# INRIA, Evaluation of Theme
# Stochastic Models and Methods

Project-team ASPI

March 2014

**Project-team title:**
**Applications of Interacting Particle Systems to Statistics**

**Scientific leader:** François Le Gland

**Research center:** Rennes Bretagne—Atlantique

**Common project-team with:**
université de Rennes 1, université Rennes 2, CNRS

## 1 Personnel

### Personnel (March 2010)

| | Misc. | INRIA | CNRS | University | **Total** |
|---|---|---|---|---|---|
| DR/Professors | | 1 | | | **1** |
| CR/Assistant Professors | | 1 | | 1 | **2** |
| Permanent Engineers | | | | | |
| Temporary Engineers | | | | | |
| PhD Students | 3 | | | | **3** |
| Post-Doc. | | 1 | | | **1** |
| **Total** | **3** | **3** | | **1** | **7** |
| External Collaborators | | | | | |
| Visitors (> 1 month) | | | | | |

DR (directeur de recherche): senior research scientist
CR (chargé de Recherche): junior research scientist
Permanent Engineers: on a permanent position with CNRS, INRIA, etc.
Temporary Engineers (ingénieur expert, ingénieur associé): supported by a contract

### Personnel (March 2014)

| | Misc. | INRIA | CNRS | University | **Total** |
|---|---|---|---|---|---|
| DR/Professors | | 1 | | 1 | **2** |
| CR/Assistant Professors | | 1 | | 1 | **2** |
| Permanent Engineers | | | | | |
| Temporary Engineers | | | | | |
| PhD Students | 2 | | | 1 | **3** |
| Post-Doc. | | | | | |
| **Total** | **2** | **2** | | **3** | **7** |
| External Collaborators | | | | | |
| Visitors (> 1 month) | | | | | |

### Changes in staff

| DR/Professors<br>CR/Assistant Professors | Misc. | INRIA | CNRS | University | **Total** |
|---|---|---|---|---|---|
| Arrival | | | | +2 | **+2** |
| Leaving | | | | −1 | **−1** |

### Comments

Since January 2012, ASPI is a joint (INRIA, université de Rennes 1, université Rennes 2, CNRS) project–team, and is part of IRMAR (institut de recherche mathématique de Rennes, UMR 6625).
Florent Malrieu (assistant professor at université de Rennes 1) joined the team in January 2012. He has obtained a promotion (as a professor at université François Rabelais, Tours) and left the team in September 2013.
Valérie Monbet (professor at université de Rennes 1) joined the team in January 2012.

### Current composition of the project-team (March 2014):

Permanent staff

- Frédéric Cérou, chargé de recherche, INRIA.
- Arnaud Guyader, maître de conférence, université Rennes 2.
- François Le Gland, directeur de recherche, INRIA.
- Valérie Monbet, professeur, université de Rennes 1.

Project assistant

- Fabienne Cuyollaa, INRIA.

Non–permanent staff

- Julie Bessac, PhD student, université de Rennes 1.
- Damien–Barthélémy Jacquemart, PhD student, ONERA Palaiseau.
- Alexandre Lepoutre, PhD student, ONERA Palaiseau.

**Current position of former project-team members (including PhD students during the March 2010–March 2014 period):**

- Florent Malrieu, professor at université François Rabelais, Tours, since September 2013.
- Paul Bui Quang, PhD student, defense in July 2013 — post–doc at CEA DAM in Bruyères–le–Chatel since October 2013.
- Nicolas Jégou, PhD student, defense in November 2012 — PRCE, université Rennes 2.
- Rudy Pastel, PhD student, defense in February 2012 — post–doc at BASF in Ludwigshafen from March 2012 to April 2013 — technology consultant at Zielpuls GmbH in Munich since November 2013.
- Anindya Goswami, post–doc until mid August 2010 — post–doc at the Technion from September 2010 to August 2011 — assistant professor at IISER in Pune since September 2011.
- Nordine El Baraka, PhD student, dropped in August 2010 (after 36 months) without defending his thesis.
- Adrien Ickowicz, PhD student, defense in May 2010 — ATER at université Paris Dauphine from September 2010 to August 2011 — post–doc at Télécom Lille from September 2011 to August 2012 — post–doc at CSIRO Mathematics, Informatics and Statistics in Sydney since September 2012.

### Last INRIA enlistments

No INRIA enlistment has occurred during the evaluation period.

# 2 Work progress

## 2.1 Keywords

sequential Monte Carlo (SMC) method, interacting particle system, stochastic hybrid system, Bayesian filtering, particle filtering, localization, navigation and tracking, sequential data assimilation, ensemble Kalman filter, stochastic weather generator, rare event simulation, risk evaluation, multilevel splitting, collision risk, fingerprinting and watermarking, global optimization, sensor management, machine learning, nearest neighbor estimation and classification, regression estimation, non–parametric statistics, approximate Bayesian computation (ABC).

## 2.2 Context and overall goal of the project

The scientific objectives of the ASPI team are the design, practical implementation and mathematical analysis of interacting Monte Carlo methods, also known as particle methods or sequential Monte Carlo (SMC) methods, with focus on

- statistical inference in hidden Markov models and particle filtering,
- risk evaluation and simulation of rare events,
- global optimization.

Intuitively speaking, interacting Monte Carlo methods are sequential simulation methods, in which particles

- *explore* the state space by mimicking the evolution of an underlying random process,
- *learn* the environment by evaluating a fitness function,
- and *interact* so that only the most successful particles (in view of the value of the fitness function) are allowed to survive and to get offsprings at the next generation.

The effect of this mutation/selection mechanism is to automatically concentrate particles (i.e. the available computing power) in regions of interest of the state space. The whole context is multidisciplinary, not only because of the many scientific and engineering areas in which these particle methods are used, but also because of the diversity of the scientific communities which have already contributed to establish the foundations of the field

> target tracking, interacting particle systems, empirical processes, genetic algorithms (GA), hidden Markov models and nonlinear filtering, Bayesian statistics, Markov chain Monte Carlo (MCMC) methods, etc.

In the special case of particle filtering, which has numerous applications (under the generic heading of positioning, navigation and tracking) in target tracking, computer vision, mobile robotics, ubiquitous computing and ambient intelligence, sensor networks, etc., each particle represents a possible hidden state, and is multiplied or terminated at the next generation on the basis of its consistency with the current observation, as quantified by the likelihood function. With these genetic–type algorithms, it becomes easy to efficiently combine a prior model of displacement with or without constraints, sensor–based measurements, and a base of reference measurements, for example in the form of a digital map (digital elevation map, attenuation map, etc.).

In the special case of splitting or branching methods for the estimation of the small probability of a rare but critical event, for instance the probability that a random process reaches a critical region of the state space (a major issue in industrial areas such as nuclear power plants, telecommunication networks, protection of digital documents, finance and insurance industry, defence systems, air traffic management, etc.), intermediate less critical regions are introduced, and trajectories hitting an intermediate region, i.e. going potentially towards the most critical region, are given offsprings, thus increasing the number of trajectories that eventually hit the most critical region. The same strategy also applies in static cases, for the estimation of the small probability that a random variable, defined as the output of some (large, complex, costly to execute) numerical code with random input, exceeds some extreme level.

In the most general case, particle methods provide approximations of Feynman–Kac distributions, a pathwise generalization of Gibbs–Boltzmann distributions, by means of the weighted empirical probability distribution associated with an interacting particle system, which provides a very powerful mathematical framework for the mathematical analysis of practical algorithms, mostly in the large sample asymptotics, i.e. as the number of particles grows to infinity:

> convergence in $p$–th mean, uniformly over time or over a class of functions, central limit theorem, propagation of chaos, probability of deviation, etc.

Issues to be considered here are:

> how to resample, when to resample, which Markov model and selection functions to use (as in importance sampling), how many particles to use, etc.

and what is the impact of these different designs on the approximation. Many results in this direction have been obtained by Pierre Del Moral (EPI ALEA, INRIA Bordeaux — Sud Ouest), and can be found in the textbooks [13], [14] and in more recent papers. Yet, ASPI has also contributed and is still contributing on its own to this theory.

In summary, ASPI essentially carries methodological research activities, with the objective to obtain generic results with high potential for applications, and to implement these results on a few practical examples, through continuing collaboration and partnership with public research institutions or with industrial partners, such as

ONERA (French aerospace lab) and DGA (French defence agency).

Applications currently or recently considered are

> geolocalization and tracking of mobile terminals, terrain–aided navigation, information fusion for indoor localization, optimization of sensors position and activation, risk assessment in air traffic management, collision risk, and protection of digital documents.

### 2.3 Objectives for the evaluation period

The objectives for the previous evaluation period were formulated in the proposal submitted in March 2004:

1. Methodology of particle methods.
2. Simulation of rare events.
3. Particle filtering and statistics of hidden Markov models.

An additional topic, that was not present in the initial list of objectives, has emerged during the previous evaluation period:

4. Functional data analysis and classification.

Based on the written output (comments, critics, advice) provided by the experts in their report after the last evaluation held in March 2010, the following answer was formulated by ASPI.

*We see in this evaluation report a confirmation of the organization of our activities along three axes: (i) mathematical analysis of particle methods, (ii) rare event simulation, described as the winner in our portfolio, and more traditionally (iii) particle filtering, with its applications in localization, navigation and tracking.*

*We do agree with the experts that the practical implications of rare event simulation appear in all kind of domains, and that our efforts should be concentrated in this direction. Besides, we believe that the paradigm of multilevel splitting can also be successfully applied to other areas connected with, but not limited to, rare events. One such area, already listed in our synthesis document, is global optimization of some criterion. The idea here is to generate a population of particles living in intermediate regions defined by the criterion exceeding some level, and to adaptively increase this level until the maximum of the criterion is reached. Another such area, that was not yet listed in our synthesis document, but which has already generated intense activity since then, in collaboration with Tony Lelièvre (CERMICS and ENPC, Marne–la–Vallée and EPI MICMAC, INRIA Paris — Rocquencourt), is molecular simulation. The idea here is to study the transition between* *two metastable states or configurations of a complex molecule, defined in terms of a potential landscape. This transition can be seen as a rare event, and one is not only interested in the accurate evaluation of the probability of the transition to occur, but also in learning which are the typical paths or transformations that bring the molecule from one metastable state or configuration to the other. While importance sampling is the popular method to answer the first question, it fails to address the second question, which leaves some hope for our splitting methods.*

*More traditional applications of rare event simulation will also be considered as well. This includes the important question, in a static context, of evaluating the probability that the output variable of a black–box system, described in terms of a complicated transformation or just a computer program, exceeds some level when the input variables are random, with a known probability distribution.*

*Another issue that has been discussed with the experts, and also inside the team after the evaluation seminar, was the need to continue with non–parametric statistics as another objective for the team. Indeed, some important results on the convergence and convergence rate of the nearest neighbour estimator in infinite dimension have been obtained recently by Frédéric Cérou and Arnaud Guyader, in collaboration with Gérard Biau (Université Pierre et Marie Curie and ENS Paris, and EPI CLASSIC, INRIA Paris — Rocquencourt). These are essentially theoretical results, but a practical application is missing so far. Investigating the statistical properties of recommendation systems has been considered for a while as a potential suitable application. Our own conclusions, and the recommendations of the experts as well, are that we should not give up our research program on particle methods, which also means, as a counterpart, that we do not plan to invest much more on applications of functional data classification to recommendation systems.*

As a result, the outcome of the last evaluation held in March 2010, was a slight redefinition of the research objectives of ASPI for the current evaluation period:

1. Simulation of rare events, with coming up activity in
   - molecular simulation,
   - global optimization.
2. Particle filtering, with focus on
   - localization, navigation and tracking,
   - sequential data assimilation.
3. Non–parametric statistics.

