



IN PARTNERSHIP WITH:  
**CNRS**

**Université Rennes 1**

**Université Haute Bretagne  
(Rennes 2)**

Activity Report 2011

## **Project-Team ASPI**

Applications of interacting particle systems to  
statistics

IN COLLABORATION WITH: Institut de recherche mathématique de Rennes (IRMAR)

RESEARCH CENTER  
**Rennes - Bretagne-Atlantique**

THEME  
**Stochastic Methods and Models**



## Table of contents

<b>1. Members</b> .....	<b>1</b>
<b>2. Overall Objectives</b> .....	<b>1</b>
<b>3. Scientific Foundations</b> .....	<b>2</b>
3.1. Interacting Monte Carlo methods and particle approximation of Feynman–Kac distributions	2
3.2. Statistics of HMM	3
3.3. Multilevel splitting for rare event simulation	5
3.4. Nearest neighbor estimates	7
<b>4. Application Domains</b> .....	<b>8</b>
4.1. Localisation, navigation and tracking	8
4.2. Rare event simulation	8
<b>5. New Results</b> .....	<b>9</b>
5.1. Reactive trajectories for molecular dynamics	9
5.2. Learning the optimal importance distribution	9
5.3. Impact of dimension in particle filtering and the Laplace method	9
5.4. Stochastic wind–wave modelling	10
5.5. Sequential data assimilation: ensemble Kalman filter vs. particle filter	10
<b>6. Contracts and Grants with Industry</b> .....	<b>10</b>
6.1. Safety, complexity and responsibility based design and validation of highly automated air traffic management (iFLY) — FP6 Aerospace	10
6.2. Optimization of sensors location and activation — contract with DGA / Techniques navales	10
6.3. Information fusion for localisation (FIL) — ANR Télécommunications	11
6.4. Ensemble methods for prediction and data assimilation (PREVASSEMBLE) — ANR Conception et Simulation	12
6.5. Slightly–multimodal transmission and detection (STRADE) — ANR Réseaux du Futur et Services	12
<b>7. Dissemination</b> .....	<b>12</b>
7.1. Scientific animation	12
7.2. Teaching	12
7.3. PhD and Habilitation Theses	13
7.4. Participation in workshops, seminars, lectures, etc.	13
<b>8. Bibliography</b> .....	<b>14</b>



## Project-Team ASPI

**Keywords:** Monte Carlo Methods, Markovian Model, Rare Events, Particle Filtering, Tracking, Data Assimilation

### 1. Members

#### Research Scientists

Frédéric Cérou [CR INRIA]  
François Le Gland [DR INRIA, team leader]

#### Faculty Members

Arnaud Guyader [assistant professor, université de Rennes 2, HdR]  
Florent Malrieu [assistant professor, université de Rennes 1, délégation INRIA until August 31, 2011, HdR]  
Valérie Monbet [professor, université de Rennes 1, HdR]

#### PhD Students

Paul Bui–Quang [université de Rennes 1, ONERA grant]  
Damien Jacquemart [université de Rennes 1, DGA / ONERA grant, since October 1, 2011]  
Alexandre Lepoutre [université de Rennes 1, ONERA grant]  
Rudy Pastel [université de Rennes 1, ONERA grant]

#### Administrative Assistant

Fabienne Cuyollaa

### 2. Overall Objectives

#### 2.1. Overall Objectives

The scientific objectives of ASPI are the design, analysis and implementation of interacting Monte Carlo methods, also known as particle methods, with focus on

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

The whole problematic is multidisciplinary, not only because of the many scientific and engineering areas in which 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.

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. 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, wireless communications, 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 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, with applications that go far beyond filtering, in

simulation of rare events, simulation of conditioned or constrained random variables, interacting MCMC methods, molecular simulation, etc.

The main applications currently considered are geolocalisation and tracking of mobile terminals, terrain-aided navigation, data fusion for indoor localisation, optimization of sensors location and activation, risk assessment in air traffic management, protection of digital documents.

