natural language using the rule-based programming paradigm. The left hand side of a rule is composed of events and conditions, and the right hand side specifies the actions to be taken when the left hand-side is true or becomes true. A program may include

several rules that can be executed either in parallel or sequentially. Once entered, programs are translated into syntactic abstract trees. The interpreter, executes end-user programs, using the corresponding syntactic abstract trees as input. SPOK is implemented as a mix of OSGi and ApAM components where ApAM is in turn a middleware that runs on top of OSGi.

5.4. Qualitative approaches for building energy management

Participant: Patrick Reignier.

Reducing housing energy costs is a major challenge of the 21st century. In the near future, the main issue for building construction is the thermal insulation, but in the longer term, the issues are those of renewable energy (solar, wind, etc.) and smart buildings. Home automation system basically consists of household appliances linked via a communication network allowing interactions for control purposes. Thanks to this network, a load management mechanism can be carried out: it is called distributed control. An optimal home energy management system is still a goal to aim for, because lots of aspects are still not completely fulfilled. Most of the energy systems respect only the energy needs, but they don't tackle the user needs or satisfaction. Energy systems also have a lack when it comes to the dynamicity of the environments (the system ability to adapt).The problem is similar for the existing HMI (Human User Interface) of those Home Automation Systems where only experts can understand the data coming from the sensors and most important, the energy plan coming from management system (How? and Why?). The goal of this study is to propose a house energy model that can be both used to predict at some level energy evolution and that can be understood by the end user. The house energy model is based on Fuzzy Cognitive Maps representing cause-effects relations. It is first designed by an expert and then automatically tuned to a particular house using machine learning approaches. Preliminary experiments have been done this year using the Predis datasets.