The recent activity of ASPI will be described below along these three research directions.

### 2.4 Objective 1: Simulation of rare events (executive summary)

The estimation of the small probability of a rare but critical event, is a crucial issue in industrial areas such as

> nuclear power plants, telecommunication networks, protection of digital documents, finance and insurance industry, defence systems, air traffic management, etc.

In such complex systems, analytical methods cannot be used, and naive Monte Carlo methods are clearly inefficient to estimate accurately very small probabilities. Besides importance sampling, an alternate widespread technique consists in multilevel splitting [18, 19, 25], where trajectories going towards the critical set are given offsprings, thus increasing the number of trajectories that eventually reach the critical set. As shown in [7], the Feynman–Kac formalism is well suited for the design and analysis of splitting algorithms for rare event simulation.

**Dynamical case** Multilevel splitting can be used for rare event simulation problems in dynamical situations. To be specific, consider a complex dynamical system modelled as a Markov process, whose state can possibly contain continuous components and finite components (mode, regime, etc.), and the objective is to compute the probability, hopefully very small, that a critical region of the state space is reached by the Markov process before a final time $T$, which can be deterministic and fixed, or random (for instance the time of return to a recurrent set, corresponding to a nominal behaviour).

**Static case** Multilevel splitting can also be used for rare event simulation (also called propagation of uncertainty) problems in static situations. Here, the objective is to learn the probability distribution of an output random variable $Y = F(X)$, where the function $F$ is only defined pointwise for instance by a computer programme, and where the probability distribution of the input random variable $X$ is known and easy to simulate from. More specifically, the objective could be to compute the probability of the output random variable exceeding a threshold, or more generally to evaluate the cumulative distribution function of the output random variable for different output thresholds. This problem is characterized by the lack of an analytical expression for the function $F$, the computational cost of a single pointwise evaluation of the function, which means that the number of calls to the function should be limited as much as possible.

The starting point for the growing interest of ASPI in rare event simulation has been air traffic management, considered in two European projects (HYBRIDGE and iFLY) coordinated by NLR. During the previous evaluation period, a major application area of rare event simulation has been the reliability of protection mechanisms for digital documents, i.e. the evaluation of (small) probabilities of false alarm in watermarking or fingerprinting schemes. This activity was considered in the ANR project NEBBIANO, within the SETIN (Sécurité et Informatique) programme. Applications to air traffic (and more generally to aeronautics and space, in collaboration with ONERA/DCPS Palaiseau), and to protection of digital documents, are still actively considered in ASPI. Applications to molecular simulation and to global optimization have been included, as a result of the last evaluation, in the objectives of ASPI for the current evaluation period. The interest in molecular simulation originated from contacts that turned into active collaboration with Tony Lelièvre and his group (CERMICS, ENPC and EPI MATHERIALS, INRIA Paris — Rocquencourt). The interest in global optimization is a legacy of Jean–Pierre Le Cadre († in July 2009). It has grown during a first contract with DGA/Techniques Navales (previously known as CTSN), and continues with DUCATI, a second contract that also involves ONERA/DTIM Palaiseau.

### 2.4.1 Personnel

Permanent researchers: Frédéric Cérou, Arnaud Guyader, François Le Gland, Florent Malrieu.

PhD students: Rudy Pastel, Damien Jacquemart.

### 2.4.2 Project-team positioning

The question of rare event simulation is very large and occurs in many different domains, hence the picture given here is certainly incomplete. However, important contributors definitely include Paul Glasserman (Columbia University) for his early mathematical analysis of multilevel splitting, Zdravko Botev and Dirk Kroese (University of Queensland) for their contribution to multilevel splitting and their introduction of the cross–entropy method together with Reuven Rubinstein († in December 2012), and Paul Dupuis (Brown University) for his construction of efficient importance distributions (for importance sampling) and importance functions (for multilevel splitting), using large deviations arguments. Within INRIA, Bruno Tuffin (EPI DIONYSOS, INRIA Rennes — Bretagne Atlantique) is also a very active contributor, especially through collaboration with Pierre L'Écuyer (université de Montréal), and Tony Lelièvre and his group (CERMICS, ENPC and EPI MATHERIALS, INRIA Paris — Rocquencourt) is a leading expert in Monte Carlo methods in molecular simulation.

### 2.4.3 Scientific achievements

During the current evaluation period, a major application area of rare event simulation was molecular simulation. The idea here is to study the transition between two metastable states or configurations of a complex molecule, defined in terms of a potential landscape. This transition can be seen as a rare event, and one is not only interested in the accurate evaluation of the probability of the transition to occur, but also in learning which are the typical paths or transformations that bring the molecule from one metastable state or configuration to the other, how long is the duration of such a transformation, assuming that it occurs. While importance sampling is the popular method to answer the first question, it fails to address the second question, which leaves some hope for the splitting methods developped in ASPI.

**A multiple replica approach to simulate reactive trajectories** A method to generate reactive trajectories, namely equilibrium trajectories leaving a metastable state and ending in another one, has been proposed in [14]. The algorithm is based on simulating in parallel many copies of the system, and selecting the replicas which have reached the highest values along a chosen one–dimensional reaction coordinate. This reaction coordinate does not need to precisely describe all the metastabilities of the system for the method to give reliable results. An extension of the algorithm to compute transition times from one metastable state to another one is also presented. We demonstrate the interest of the method on two simple cases: a one–dimensional two–well potential and a two–dimensional potential exhibiting two channels to pass from one metastable state to another one.

Collaboration with Tony Lelièvre (CERMICS, ENPC and EPI MATHERIALS, INRIA Paris — Rocquencourt).

**On the length of one–dimensional reactive paths** Motivated by some numerical observations on molecular dynamics simulations, metastable trajectories are studied in a very simple setting, namely paths generated by a one–dimensional overdamped Langevin equation for a double well potential. More precisely, interest is in so–called reactive paths, namely trajectories which leave definitely one well and reach the other one. The objective in [13] is to precisely analyze the distribution of the lengths of reactive paths in the limit of small temperature, and to compare the theoretical results to numerical results obtained by a Monte Carlo method, namely the multilevel splitting approach.

Collaboration with Tony Lelièvre (CERMICS, ENPC and EPI MATHERIALS, INRIA Paris — Rocquencourt).

The next result is a continuation of the activity started in the previous evaluation period, on applications of rare event simulation to the reliability of protection mechanisms for digital documents.

**Decoding fingerprints using the MCMC method** A new fingerprinting decoder is proposed in [7] based on the Markov chain Monte Carlo (MCMC) method. A Gibbs sampler generates groups of users according to the posterior probability that these users could have forged the sequence extracted from the pirated content. The marginal probability that a given user pertains to the collusion is then estimated by a Monte Carlo method. The users having the biggest empirical marginal probabilities are accused. This MCMC method can decode any type of fingerprinting codes. It is in the spirit of the *Learn and Match* decoding strategy: it assumes that the collusion attack belongs to a family of models. The EM algorithm estimates the parameters of the collusion model from the extracted sequence. This part of the algorithm is described for the binary Tardos code and with the exploitation of the soft outputs of the watermarking decoder. The experimental body considers some extreme setups where the fingerprinting code lengths are very small. It reveals that the weak link of our approach is the estimation part. This is a clear warning to the *Learn and Match* decoding strategy.

Collaboration with Teddy Furon (EPI TEXMEX, INRIA Rennes — Bretagne Atlantique).

The remaining part is devoted to presenting scientific achievements in the methodology of splitting methods.

**Rare event simulation for a static distribution** The problem considered here is to evaluate the probability that the output variable $Y = F(X)$ of a black–box system exceeds a given critical threshold. In its simplest (non–adaptive) version, the proposed splitting methods consists in (i) introducing a sequence of intermediate regions in the input space, implicitly defined by exceeding an increasing sequence of thresholds or levels, (ii) counting the fraction of samples that reach a level given that the previous level has been reached already, and (iii) improving the diversity of the selected samples, usually using an artificial Markovian dynamics. In this way, the algorithm learns

- the transition probability between successive levels, hence the probability of reaching each intermediate level,
- and the probability distribution of the input random variable, conditioned on the output variable reaching each intermediate level.

An adaptive version of the algorithm, where the intermediate thresholds are defined as quantiles of the (random) output variable is also considered. The contribution of [11] has been to study the impact of design parameters, such as the intermediate levels or the Metropolis kernel introduced in the mutation step, on the asymptotic variance obtained through a central limit theorem.

Collaboration with Pierre Del Moral (EPI ALEA, INRIA Bordeaux — Sud Ouest) and with Teddy Furon (EPI TEXMEX, INRIA Rennes — Bretagne Atlantique).

An improved version of the algorithm has been considered in [20], in which the next intermediate level is defined as the minimum of the sampled output variable, and only the sample that achieves the minimum is removed and resampled. This algorithm improves upon existing multilevel splitting methods and can be analyzed using Poisson process tools that lead to exact and nonasymptotic description of the distribution of the estimated probabilities and quantiles. The performance of the algorithm is demonstrated in a problem related to digital watermarking.

Collaboration with Nicolas Hengartner (LANL) and Éric Matzner–Løber (université Rennes 2).

Several applications to space and aeronautics have been considered by Rudy Pastel in his PhD thesis [6]: evaluation of the collision risk between satellite and space debris in [29], evaluation of the deviation from the nominal impact position of a missile or a launcher in [28].

Collaboration with Jérôme Morio (ONERA/DCPS Palaiseau).

**Rare event simulation in stochastic hybrid systems** A rare but critical event (such as a scalar function of the continuous component of the state exceeding a given threshold) can occur under different circumstances:

- the process can remain in *nominal* mode, where the critical event is very unlikely to occur,
- or the process can switch in some *degraded* mode, where the critical event is much more likely to occur, but the switching itself is very unlikely to occur.