## 3. Scientific Foundations

### 3.1. Interacting Monte Carlo methods and particle approximation of Feynman–Kac distributions

Monte Carlo methods are numerical methods that are widely used in situations where (i) a stochastic (usually Markovian) model is given for some underlying process, and (ii) some quantity of interest should be evaluated, that can be expressed in terms of the expected value of a functional of the process trajectory, which includes as an important special case the probability that a given event has occurred. Numerous examples can be found, e.g. in financial engineering (pricing of options and derivative securities) [46], in performance evaluation of communication networks (probability of buffer overflow), in statistics of hidden Markov models (state estimation, evaluation of contrast and score functions), etc. Very often in practice, no analytical expression is available for the quantity of interest, but it is possible to simulate trajectories of the underlying process. The idea behind Monte Carlo methods is to generate independent trajectories of this process or of an alternate instrumental process, and to build an approximation (estimator) of the quantity of interest in terms of the weighted empirical probability distribution associated with the resulting independent sample. By the law of large numbers, the above estimator converges as the size  $N$  of the sample goes to infinity, with rate  $1/\sqrt{N}$  and the asymptotic variance can be estimated using an appropriate central limit theorem. To reduce the variance of the estimator, many variance reduction techniques have been proposed. Still, running independent Monte Carlo simulations can lead to very poor results, because trajectories are generated *blindly*, and only afterwards are the corresponding weights evaluated. Some of the weights can happen to be negligible, in which case the corresponding trajectories are not going to contribute to the estimator, i.e. computing power has been wasted.

A recent and major breakthrough, has been the introduction of interacting Monte Carlo methods, also known as sequential Monte Carlo (SMC) methods, in which a whole (possibly weighted) sample, called *system of particles*, is propagated in time, where the particles

- *explore* the state space under the effect of a *mutation* mechanism which mimics the evolution of the underlying process,
- and are *replicated* or *terminated*, under the effect of a *selection* mechanism which automatically concentrates the particles, i.e. the available computing power, into regions of interest of the state space.

In full generality, the underlying process is a discrete-time Markov chain, whose state space can be

finite, continuous, hybrid (continuous / discrete), graphical, constrained, time varying, pathwise, etc.,

the only condition being that it can easily be *simulated*.

In the special case of particle filtering, originally developed within the tracking community, the algorithms yield a numerical approximation of the optimal Bayesian filter, i.e. of the conditional probability distribution of the hidden state given the past observations, as a (possibly weighted) empirical probability distribution of the system of particles. In its simplest version, introduced in several different scientific communities under the name of *bootstrap filter* [48], *Monte Carlo filter* [53] or *condensation* (conditional density propagation) algorithm [50], and which historically has been the first algorithm to include a redistribution step, the selection mechanism is governed by the likelihood function: at each time step, a particle is more likely to survive and to replicate at the next generation if it is consistent with the current observation. The algorithms also provide as a by-product a numerical approximation of the likelihood function, and of many other contrast functions for parameter estimation in hidden Markov models, such as the prediction error or the conditional least-squares criterion.

Particle methods are currently being used in many scientific and engineering areas

positioning, navigation, and tracking [49], [43], visual tracking [50], mobile robotics [44], [67], ubiquitous computing and ambient intelligence, sensor networks, risk evaluation and simulation of rare events [47], genetics, molecular simulation [45], etc.

Other examples of the many applications of particle filtering can be found in the contributed volume [26] and in the special issue of *IEEE Transactions on Signal Processing* devoted to *Monte Carlo Methods for Statistical Signal Processing* in February 2002, where the tutorial paper [27] can be found, and in the textbook [64] devoted to applications in target tracking. Applications of sequential Monte Carlo methods to other areas, beyond signal and image processing, e.g. to genetics, can be found in [63]. A recent overview can also be found in [31].

Particle methods are very easy to implement, since it is sufficient in principle to simulate independent trajectories of the underlying process. The whole problematic is multidisciplinary, not only because of the already mentioned diversity of the scientific and engineering areas in which particle methods are used, but also because of the diversity of the scientific communities which have 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.

These algorithms can be interpreted as numerical approximation schemes for Feynman–Kac distributions, a pathwise generalization of Gibbs–Boltzmann distributions, in terms of the weighted empirical probability distribution associated with a system of particles. This abstract point of view [36], [34], has proved to be extremely fruitful in providing a very general framework to the design and analysis of numerical approximation schemes, based on systems of branching and / or interacting particles, for nonlinear dynamical systems with values in the space of probability distributions, associated with Feynman–Kac distributions. Many asymptotic results have been proved as the number  $N$  of particles (sample size) goes to infinity, using techniques coming from applied probability (interacting particle systems, empirical processes [70]), see e.g. the survey article [36] or the recent textbook [34], and references therein

convergence in  $L^p$ , convergence as empirical processes indexed by classes of functions, uniform convergence in time, see also [59], [60], central limit theorem, see also [56], propagation of chaos, large deviations principle, etc.