Not only is it important to accurately estimate the (very small) probability that the critical event occurs before some fixed final time, but it is also important to have an accurate account on the reason why it occured, or in other words to estimate the probability of the different modes in the critical regime. A classical implementation of the multilevel splitting would not be efficient. Indeed, as soon as (even a few) sample paths switch to a *degraded* mode, these sample paths will dominate and it will not be possible to estimate the contribution of samples paths in the *nominal* mode. Moreover, sampling the finite component of the state is not efficient to accurately estimate the (very small) probability of rare but critical modes. Indeed, making sure that sufficiently many trajectories are living on each of the many different modes may require a huge number of trajectories, making the method impractical. Several variance reduction methods have already been proposed by Jaroslav Krystul in his PhD thesis [23] and in [2], and some new methods have been studied to handle this situation and reduce the variance:

- stratification can be used to make sure that at each resampling stage, a fixed number of trajectories is allocated to each mode, resulting in *sampling per mode* strategy,
- conditioning, or Rao–Blackwell marginalization, can be used to avoid sending trajectories exploring the many different modes, and to use instead more efficiently the analytical expression, provided by a Wonham filter, for the conditional probability vector of the mode given the continuous component of the state.

An algorithm implementing *sampling per mode* has been studied in [23]: convergence results and central limit theorems are proved, and the expression of the asymptotic variance is discussed in each case, so as to provide guidelines for the choice of design parameters (sample size per mode, mode–dependent intermediate regions, etc.). An algorithm implementing marginalization has been presented in [8], but its theoretical study remains to be completed and belongs to the future objectives for the next evaluation period.

Collaboration with Pascal Lezaud (DGAC/DSNA and ENAC), Jaroslav Krystul (Twente University) and Anindya Goswami (IISER Pune). It was initiated during the European project iFLY coordinated by Henk Blom (NLR).

**Combining importance sampling and multilevel splitting** Consider the problem of accurately estimating the (very small) probability that a rare but critical event (such as a scalar function of the state of a Markov process exceeding a given threshold) occurs before some fixed final time. Multilevel splitting is in principle a very efficient solution, in which sample paths are propagated and are replicated when some intermediate events occur. However, intermediate events that are defined in terms of the state variable only (such as a scalar function of the state exceeding an intermediate threshold) are not a good design. A more efficient but more complicated design would be to let the intermediate events depend also on time. An alternative design studied in [9] is to keep intermediate events simple, defined in terms of the state variable only, and to make sure that samples that exceed the threshold early are replicated more than samples that exceed the same threshold later.

Collaboration with Jérôme Morio (ONERA/DCPS Palaiseau).

**Multilevel splitting for global optimization** An issue closely related to rare event simulation is global optimization. The idea is to replace the difficult problem of finding the set of global maxima of a real–valued cost function by the apparently simpler problem of sampling a population from a reference probability distribution conditionned on the cost function exceeding a threshold. Asymptotically as the threshold grows, the population concentrates on the set of global maxima of the cost function. An early approach has been the cross–entropy method [29, 12], which relies on learning the optimal importance distribution within a prescribed parametric family, and new algorithms based on multilevel splitting have been proposed recently [22, 3, 28, 4].

Theoretical study of these new algorithms belong to the future objectives for the next evaluation period. During the present evaluation period, preliminary results and extensive numerical experience, including for multi–objective optimization problems, have already been obtained by Mathieu Chouchane in his PhD thesis [11], with application to the optimization of sensors position and activation [16].

Collaboration with Mathieu Chouchane and Sébastien Paris (université Paul Cézanne and LSIS).

**Smoothing splitting method for counting** An improved version of the splitting method, called the smoothed splitting method (SSM), has been developped in [15] for counting problems associated with complex sets, in particular for counting the number of satisfiability assignments. A satisfiability problem consists in several logical clauses involving several Boolean variables (typically several hundreds or thousands each). The goal is to find (if any) all the instances of the variables (0 or 1) which make all the clauses true. This is well known as a NP–hard problem if we want to solve it exactly. An approximate new solver is proposed, based on rare event simulation techniques. Like the conventional splitting algorithms, the proposed algorithm uses a sequential sampling plan to decompose a *difficult* problem into a sequence of *easy* ones. The main difference between SSM and conventional splitting is that it works with an auxiliary sequence of continuous sets instead of the original discrete sets. The rationale of doing so is that continuous sets are easier to handle. We have shown on several examples that while SSM and its conventional splitting counterpart are similar in their CPU time and variability, SSM is more robust and more flexible. In particular, it makes it simpler to tune the parameters.

Collaboration with Reuven Rubinstein († in December 2012) and Radislav Vaisman (Technion, Israel Institute of Technology).

### 2.4.4 Collaborations

Collaboration with Tony Lelièvre (CERMICS, ENPC and EPI MATHERIALS, INRIA Paris — Rocquencourt) on molecular simulation, with Teddy Furon (EPI TEXMEX, INRIA Rennes — Bretagne Atlantique) on reliability of watermarking and fingerprinting, with Pierre Del Moral (EPI ALEA, INRIA Bordeaux — Sud Ouest) on rare event simulation for a static distribution, with Nicolas Hengartner (LANL) and Éric Matzner–Løber (université Rennes 2) on Monte Carlo estimation of extreme quantiles of posterior probabilities, with Pascal Lezaud (DGAC/DSNA and ENAC), Henk Blom (NLR), Jaroslav Krystul (Twente University) and Anindya Goswami (IISER Pune) on rare event simulation in stochastic hybrid systems, with Jérôme Morio (ONERA/DCPS Palaiseau) on rare event probabilities in complex aerospace systems, with Mathieu Chouchane and Sébastien Paris (université Paul Cézanne and LSIS) and Christian Musso (ONERA/DTIM Palaiseau) on splitting methods for global optimization, with Reuven Rubinstein († in December 2012) and Radislav Vaisman (Technion, Israel Institute of Technology) on multilevel splitting for counting.

### 2.4.5 External support

Research on splitting methods is/has been supported by the European project iFLY on highly automated air traffic management, within the FP6 Aeronautics and Space programme, and by ONERA/DCPS through two scholarships, one on static rare event simulation with aeronautics and space applications, and the second on dynamical rare event simulation for the evaluation of collision risk.

Applications to global optimization is supported by DGA/Techniques Navales under a series of two contracts on optimization of sensors position and activation (the second contract is another collaboration with ONERA/DTIM).

### 2.4.6 Self assessment

Of the three objectives presented here, this is the one where more effort was put during the evaluation period. Four years ago, one could have argued that ASPI was a newcomer and an outsider in this area, but this is not quite the case today and some of our contributions are now well–recognized. Not only ASPI has already produced important and recognized contributions, but it has also provided a very efficient and general–purpose methodology to analyze the performance of already existing or new and original splitting methods, under very general assumptions. The effort put here should be continued, especially in the same two directions: molecular simulation, where a fruitful collaboration is already effective with leading experts, and global optimization, where a complex real–world application is already available.

### 2.5 Objective 2: Particle filtering (executive summary)

During the evaluation period, two different application domains of particle filtering have been investigated in ASPI:

- localization, navigation and tracking,
- sequential data assimilation.

The objective in localization, navigation and tracking, see [21, 27, 32] for instance, is to estimate the position (and also velocity, attitude, etc.) of a mobile object, from the combination of different sources of information, including

- a prior dynamical model of typical evolutions of the mobile, such as inertial estimates and prior model for inertial errors,
- measurements provided by sensors,
- and possibly a digital map providing some useful feature (terrain altitude, power attenuation, etc.) at each possible position.

In some applications, another useful source of information is provided by

- a map of constrained admissible displacements, for instance in the form of an indoor building map,

which particle methods can easily handle (map-matching). This Bayesian dynamical estimation problem is also called filtering, and its numerical implementation using particle methods, known as particle filtering, has been introduced in the early 90's by the target tracking community [20], which has already contributed to many of the most interesting algorithmic improvements and is still very active, and has found applications in

> target tracking, integrated navigation, points and/or objects tracking in video sequences, mobile robotics, wireless communications, ubiquitous computing and ambient intelligence, sensor networks, etc.

ASPI is contributing to several applications of particle filtering in localization, navigation and tracking, such as

> geolocalization and tracking of mobile terminals, terrain–aided navigation and information fusion for indoor localization.

Sequential data assimilation deals essentially with the same estimation problem, with applications focused in oceanography and meteorology, and with some specific features related with high–dimension. The traditionnal trend in geophysical sciences is to use as prior information some deterministic models formulated in terms of PDEs and reflecting as much as possible the physics. An emerging trend is to use also simpler statistical models obtained by processing time series data collected in situ, or by processing image sequences provided by satellite observations.

The continuing interest in localization, navigation and tracking finds its roots in the well–established expertise of ASPI. The growing interest in sequential data assimilation has started on the occasion of the ARC project ADOQA coordinated by Isabelle Herlin (EPI CLIME, INRIA Paris — Rocquencourt), followed by the ANR project PREVASSEMBLE. This activity became even more important when Valérie Monbet obtained a position at université de Rennes 1 in September 2010, and joined ASPI later on, in January 2012.

### 2.5.1 Personnel

Permanent researchers: François Le Gland, Valérie Monbet.

PhD students: Julie Bessac, Paul Bui–Quang, Alexandre Lepoutre.

### 2.5.2 Project-team positioning

Particle filtering and its applications to localization, navigation and tracking, has raised huge interest worldwide, across many scientific domains, such as target tracking, video tracking, mobile robotics, etc., with main contributors Fredrik Gustafsson and his group (Linköping University), Simon Maskell and David Salmond (QinetiQ, Malvern and Farnborough), Neil Gordon (DSTO, Adelaide), Yvo Boers and Hans Driessen (Thalès Nederland, Hengelo), Ba–Ngu Vo (University of Melbourne), Christian Musso (ONERA/DTIM, Palaiseau), Petar Djurić (SUNY at Stony Brook), Dieter Fox (University of Washington), Sebastian Thrun (Stanford University). Within INRIA, François Caron and Pierre Del Moral (EPI ALEA, INRIA Bordeaux — Sud Ouest) are very active contributors.

Some of the main contributors in sequential data assimilation and especially in ensemble Kalman filtering are Geir Evensen (Hydro Research Centre, Bergen) its inventor, Laurent Bertino (NERSC, Bergen), Peter Jan van Leeeuwen (University of Reading), Sebastian Reich (Universität Potsdam and University of Reading), Jeffrey L. Anderson and Chris Snyder (NCAR, Boulder) and Thomas Bengtsson (Bell Labs, Statistics and Data Mining Department, Murray Hill). Within INRIA, Étienne Mémin and his group (EPI FLUMINANCE, INRIA Rennes — Bretagne Atlantique) is also a very active contributor.

### 2.5.3 Scientific achievements

During the current (and the previous) evaluation period, several applications in localization, navigation and tracking have been studied:

- information fusion for indoor localization (FIL),
- geolocalization and tracking in a wireless network,
- terrain–aided navigation,

and a presentation of the first application can be found in Section 4.1. Other issues in particle filtering have also been addressed, in collaboration with ONERA, with the following contributions.

**Laplace method in particle filtering** The Laplace method is an asymptotic method used in a static framework to compute integrals, valid in theory as soon as the function to be integrated exhibits an increasingly dominating maximum point, which brings the essential contribution to the integral. On the other hand, it is well–known that sequential Monte Carlo methods based on importance sampling are inefficient when the weighting function (here, the likelihood function) is too much spatially localized, e.g. when the variance of the observation noise is too small, whereas this is precisely the situation where the Laplace method is efficient and theoretically justified, hence the natural idea of combining the two approaches. An algorithm associating the Laplace method and particle filtering, called the Laplace particle filter has been proposed by Paul Bui Quang in his PhD thesis [1].

Separately, an approximation of the Bayesian filter based on the Laplace method only (i.e. without any generation of random samples) has been proposed and the propagation of the approximation error from one time step to the next time step, has been analyzed in an appropriate asymptotic framework, e.g. when the variance of the observation noise goes to zero, or when the variances of the model noise and of the observation noise jointly go (with the same rate) to zero, or more generally when the information contained in the system goes to infinity, with an interpretation in terms of identifiability.

Collaboration with Christian Musso (ONERA/DTIM Palaiseau).

**TBD (track–before–detect)** Track–before–detect refers to situations where the target SNR is so low that it is practically impossible to detect the presence of a target, using a simple thresholding rule. In such situations, the solution is to keep all the information available in the raw radar data and to address directly the tracking problem, using a particle filter with a binary Markov variable that models the presence or absence of the target.

The choice of the proposal distribution is crucial here, and an efficient particle filter is proposed in [11] that is based on a relevant proposal distribution built from detection and estimation considerations, that aims at extracting all the available information from the measurements. The proposed filter leads to a dramatically improved performance as compared with particle filters based on the classical proposal distribution, both in terms of detection and estimation.

A further improvement, in terms of detection performance, is to model the problem as a quickest change detection problem [31] in a Bayesian framework. In this context, the posterior distribution of the first time of appearance of the target is a mixture where each component represents the hypothesis that the target appeared at a given time. The posterior distribution is intractable in practice, and it is proposed in [10] to approximate each component of the mixture by a particle filter. It turns out that the mixture weights can be computed recursively in terms of quantities that are provided by the different particle filters. The overall filter yields good performance as compared with classical particle filters for track–before–detect.

Collaboration with Olivier Rabaste (ONERA/DEMR Palaiseau).

The remaining part is devoted to presenting scientific achievements in sequential data assimilation.

**Asymptotics of the ensemble Kalman filter** Very little was known about the asymptotic behaviour of the ensemble Kalman filter [15], whereas on the other hand, the asymptotic behaviour of many different classes of particle filters is well understood, as the number of particles goes to infinity. Interpreting the ensemble elements as a population of particles with mean–field interactions, and not only as an instrumental device producing an estimation of the hidden state as the ensemble mean value, it has been possible to prove the convergence of the ensemble Kalman filter, with a rate of order $1/\sqrt{N}$, as the number $N$ of ensemble elements increases to infinity [24]. In addition, the limit of the empirical distribution of the ensemble elements has been exhibited, which differs in general from the Bayesian filter.

The most recent contribution has been to prove (by induction) the asymptotic normality of the estimation error, i.e. to prove a central limit theorem for the ensemble Kalman filter. Explicit expression of the asymptotic variance has been obtained for linear Gaussian systems (where the exact solution is known, and where EnKF is unbiased). This expression has been compared with explicit expressions of the asymptotic variance for two popular particle filters: the bootstrap particle filter and the so–called optimal particle filter, that uses the next observation in the importance distribution.

### 2.5.4 Collaborations

Collaboration on applications in localization, navigation and tracking with Christian Musso (ONERA/DTIM Palaiseau) on the Laplace method and on high–dimensional particle filtering, and with Olivier Rabaste (ONERA/DEMR Palaiseau) on Monte Carlo methods in track–before–detect, respectively.

Collaboration on applications in sequential data assimilation with Étienne Mémin (EPI FLUMINANCE, INRIA Rennes — Bretagne Atlantique), and with Christophe Baehr (Météo–France, centre national de recherche météorologique, Toulouse).

### 2.5.5 External support

Applications in localization, navigation and tracking are/have been supported by ANR under the project FIL on information fusion for indoor localization, within the Télécommunications programme, by ONERA/DTIM trough a scholarship on the Laplace method and high–dimensional particle filtering, and by ONERA/DEMR through another scholarship on Monte Carlo methods in track–before–detect.

Applications in sequential data assimilation have been supported by ANR under the project PREVASSEMBLE on ensemble methods for prediction and data assimilation, within the COSINUS (conception and simulation) programme.

### 2.5.6 Self assessment

ASPI members have been early contributors to particle filtering, which by now has attracted huge interest worldwide and across various domains, with many applications in localization, navigation and tracking. What makes ASPI visible here is not so much applications of particle methods, which many other groups throughout the world do very well within their own scientific domain, but rather the rigorous mathematical analysis of some of the most popular particle algorithms.

ASPI for a while was a complete outsider in the area of sequential data assimilation. With Valérie Monbet joining ASPI recently, and bringing her modelling expertise in oceanography and in meteorology, the situation has improved a lot. Many interesting issues could be addressed in the next evaluation period, in modelling, design of stochastic weather generators, large sample asymptotics, etc.

## 2.6 Objective 3: Nonparametric statistics (executive summary)

In pattern recognition and statistical learning, also known as machine learning, nearest neighbor (NN) algorithms are amongst the simplest but also very powerful algorithms available. Basically, given a training set of data, i.e. an $N$–sample of i.i.d. object–feature pairs, with real–valued features, the question is how to generalize, that is how to guess the feature associated with any new object. To achieve this, one chooses some integer $k$ smaller than $N$, and takes the mean–value of the $k$ features associated with the $k$ objects that are nearest to the new object, for some given metric. In general, there is no way to guess exactly the value of the feature associated with the new object, and the minimal error that can be done is that of the Bayes estimator, which cannot be computed by lack of knowledge of the distribution of the object–feature pair, but the Bayes estimator can be useful to characterize the strength of the method. So the best that can be expected is that the NN estimator converges, say when the sample size $N$ grows, to the Bayes estimator. This is what has been proved in great generality by Stone [30] for the mean square convergence, provided that the object is a finite–dimensional random variable, the feature is a square–integrable random variable, and the ratio $k/N$ goes to 0. Nearest neighbor estimator is not the only local averaging estimator with this property, but it is arguably the simplest. The asymptotic behavior when the sample size grows is well understood in finite dimension, but the situation is radically different in general infinite dimensional spaces, when the objects to be classified are functions, images, etc. Several theoretical advances on this problem have been obtained in [9] and more recently in [9] in collaboration with Gérard Biau (université Pierre et Marie Curie, ENS Paris and EPI CLASSIC, INRIA Paris—Rocquencourt).

### 2.6.1 Personnel

Permanent researchers: Frédéric Cérou, Arnaud Guyader.

### 2.6.2 Project-team positioning

Main contributors include Gérard Biau (université Pierre et Marie Curie, ENS Paris and EPI CLASSIC, INRIA Paris — Rocquencourt), Luc Devroye (McGill University), Gábor Lugosi (Pompeu Fabra University), László Györfi (Budapest University of Technology and Economics), Adam Krzyżak (Concordia University).

### 2.6.3 Scientific achievements

During the current evaluation period, several issues in non–parametric estimation have been addressed, with the following contributions.

**ABC** Approximate Bayesian computation (ABC for short) is a family of computational techniques which offer an almost automated solution in situations where evaluation of the posterior likelihood is computationally prohibitive, or whenever suitable likelihoods are not available. In [7], the procedure is analyzed from the point of view of $k$–nearest neighbor theory and the statistical properties of its outputs is explored. In particular some asymptotic features of the genuine conditional density estimate associated with ABC, which is a new interesting hybrid between a $k$–nearest neighbor and a kernel method, is discussed. In other words, this result also pertains to particle filtering. These are among the very few results on the convergence of ABC, and the assumptions on the underlying probability distribution are minimal.

Collaboration with Gérard Biau (université Pierre et Marie Curie, ENS Paris and EPI CLASSIC, INRIA Paris — Rocquencourt).

**Mutual nearest neighbors** Motivated by promising experimental results, the theoretical properties have been investigated of a recently proposed nonparametric estimator, called the MNN (mutual nearest neighbors) rule, which estimates the regression function $m(x) = E[Y \mid X = x]$ as follows: first identify the $k$ nearest neighbors of $x$ in the sample, then keep only those for which $x$ is itself one of the $k$ nearest neighbors, and finally take the average over the corresponding response variables. It is proved in [18] that this estimator is consistent and that its rate of convergence is optimal. Since the estimate with the optimal rate of convergence depends on the unknown distribution of the observations, adaptation results also hold by data–splitting.

Collaboration with Nicolas Hengartner (LANL).

**Iterative isotone regression** Some theoretical aspects of a recent nonparametric method for estimating a univariate regression function of bounded variation are explored. The method exploits the Jordan decomposition which states that a function of bounded variation can be decomposed as the sum of a non–decreasing function and a non–increasing function. This suggests combining the backfitting algorithm for estimating additive functions with isotonic regression for estimating monotone functions. The resulting iterative algorithm is called IIR (iterative isotonic regression). The main result in [19] states that the estimator is consistent if the number of iterations $k$ grows appropriately with the sample size $N$. The proof requires two auxiliary results that are of interest in and by themselves: the first result generalizes the well–known consistency property of isotonic regression to the framework of a non–monotone regression function, and the second result relates the backfitting algorithm to the von Neumann algorithm in convex analysis. With the geometrical interpretation linking this iterative method with the von Neumann algorithm, and making a connection with the general property of isotonicity of projection onto convex cones, another equivalent algorithm is derived in [21].

Collaborations with Nicolas Hengartner (LANL), Nicolas Jégou and Eric Matzner–Løber (université Rennes 2), and with Alexander B. Németh (Babeş Bolyai University) and Sándor Z. Németh (University of Birmingham).

### 2.6.4 Collaborations

Collaboration with Gérard Biau (université Pierre et Marie Curie, ENS Paris and EPI CLASSIC, INRIA Paris — Rocquencourt) on approximate Bayesian computation (ABC), with Nicolas Hengartner (LANL) on mutual nearest neighbors, with Nicolas Hengartner (LANL) again, Nicolas Jégou and Eric Matzner–Løber (université Rennes 2), Alexander B. Németh (Babeş Bolyai University) and Sándor Z. Németh (University of Birmingham) on iterative isotone regression.

### 2.6.5 External support

No external support was obtained so far. A proposal has been submitted in 2013 to the Associated Team programme (REPAIR, with Los Alamos National Laboratory).

### 2.6.6 Self assessment

Some deep and original results have been obtained here, e.g. in the analysis of nearest neighbors estimation and classification in infinite dimension, or in the analysis of ABC. The effort put here mainly concerns Arnaud Guyader, who has been offered several secondment positions in the past, and has set fruitful collaborations with some major contributors in the domain. This is also a potential threat on this activity, that would probably not survive, should he obtain a (deserved) promotion and leave Rennes in the near future.

## 2.7 Note on stochastic hybrid systems

Stochastic hybrid systems, with continuous/discrete state space, have very powerful modelling abilities and raise many interesting issues. Several examples of these models have been studied during the current evaluation period:

- estimation, simulation and asymptotic behavior of piecewise–deterministic Markov processes (PDMP) is the subject of the ANR project PIECE coordinated by Florent Malrieu,

- stability and long–time behavior of switching diffusions have been recently addressed by Florent Malrieu in [5], [6], [4] and [17], in collaboration with Michel Benaïm (université de Neuchâtel) and Pierre André Zitt (université Paris Est — Marne–la–Vallée),
- rare event simulation in switching diffusions has been studied by François Le Gland in [23] and [8], in collaboration with Pascal Lezaud (DGAC/DSNA and ENAC), Jaroslav Krystul (Twente University) and Anindya Goswami (IISER Pune),
- Markov–switching autoregressive models for wind time series have been studied by Valérie Monbet in [3], in collaboration with Pierre Ailliot (université de Bretagne occidentale).

# 3 Knowledge dissemination

## 3.1 Publications

| | 2010 | 2011 | 2012 | 2013 | 2014 | **Total** |
|---|---|---|---|---|---|---|
| PhD Thesis | 2 | | 2 | 1 | | **5** |
| HDR | | 1 | | | | **1** |
| Journal | 3 | 6 | 6 | 6 | 6 | **27** |
| Book chapter | | 1 | 1 | | | **2** |
| Conference proceedings | 2 | 3 | 6 | 4 | | **15** |
| Book (written) | | | 1 | | | **1** |

Indicate the major journals in the field and, for each, indicate the number of papers coauthored by members of the project-team that have been accepted during the evaluation period.

1. Annales de l'IHP, Probabilités et Statistiques: 2
2. Statistics and Computing: 1
3. The Annals of Applied Probability: 1
4. Stochastic Processes and their Applications: 1
5. Journal of Machine Learning Research: 2
6. IEEE Transactions on Information Theory: 1

Indicate the major conferences in the field and, for each, indicate the number of papers coauthored by members of the project-team that have been accepted during the evaluation period.

1. International Conference on Information Fusion: 5
2. Winter Simulation Conference: 1

## 3.2 Software

**Rare event MATLAB toolbox** This toolbox [link] is available on Matlab Central. It proposes two algorithms [11] and [20] developped by ASPI for the estimation of probability and quantile related to rare event problems, and described in Section 2.4.3. The practical implementation was done by Teddy Furon (EPI TEXMEX, INRIA Rennes — Bretagne Atlantique).

### 3.3 Valorization and technology transfert (socio–economic impact and transfer)

The application of adaptive multilevel splitting to experimental assessment of reliability for watermarking and fingerprinting schemes [8] has been protected by several patents, owned jointly by INRIA and université Rennes 2: a French patent FR2935058 [link] submitted in August 2008 and published in February 2010, an international patent WO2010018313 [link], a European patent EP2318961 [link] and a US patent US2012060061 [link] all submitted in July 2009 and published in February 2010, in May 2011 and in March 2012, respectively. The inventors are Frédéric Cérou, Teddy Furon (EPI TEXMEX, INRIA Rennes — Bretagne Atlantique) and Arnaud Guyader.

### 3.4 Teaching

As an associate professor, Arnaud Guyader (yearly total, 192 hours) is teaching analysis, statistics and probability at université Rennes 2, and he is also a member of the committee of *oraux blancs d'agrégation de mathématiques* for ENS de Rennes (previously known as *antenne de Bretagne de l'ENS Cachan*).

François Le Gland (yearly total, 59 hours) gives a course on Bayesian filtering and particle approximation [link] at ENSTA (École Nationale Supérieure de Techniques Avancées, 21 hours), a course on hidden Markov models [link] at Télécom Bretagne (6 hours), a course on linear and nonlinear filtering [link] at ENSAI (École Nationale de la Statistique et de l'Analyse de l'Information, 12 hours) and a course on Kalman filtering and hidden Markov models [link] at université de Rennes 1, master in electronics and telecommunications (ED 359), track in signal and image processing, embedded systems, and control (20 hours). He has also organized from November 2010 to February 2011, a working group at ONERA Palaiseau on particle methods and their applications to Bayesian filtering and to rare event simulation, and he has organized in February 2012, a thematic school on particle filtering [link], proposed as a complementary scientific training to PhD students of école doctorale MATISSE (ED 359).

As an associate professor, Florent Malrieu (yearly total, 192 hours) has been teaching analysis, statistics and probability at université de Rennes 1. He has also given in February 2013, a doctoral course on piecewise deterministic Markov processes (PDMP) proposed as a complementary scientific training to PhD students of école doctorale MATISSE (ED 359).

As a professor, Valérie Monbet (yearly total, 192 hours) gives several courses on data analysis, on time series, and on mathematical statistics, all at université de Rennes 1 within the master on statistics and econometrics (ED 505). She is also the director of the master on statistics and econometry at université de Rennes 1, école doctorale SHOS (ED 505).

### 3.5 Visibility

Arnaud Guyader and Frédéric Cérou have coorganized the workshop on *Computation of Transition Trajectories and Rare Events in Non-Equilibrium Systems*, held in ENS de Lyon in June 2012 [link]. Arnaud Guyader has organized the session on rare events simulation at the *Journées MAS de la SMAI*, held in Clermont-Ferrand in August 2012. He has also co-organized the 2012 edition of the *Journées de Statistiques Rennaises*, held in Rennes in October 2012 [link].

François Le Gland has been a member of the scientific and organizing committees for the international conference on *Ensemble Methods in Geophysical Sciences*, held in

Toulouse in November 2012 [link], an event organized within the ANR project PREVASSEMBLE. He is a member of the organizing committee of the *46èmes Journées de Statistique*, to be held in Rennes in June 2014 [link].

Florent Malrieu has co-organized the 2012 edition of the *Journées de Probabilités*, held in Roscoff in June 2012 [link]. He has coordinated the spring semester of the Labex Henri Lebesgue on *Perspectives in Analysis and Probability*, from April to September 2013 [link], and he has co–organized the workshop on *Piecewise Deterministic Markov Processes* in May 2013 [link]. He is also the coordinator of the ANR project PIECE (programme Jeunes Chercheuses et Jeunes Chercheurs) [link]. He is the coordinator, with Tony Lelièvre (CERMICS and ENPC, Marne-la-Vallée and EPI MATHERIALS, INRIA Paris — Rocquencourt), of the SMAI meeting series *EDP/Probabilités* [link]. He his a member of the scientific and organizing committee of the conference in honour of the 60th birthday of Dominique Bakry, to be held in Toulouse in December 2014.

Valérie Monbet has co-organized the first international workshop on *Stochastic Weather Generators*, held in Roscoff in May 2012 [link], an event attended by most of the major teams working on WGs. She has co-organized with Pierre Ailliot (université de Bretagne occidentale, Brest) the two workshops *Space-Time Data Analysis in Oceanography and Meteorology* held in Berder in May 2013 [link] and in Landéda in November 2013 [link].

# 4 Funding

### Funding external to INRIA

| (K€ per fiscal year) | 2010 | 2011 | 2012 | 2013 | 2014 |
|---|---|---|---|---|---|
| National initiatives | | | | | |
| ANR FIL | 28 | | | | |
| ANR PREVASSEMBLE | 5 | 5 | 5 | | |
| ANR NEBBIANO | 12 | | | | |
| European projects | | | | | |
| FP6 iFLY | 14 | | 8.5 | | |
| Industrial contracts | | | | | |
| Thalès Communications | 50 | | | | |
| DGA/Techniques navales | 10 | 10 | | | |
| DGA/Techniques navales (DUCATI) | | | | 39 | 39 |
| Canon Research | | | | | 2.5 |
| Total | 119 | 15 | 13.5 | 39 | 41.5 |

| (number per academic year) | 10/11 | 11/12 | 12/13 | 13/14 |
|---|---|---|---|---|
| Scholarships | | | | |
| ONERA/DCPS PhD | 1 | | | |
| ONERA/DTIM PhD | 1 | 1 | | |
| ONERA/DEMR PhD | 1 | 1 | 1 | |
| DGA+ONERA/DCPS PhD | | 1 | 1 | 1 |
| Total | 3 | 3 | 2 | 1 |

### INRIA competitive funding

No competetive funding by INRIA has been obtained during the evaluation period. However, several secondment positions have been obtained and are gratefully acknowledged.

| (K€ per fiscal year) | 2010 | 2011 | 2012 | 2013 | 2014 |
|---|---|---|---|---|---|
| INRIA research initiatives | | | | | |
| ARC | | | | | |
| AE | | | | | |
| Associated teams | | | | | |
| EA | | | | | |
| Scholarships | | | | | |
| Internship | | | | | |
| PhD | | | | | |
| Post Doc | | | | | |
| Technological development | | | | | |
| AI | | | | | |
| ADT | | | | | |
| ODL | | | | | |
| Total | | | | | |

| (number per academic year) | 10/11 | 11/12 | 12/13 | 13/14 |
|---|---|---|---|---|
| Other INRIA competitive funding | | | | |
| Délégation F. Malrieu | 1 | | | |
| Délégation A. Guyader | | | 1 | 0.5 |
| Total | 1 | | 1 | 0.5 |

ARC (action de recherche coopérative): INRIA cooperative research initiative
AE (action d'envergure nationale): large–scale nationwide initiative action
EA (équipe associée): associated team
PhD (CORDI-S, doctorat INRIA sur subvention): INRIA doctoral research contract
Post–Doc (post-doc INRIA sur subvention): INRIA postdoctoral research contract
AI: junior engineer supported by INRIA
ADT (action de développement technologique): technological development action
ODL (opération de développement logiciel): software development operation
Délégation: secondment position

## 4.1 National initiatives

**FIL** (ALLOC 2856, from January 2008 to June 2011). This ANR project within the Télécommunications programme has been coordinated by Thalès Alenia Space. Academic partners are LAAS (laboratoire d'architecture et d'analyse des systèmes), TeSA consortium including ENAC (école nationale de l'aviation civile) and ISAE (institut des sciences de l'aéronautique et de l'espace). Industrial partners were Microtec and Silicom.

The overall objective was to study and demonstrate information fusion algorithms for localization of pedestrian users in an indoor environment, where GPS solution cannot be used. The design combined

- a pedestrian dead–reckoning (PDR) unit, providing noisy estimates of the linear displacement, angular turn, and possibly of the level change through an additional pression sensor,

- range and/or proximity measurements provided by beacons at fixed and known locations, and possibly indirect distance measurements to access points, through a measure of the power signal attenuation,
- constraints provided by an indoor map of the building (map-matching),
- collaborative localization when two (or more) users meet and exchange their respective position estimates.

Particle methods have been proposed in [1] and [24] as the basic information fusion algorithm for the centralized server–based version, including adaption of the sample size using KLD–sampling [16], or collaboration between users [17].

**PREVASSEMBLE** (ALLOC 3767, from January 2009 to December 2011). This ANR project within the COSINUS (conception and simulation) programme, has been coordinated by École Normale Supérieure, Paris. The other (industrial) partner was Météo–France. This has been a collaboration with Étienne Mémin (EPI FLUMINANCE, INRIA Rennes — Bretagne Atlantique) and with Anne Cuzol and Valérie Monbet (now a professor at université de Rennes 1, and a member of ASPI) from université de Bretagne Sud in Vannes.

The contribution of ASPI to this project has been to compare in [24] sequential data assimilation methods such as the ensemble Kalman filter (EnKF) with particle filters. This comparison has been made on the basis of asymptotic variances. Even though the project is now terminated, research activity still continues on these issues.

**PIECE** (from January 2013 to December 2016) [link]. This ANR project within the JCJC (jeunes chercheuses et jeunes chercheurs) programme, is coordinated by Florent Malrieu (a member of ASPI until September 2013, now a professor at université François Rabelais, Tours).

Piecewise deterministic Markov processes (PDMP) are non–diffusive stochastic processes which naturally appear in many areas of applications as communication networks, neuron activities, biological populations or reliability of complex systems. Their mathematical study has been intensively carried out in the past two decades but many challenging problems remain completely open. This project aims at federating a group of experts with different backgrounds (probability, statistics, analysis, partial differential equations) to study models coming from different fields (biology, physics, computer science, etc.). The main lines of the project relate to estimation, simulation and asymptotic behaviors (long time, large populations, multi–scale problems) in the various contexts of application.

## 4.2 European projects

The project presented below ended shortly after the beginning of the evaluation period. It should be mentionned that (together with the previous FP5 project HYBRIDGE) it has been the starting point for the growing interest of ASPI in multilevel splitting for rare event simulation.

**iFLY** (ALLOC 2399, from May 2007 to April 2011) [link]. This FP6 project within the Aeronautics and Space programme has been coordinated by National Aerospace Laboratory (NLR) (The Netherlands), and can be seen as a follow–up of the previous FP5 project HYBRIDGE within the IST programme. Eighteen academic and industrial partners were involved, and actual collaboration occurred with Politecnico di

Milano (Italy), University of Twente (The Netherlands), ETH Zürich (Switzerland), DGAC/DSNA and ENAC (France), and Honeywell (Czech Republic).

The objective of iFLY was to develop both an advanced airborne self separation design and a highly automated air traffic management (ATM) design for en–route traffic, which takes advantage of autonomous aircraft operation capabilities and which is aimed to manage a three to six times increase in current en–route traffic levels. The project combined expertise in air transport human factors, safety and economics with analytical and Monte Carlo simulation methodologies. The contribution of ASPI to this project concerned the work package on accident risk assessment methods and their implementation using conditional Monte Carlo methods, especially for large scale stochastic hybrid systems: designing and studying variants [2] suited for hybrid state space (resampling per mode, marginalization) have been investigated in [23] and [8]. Even though the project is now terminated, research activity still continues on these issues.

### 4.3 Industrial contracts

The contract presented below ended shortly after the beginning of the evaluation period. It should be mentionned that (together with the follow–up contract DUCATI) it has been the starting point for the growing interest of ASPI in multilevel splitting for global optimization.

**Optimization of sensors position and activation** (ALLOC 4233, from April 2009 to March 2011). This project has been supported by CTSN (Centre Technique des Systèmes Navals), a DGA (Délégation Générale à l'Armement) entity. This collaboration with Sébastien Paris, from université Paul Cézanne, is related with the supervision of the PhD thesis of Mathieu Chouchane, and was initiated by Jean–Pierre Le Cadre.

The objective of this project was to optimize the position and activation times of a few sensors deployed by a platform over a search zone, so as to maximize the probability of detecting a moving target. The difficulty here is that the target can detect an activated sensor before it is detected itself, and it can then modify its own trajectory to escape from the sensor. This makes the optimization problem a spatio–temporal problem. Indeed, if a sensor would be activated at all times, then it would be detected by a target before it could itself detect the target. Following the multilevel splitting point of view, the idea has been to replace the problem of maximizing a cost function (the probability of detection) over the possible configurations (admissible position and activation times) by the apparently simpler problem of sampling a population from a reference probability distribution conditionned on the cost function exceeding a threshold. Asymptotically as the threshold grows, the population concentrates on the set of global maxima of the cost function [16]. Multi–objective optimization problems have also been considered by Mathieu Chouchane in his PhD thesis [11].

**DUCATI** (ALLOC 7326, from April 2013 to December 2016) This project is supported by DGA/Techniques Navales (previously known as CTSN), a DGA (Délégation Générale à l'Armement) entity. This collaboration with Sébastien Paris, from université Paul Cézanne, and with Christian Musso (ONERA Palaiseau), is related with the supervision of the PhD thesis of Yannick Kenné, and is a continuation of the project mentionned above.

The activity in the beginning of this new project is to study different ways to merge two different solutions to the optimization problem: a fast, though suboptimal, solution developped by ONERA in which sensors are deployed where and when the probability of presence of a target is considered high enough, and the optimal population–based solution developped by LSIS and by INRIA.

### 4.4 INRIA research initiatives

A few proposals have been submitted to the INRIA research initiatives programmes during the evaluation period, and none has been accepted: one submission in 2011 to the ARC programme (REACT, on rare events applied to chemistry and turbulence), and one submission in 2013 to the Associated Team programme (REPAIR, with Los Alamos National Laboratory).

## 5 Objectives for the next four years

Here is the proposition for the redefinition of the research objectives of ASPI for the next four years. This proposition is based on self assessment of the current objectives and discussions inside the team.

### 5.1 Rare events simulation

The effort put here should be continued, especially in the same two directions: molecular simulation, where a fruitful collaboration is already effective with leading experts, and global optimization, where a complex real–world application is already available.

**Molecular simulation** This is already a successful activity, which ASPI is willing to continue and further develop, in collaboration with Tony Lelièvre and his group (CERMICS, ENPC and EPI MATHERIALS, INRIA Paris — Rocquencourt), especially Mathias Rousset (CR INRIA) and Charles–Édouard Bréhier (post–doc). To give just an example [5], reactive trajectories have been recently studied in the relatively simple model of a stochastic Allen–Cahn PDE, modelling the interface between two chemical species and interpreted as the limit, under an appropriate scaling, of a large number of interacting overdamped Langevin equations. This model generalizes the finite–dimensional model considered in [14], and requires further investigation.

**Global optimization** During the current evaluation period, multilevel splitting has been successfully applied to the optimization of sensors position and activation, with the support of DGA/Techniques Navales. Extensive numerical experience, including for multi–objective optimization problems, has been obtained and the objective for the next evaluation period is to study and analyze the performance of these algorithms and their large–sample asymptotics.

**Methodology** There is still a lot to be done here, especially about the convergence issues, including a central limit theorem (CLT) and explicit expressions for the asymptotic variance, for the adaptive version (where the intermediate levels are learned on–line) of the multilevel splitting algorithm. In the static case, the problem has been solved [11] only for an idealized version of the algorithm. In the dynamic case, partial answers [10] are available only for a one–dimensional situation with a random finite time.

Rare event simulation in stochastic hybrid systems requires further investigation. To be specific, an algorithm implementing marginalization has been presented in [8], and its thorough analysis is still to be done.

### 5.2 Modelling and estimation (data assimilation) in oceano–meteo

With the modelling expertise in oceanography (and in meteorology) brought by Valérie Monbet, some interesting issues could be addressed in modelling, design of stochastic weather generators, large sample asymptotics, etc.

**Probabilistic modelling** In full generality, stochastic weather generators simulate sequences of daily weather and climate consistent with specific aspects of climate variability and change. The simulated sequences of meteorological variables (rainfall, wind, temperature, etc.) are typically used as input into complex environmental and ecosystem models. They have potentially a wide range of applications in hydrology, agriculture and environmental management. In principle, they could also be used as input to black–box model, e.g. to evaluate the response of a structure to extreme or severe weather conditions. A practical objective here is to design stochastic state space models for wind fields and sea state, including Markov–switching modes to account for the weather type, so as to generate realistic wind conditions and sea state.

**Large–sample asymptotics in sequential data assimilation** Convergence issues, including a central limit theorem (CLT) and explicit expression for the asymptotic variance have been studied for the ensemble Kalman filter (EnKF), and a comparison has been made with two popular particle filters on the basis of the respective expression of their asymptotic variance (at least in the linear Gaussian case). Similar work remains to be done for the weighted ensemble Kalman filter (WEnKF), an unbiased version of EnKF proposed in [26].

## 6 Bibliography of the project-team

### Books and monographs

[1] Pierre-André Cornillon, Arnaud Guyader, François Husson, Nicolas Jégou, Julie Josse, Maela Kloareg, Eric Matzner-Løber, and Laurent Rouvière. *R for statistics.* Chapman and Hall/CRC Press, 2012.

### PhD and habilitation theses

[1] Paul Bui Quang. *Approximation particulaire et méthode de Laplace pour le filtrage bayésien.* Thèse, Université de Rennes 1, July 2013.

[2] Francis Celeste. *Cartes incertaines et planification optimale pour la localisation d'un engin autonome.* Thèse, Université de Rennes 1, February 2010.

[3] Arnaud Guyader. *Contributions à l'estimation non paramétrique et à la simulation d'évènements rares.* Habilitation à diriger des recherches, Université Rennes 2, December 2011.

[4] Adrien Ickowicz. *Méthodes d'estimation statistique pour le suivi de cibles à l'aide d'un réseau de capteurs.* Thèse, Université de Rennes 1, May 2010.

[5] Nicolas Jégou. *Régression isotonique itérée.* Thèse, Université Rennes 2, Rennes, November 2012.

[6] Rudy Pastel. *Estimation de probabilités d'évènements rares et de quantiles extrêmes. Applications dans le domaine aérospatial.* Thèse, Université de Rennes 1, February 2012.

### Articles in referred journals and in contributed volumes

[1] Pierre Ailliot, Anastassia Baxevani, Anne Cuzol, Valérie Monbet, and Nicolas Raillard. Space–time models for moving fields. application to significant wave height. *Environmetrics*, 22(3):354–369, May 2011.

[2] Pierre Ailliot, Christophe Maisondieu, and Valérie Monbet. Dynamical partitioning of directional ocean wave spectra. *Probabilistic Engineering Mechanics*, 33:95–102, July 2013.

[3] Pierre Ailliot and Valérie Monbet. Markov switching autoregressive models for wind time series. *Environmental Modelling & Software*, 30:92–101, April 2012.

[4] Jean-Baptiste Bardet, Alejandra Christen, Arnaud Guillin, Florent Malrieu, and Pierre-André Zitt. Total variation estimates for the TCP process. *Electronic Journal of Probability*, 18(10):1–21, 2013.

[5] Michel Benaïm, Stéphane Le Borgne, Florent Malrieu, and Pierre-André Zitt. Quantitative ergodicity for some switched dynamical systems. *Electronic Communications in Probability*, 17(56):1–14, 2012.

[6] Michel Benaïm, Stéphane Le Borgne, Florent Malrieu, and Pierre-André Zitt. On the stability of planar randomly switched systems. *The Annals of Applied Probability*, 24(1):292–311, February 2014.

[7] Gérard Biau, Frédéric Cérou, and Arnaud Guyader. New insights into approximate Bayesian computation. *Annales de l'Institut Henri Poincaré, Probabilités et Statistiques.* accepted for publication.

[8] Gérard Biau, Frédéric Cérou, and Arnaud Guyader. On the rate of convergence of the bagged nearest neighbor estimate. *Journal of Machine Learning Research*, 11:687–712, 2010.

[9] Gérard Biau, Frédéric Cérou, and Arnaud Guyader. On the rate of convergence of the functional $k$–nearest neighbor estimates. *IEEE Transactions on Information Theory*, IT-56(4):2034–2040, 2010.

[10] Virgile Caron, Arnaud Guyader, Miguel Munoz Zuniga, and Bruno Tuffin. Some recent results in rare event estimation. *ESAIM Proceedings*, 44:239–259, 2014.

[11] Frédéric Cérou, Pierre Del Moral, Teddy Furon, and Arnaud Guyader. Sequential Monte Carlo for rare event estimation. *Statistics and Computing*, 22(3):795–908, 2012.

[12] Frédéric Cérou, Pierre Del Moral, and Arnaud Guyader. A nonasymptotic theorem for unnormalized Feynman-Kac particle models. *Annales de l'Institut Henri Poincaré, Probabilités et Statistiques*, 47(3):629–649, 2011.

[13] Frédéric Cérou, Arnaud Guyader, Tony Lelièvre, and Florent Malrieu. On the length of one-dimensional reactive paths. *ALEA. Latin American Journal of Probability and Mathematical Statistics*, 10(1):359–389, 2013.

[14] Frédéric Cérou, Arnaud Guyader, Tony Lelièvre, and David Pommier. A multiple replica approach to simulate reactive trajectories. *Journal of Chemical Physics*, 134(5):054108 (16 pages), 2011.

[15] Frédéric Cérou, Arnaud Guyader, Reuven Rubinstein, and Radislav Vaisman. On the use of smoothing to improve the performance of the splitting method. *Stochastic Models*, 27(4):629–650, 2011.

[16] Mathieu Chouchane, Sébastien Paris, François Le Gland, and Mustapha Ouladsine. Splitting method for spatio-temporal sensors deployment in underwater systems. In Hao Jin-Kao; Middendorf Martin, editor, *Evolutionary Computation in Combinatorial Optimization. 12th European Conference, EvoCOP 2012, Málaga, Spain, April 11-13, 2012, Proceedings*, Lecture Notes in Computer Science, Vol. 7245, pages 243–254. Springer, 2012.

[17] Joaquin Fontbona, Hélène Guérin, and Florent Malrieu. Quantitative estimates for the long–time behavior of an ergodic variant of the telegraph process. *Advances in Applied Probability*, 44(4):977–994, December 2012.

[18] Arnaud Guyader and Nick Hengartner. On the mutual nearest neighbors estimate in regression. *Journal of Machine Learning Research*, 14:2361–2376, 2013.

[19] Arnaud Guyader, Nick Hengartner, Nicolas Jégou, and Eric Matzner-Lober. Iterative isotonic regression. *ESAIM: Probability and Statistics.* accepted for publication.

[20] Arnaud Guyader, Nicolas W. Hengartner, and Eric Matzner-Løber. Simulation and estimation of extreme quantiles and extreme probabilities. *Applied Mathematics and Optimization*, 64(2):171–196, 2011.

[21] Arnaud Guyader, Nicolas Jégou, Alexander B. Németh, and Sándor Z. Németh. A geometrical approach to iterative isotone regression. *Applied Mathematics and Computation*, 227:359–369, 2014.

[22] Damien Jacquemart and Jérôme Morio. Conflict probability estimation between aircraft with dynamic importance splitting. *Safety Science*, 51(1):94–100, 2013.

[23] Jaroslav Krystul, François Le Gland, and Pascal Lezaud. Sampling per mode for rare event simulation in switching diffusions. *Stochastic Processes and their Applications*, 122(7):2639–2667, 2012.

[24] François Le Gland, Valérie Monbet, and Vu-Duc Tran. Large sample asymptotics for the ensemble Kalman filter. In Boris Rozovskii Dan Crisan, editor, *The Oxford Handbook of Nonlinear Filtering*, pages 598–631. Oxford University Press, 2011.

[25] Jérôme Morio and Rudy Pastel. Plug-in estimation of $d$-dimensional density minimum volume set of a rare event in a complex system. *Proceedings of the Institution of Mechanical Engineers, Part O: Journal of Risk and Reliability*, 226(3):337–345, 2012.

[26] Jérôme Morio, Rudy Pastel, and François Le Gland. An overview of importance splitting for rare event simulation. *European Journal of Physics*, 31(5):1295–1303, 2010.

[27] Jérôme Morio, Rudy Pastel, and François Le Gland. Estimation de probabilités et de quantiles rares pour la caractérisation d'une zone de retombée d'un engin. *Journal de la Société Française de Statistique*, 152(4):1–29, 2011.

[28] Jérôme Morio, Rudy Pastel, and François Le Gland. Missile target accuracy estimation with importance splitting. *Aerospace Science and Technology*, 25:40–44, 2013.

[29] Rudy Pastel, Jérôme Morio, and François Le Gland. Improvement of satellite conflict prediction reliability through use of the adaptive splitting technique. *Proceedings of the Institution of Mechanical Engineers, Part G: Journal of Aerospace Engineering*, 228(1):77–85, January 2014.

### Publications in conferences and workshops

[1] Mélanie Bocquel, Alexandre Lepoutre, Olivier Rabaste, and François Le Gland. Optimisation d'un filtre particulaire en contexte track-before-detect. In *Actes du 23ème Colloque sur le Traitement du Signal et des Images, Bordeaux 2011*, Bordeaux, France, 2011.

[2] Paul Bui Quang, Christian Musso, and François Le Gland. An insight into the issue of dimensionality in particle filtering. In *Proceedings of the 13th International Conference on Information Fusion, Edinburgh 2010*, Edinburgh, United Kingdom, 2010. ISIF. Paper We2.1.2.

[3] Paul Bui Quang, Christian Musso, and François Le Gland. Multidimensional Laplace formulas for nonlinear Bayesian estimation. In *Proceedings of the 20th European Signal Processing Conference, Bucharest 2012*, Bucharest, Romania, 2012.

[4] Mathieu Chouchane, Sébastien Paris, François Le Gland, Christian Musso, and Dinh-Tuan Pham. On the probability distribution of a moving target. Asymptotic and non-asymptotic results. In *Proceedings of the 14th International Conference on Information Fusion, Chicago 2011*, pages 99–101, Chicago, United States, 2011. ISIF.

[5] Éric Dutisseuil, Jean-Marc Tanguy, Adrian Voicila, Rémi Laube, François Bore, Honoré Takeugming, Florent De Dinechin, Frédéric Cérou, and Gabriel Charlet. 34 Gb/s PDM-QPSK coherent receiver using SiGe ADCs and a single FPGA for digital signal processing. In *Proceedings of the 2012 Optical Fiber Communication Conference, Los Angeles 2012*, page OM3H.7, United States, March 2012.

[6] Nordine El Baraka and François Le Gland. Bayesian terrain-aided inertial navigation using an airborne laser scanner. In *Proceedings of the 2010 ION International Technical Meeting, San Diego 2010*, pages 111–116, San Diego, United States, 2010. ION.

[7] Teddy Furon, Arnaud Guyader, and Frédéric Cérou. Decoding fingerprinting using the Markov chain Monte Carlo method. In *Proceedings of the 2012 IEEE International Workshop on Information Forensics and Security, Tenerife 2012*, Tenerife, Spain, December 2012. IEEE.

[8] Anindya Goswami and François Le Gland. Marginalization for rare event simulation in switching diffusions. In *Proceedings of the International Conference on Mathematical Techniques in Engineering Applications, Dehradun 2013*, Dehradun, India, 2013.

[9] Damien Jacquemart, François Le Gland, and Jérôme Morio. A combined importance splitting and sampling algorithm for rare event estimation. In Raghu Pasupathy, Seong-He Kim, Andreas Tolk, Ray R. Hill, and Mike E. Kuhl, editors, *Proceedings of the 2013 Winter Simulation Conference, Washington 2013*, pages 1035–1046, Washington DC, United States, 2013.

[10] Alexandre Lepoutre, Olivier Rabaste, and François Le Gland. A particle filter for target arrival detection and tracking in track-before-detect. In *Proceedings of the 2012 Workshop on Sensor Data Fusion: Trends, Solutions, Applications, Bonn 2012*, pages 13–18, Bonn, Germany, 2012.

[11] Alexandre Lepoutre, Olivier Rabaste, and François Le Gland. Optimized instrumental density for particle filter in track-before-detect. In *Proceedings of the 9th IET Data Fusion Target Tracking Conference, London 2012*, London, United Kingdom, 2012.

[12] Alexandre Lepoutre, Olivier Rabaste, and François Le Gland. Exploiting amplitude spatial coherence for multi-target particle filter in track-before-detect. In *Proceedings of the 16th International Conference on Information Fusion, Istanbul 2013*, pages 319–326, Istanbul, Turkey, 2013.

[13] Alexandre Lepoutre, Olivier Rabaste, and François Le Gland. Filtres particulaires en contexte track-before-detect en présence de fluctuations d'amplitude de type Swerling 1 et 3. In *Actes du 24ème Colloque sur le Traitement du Signal et des Images, Brest 2013*, Brest, France, 2013. GRETSI.

[14] Christian Musso, Paul Bui Quang, and François Le Gland. Introducing the Laplace approximation in particle filtering. In *Proceedings of the 14th International Conference on Information Fusion, Chicago 2011*, pages 290–297, Chicago, United States, 2011.

[15] Olivier Rabaste, Christian Riché, and Alexandre Lepoutre. Long-time coherent integration for low SNR target via particle filter in track-before-detect. In *Proceedings of the 15th International Conference on Information Fusion, Singapore 2012*, pages 127–134, Singapore, Singapore, 2012. ISIF.

# A Additional references

### Additional references cited in this document

[1] Pierre Blanchart, Liyun He, and François Le Gland. Information fusion for indoor localization. In *Proceedings of the 12th International Conference on Information Fusion, Seattle 2009*, pages 2083–2090. ISIF, July 2009.

[2] Henk A. P. Blom, G. J. (Bert) Bakker, and Jaroslav Krystul. Rare event estimation for a large scale stochastic hybrid system with air traffic application. In Gerardo Rubino and Bruno Tuffin, editors, *Monte Carlo Methods for Rare Event Analysis*, chapter 9, pages 193–214. John Wiley & Sons, Chichester, 2009.

[3] Zdravko I. Botev and Dirk P. Kroese. An efficient algorithm for rare–event probability estimation, combinatorial optimization, and counting. *Methodology and Computing in Applied Probability*, 10(4):471–505, December 2008.

[4] Zdravko I. Botev and Dirk P. Kroese. The generalized cross–entropy method, with applications to probability density estimation. *Methodology and Computing in Applied Probability*, 13(1):1–27, 2011.

[5] Charles-Edouard Bréhier, Maxime Gazeau, Ludovic Goudenège, and Mathias Rousset. Analysis and simulation of rare events for SPDE. arXiv:1401.1380.

[6] Frédéric Cérou, Pierre Del Moral, François Le Gland, and Pascal Lezaud. Limit theorems for the multilevel splitting algorithm in the simulation of rare events. In Michael E. Kuhl, Natalie M. Steiger, Frank B. Armstrong, and Jeffrey A. Joines, editors, *Proceedings of the 2005 Winter Simulation Conference, Orlando 2005*, pages 682–691, December 2005.

[7] Frédéric Cérou, Pierre Del Moral, François Le Gland, and Pascal Lezaud. Genetic genealogical models in rare event analysis. *ALEA, Latin American Journal of Probability and Mathematical Statistics*, 1:181–203, 2006. Paper 01–08.

[8] Frédéric Cérou, Teddy Furon, and Arnaud Guyader. Experimental assessment of the reliability for watermarking and fingerprinting schemes. *EURASIP Journal on Information Security*, 2008. Article ID 414962.

[9] Frédéric Cérou and Arnaud Guyader. Nearest neighbor classification in infinite dimension. *ESAIM : Probability and Statistics*, 10:340–355, 2006.

[10] Frédéric Cérou and Arnaud Guyader. Adaptive multilevel splitting for rare event analysis. *Stochastic Analysis and Applications*, 25(2):417–443, March 2007.

[11] Mathieu Chouchane. *Optimisation spatio–temporelle d'efforts de recherche pour cibles manoeuvrantes et intelligentes.* Thèse de Doctorat, Université d'Aix–Marseille, Marseille, October 2013.

[12] Pieter-Tjerk De Boer, Dirk P. Kroese, Shie Mannor, and Reuven Y. Rubinstein. A tutorial on the cross–entropy method. *Annals of Operations Research*, 134(1 (Special issue on the Cross-Entropy Method for Combinatorial Optimization, Rare Event Simulation and Neural Computation)):19–67, February 2005.

[13] Pierre Del Moral. *Feynman–Kac formulae. Genealogical and interacting particle systems with applications.* Probability and its Applications. Springer–Verlag, New York, 2004.

[14] Pierre Del Moral. *Mean field simulation for Monte Carlo integration*, volume 126 of *Monographs on Statistics and Applied Probability.* Chapman & Hall / CRC Press, London, 2013.

[15] Geir Evensen. *Data assimilation. The ensemble Kalman filter.* Springer–Verlag, Berlin, 2006.

[16] Dieter Fox. Adapting the sample size in particle filters through KLD–sampling. *International Journal of Robotics Research*, 22(12):985–1003, December 2003.

[17] Dieter Fox, Wolfram Burgard, Hannes Kruppa, and Sebastian Thrun. A probabilistic approach to collaborative multi–robot localization. *Autonomous Robots*, 8(3 (Special issue on Heterogeneous Multi–Robot Systems)):325–344, June 2000.

[18] Paul Glasserman, Philip Heidelberger, Perwez Shahabuddin, and Tim Zajic. A large deviations perspective on the efficiency of multilevel splitting. *IEEE Transactions on Automatic Control*, AC–43(12):1666–1679, December 1998.

[19] Paul Glasserman, Philip Heidelberger, Perwez Shahabuddin, and Tim Zajic. Multilevel splitting for estimating rare event probabilities. *Operations Research*, 47(4):585–600, July–August 1999.

[20] Neil J. Gordon, David J. Salmond, and Adrian F. M. Smith. Novel approach to nonlinear / non–Gaussian Bayesian state estimation. *IEE Proceedings, Part F*, 140(2):107–113, April 1993.

[21] Fredrik Gustafsson, Fredrik Gunnarsson, Niclas Bergman, Urban Forssell, Jonas Jansson, Rickard Karlsson, and Per-Johan Nordlund. Particle filters for positioning, navigation, and tracking. *IEEE Transactions on Signal Processing*, SP–50(2 (Special issue on Monte Carlo Methods for Statistical Signal Processing)):425–437, February 2002.

[22] Dirk P. Kroese, Sergey Porotsky, and Reuven Y. Rubinstein. The cross–entropy method for continuous multiextremal optimization. *Methodology and Computing in Applied Probability*, 8(3):383–407, September 2006.

[23] Jaroslav Krystul. *Modelling of stochastic hybrid systems with applications to accident risk assessment*. Ph.D Thesis, Department of Applied Mathematics, University of Twente, Enschede, September 2006.

[24] Damien Kubrak, François Le Gland, Liyun He, and Yann Oster. Multi–sensor fusion for localization. Concept and simulation results. In *Proceedings of the 2009 ION Conference on Global Navigation Satellite Systems, Savannah 2009*, pages 767–777. ION, September 2009.

[25] Pierre L'Écuyer, Valérie Demers, and Bruno Tuffin. Rare events, splitting, and quasi–Monte Carlo. *ACM Transactions on Modeling and Computer Simulation*, 17(2 (Special issue honoring Perwez Shahabuddin)), April 2007. Article 9.

[26] Nicolas Papadakis, Étienne Mémin, Anne Cuzol, and Nicolas Gengembre. Data assimilation with the weighted ensemble Kalman filter. *Tellus A*, 62(5):673–697, October 2010.

[27] Branko Ristić, Sanjeev Arulampalam, and Neil J. Gordon. *Beyond the Kalman filter: Particle filters for tracking applications*. Artech House, Boston, 2004.

[28] Reuven Y. Rubinstein. The Gibbs cloner for combinatorial optimization, counting and sampling. *Methodology and Computing in Applied Probability*, 11(4):491–549, December 2009.

[29] Reuven Y. Rubinstein and Dirk P. Kroese. *The cross–entropy method. A unified approach to combinatorial optimization, Monte Carlo simulation and machine learning*. Information Science and Statistics. Springer–Verlag, New York, 2004.

[30] Charles J. Stone. Consistent nonparametric regression (with discussion). *The Annals of Statistics*, 5(4):595–645, July 1977.

[31] Alexander G. Tartakovsky and Venugopal V. Veeravalli. General asymptotic Bayesian theory of quickest change detection. *Theory of Probability and its Applications*, 49(3):458–497, 2005.

[32] Sebastian Thrun, Wolfram Burgard, and Dieter Fox. *Probabilistic robotics.* Intelligent Robotics and Autonomous Agents. The MIT Press, Cambridge, MA, 2005.

# B Acronyms

ANR: Agence Nationale de la Recherche

ATER: Attaché Temporaire d'Enseignement et de Recherche

PRCE: Professeur Certifié Affecté dans l'Enseignement Supérieur

ARC: Action de Recherche Coopérative INRIA

EPI: équipe–projet INRIA (research group)

DGA: Délégation Générale à l'Armement

- DGA/Techniques Navales (Toulon)

ONERA: Office National d'Études et Recherches Aérospatiales

- DEMR: Département Électromagnétisme et Radar
- DCPS: Département Conception et Évaluation des Performance des Systèmes
- DTIM: Département Traitement de l'Information et Modélisation

CEA: Commissariat à l'Énergie Atomique et aux Énergies Alternatives

- DAM: Direction des Applications Militaires (Bruyères–le–Chatel)
- LETI: Laboratoire d'Électronique des Technologies de l'Information (Grenoble)

IISER: Indian Institute of Science Education and Research (Pune)

CSIRO: Commonwealth Scientific and Industrial Research Organisation, Australia

- CMIS: CSIRO Mathematics, Informatics and Statistics (Sydney)

DSTO: Defence Science and Technology Organization, Australia

NLR: National Aerospace Laboratory (Amsterdam)

DGAC: Délégation Générale de l'Aviation Civile

- DSNA: Direction des Services de la Navigation Aérienne

ENAC: École Nationale de l'Aviation Civile

ENSAI: École Nationale de la Statistique et de l'Analyse de l'Information (Bruz)

- CREST: Centre de Recherche en Économie et en Statistique

ENSTA: École Nationale Supérieure de Techniques Avancées (Palaiseau)

IRMAR: Institut de Recherche Mathématique de Rennes, UMR 6625

LSIS: Laboratoire des Sciences de l'Information et des Systèmes, UMR 6168 (Marseille)

MATISSE: Mathématiques, Télécommunications, Informatique, Signal, Systèmes, Électronique (école doctorale, ED 359)

SHOS: Sciences de l'Homme, des Organisations et de la Société (école doctorale, ED 505)

ENPC: École Nationale des Ponts et Chaussées (Marne–la–Vallée)

CERMICS: Centre d'Enseignement et de Recherche en Mathématiques et Calcul Scientifique (Marne–la–Vallée)

NCAR: National Center for Atmospheric Research (Boulder)

NERSC: Nansen Environmental and Remote Sensing Center (Bergen)

LANL: Los Alamos National Laboratory

ION: Institute of Navigation

ISIF: International Society of Information Fusion