The objective here is to systematically study the impact of the many algorithmic variants on the convergence results.

## 3.2. Statistics of HMM

Hidden Markov models (HMM) form a special case of partially observed stochastic dynamical systems, in which the state of a Markov process (in discrete or continuous time, with finite or continuous state space) should be estimated from noisy observations. The conditional probability distribution of the hidden state given past observations is a well-known example of a normalized (nonlinear) Feynman–Kac distribution, see 3.1. These models are very flexible, because of the introduction of latent variables (non observed) which allows to model complex time dependent structures, to take constraints into account, etc. In addition, the underlying Markovian structure makes it possible to use numerical algorithms (particle filtering, Markov chain Monte Carlo methods (MCMC), etc.) which are computationally intensive but whose complexity is rather small. Hidden Markov models are widely used in various applied areas, such as speech recognition, alignment of biological sequences, tracking in complex environment, modeling and control of networks, digital communications, etc.

Beyond the recursive estimation of a hidden state from noisy observations, the problem arises of statistical inference of HMM with general state space [32], including estimation of model parameters, early monitoring and diagnosis of small changes in model parameters, etc.

**Large time asymptotics** A fruitful approach is the asymptotic study, when the observation time increases to infinity, of an extended Markov chain, whose state includes (i) the hidden state, (ii) the observation, (iii) the prediction filter (i.e. the conditional probability distribution of the hidden state given observations at all previous time instants), and possibly (iv) the derivative of the prediction filter with respect to the parameter. Indeed, it is easy to express the log-likelihood function, the conditional least-squares criterion, and many other classical contrast processes, as well as their derivatives with respect to the parameter, as additive functionals of the extended Markov chain.

The following general approach has been proposed

- first, prove an exponential stability property (i.e. an exponential forgetting property of the initial condition) of the prediction filter and its derivative, for a misspecified model,
- from this, deduce a geometric ergodicity property and the existence of a unique invariant probability distribution for the extended Markov chain, hence a law of large numbers and a central limit theorem for a large class of contrast processes and their derivatives, and a local asymptotic normality property,
- finally, obtain the consistency (i.e. the convergence to the set of minima of the associated contrast function), and the asymptotic normality of a large class of minimum contrast estimators.

This programme has been completed in the case of a finite state space [7], and has been generalized [37] under an uniform minoration assumption for the Markov transition kernel, which typically does only hold when the state space is compact. Clearly, the whole approach relies on the existence of an exponential stability property of the prediction filter, and the main challenge currently is to get rid of this uniform minoration assumption for the Markov transition kernel [35], [60], so as to be able to consider more interesting situations, where the state space is noncompact.

**Small noise asymptotics** Another asymptotic approach can also be used, where it is rather easy to obtain interesting explicit results, in terms close to the language of nonlinear deterministic control theory [55]. Taking the simple example where the hidden state is the solution to an ordinary differential equation, or a nonlinear state model, and where the observations are subject to additive Gaussian white noise, this approach consists in assuming that covariances matrices of the state noise and of the observation noise go simultaneously to zero. If it is reasonable in many applications to consider that noise covariances are small, this asymptotic approach is less natural than the large time asymptotics, where it is enough (provided a suitable ergodicity assumption holds) to accumulate observations and to see the expected limit laws (law of large numbers, central limit theorem, etc.). In opposition, the expressions obtained in the limit (Kullback–Leibler divergence, Fisher information matrix, asymptotic covariance matrix, etc.) take here a much more explicit form than in the large time asymptotics.



The following results have been obtained using this approach

- the consistency of the maximum likelihood estimator (i.e. the convergence to the set  $M$  of global minima of the Kullback–Leibler divergence), has been obtained using large deviations techniques, with an analytical approach [51],
- if the abovementioned set  $M$  does not reduce to the true parameter value, i.e. if the model is not identifiable, it is still possible to describe precisely the asymptotic behavior of the estimators [52]: in the simple case where the state equation is a noise-free ordinary differential equation and using a Bayesian framework, it has been shown that (i) if the rank  $r$  of the Fisher information matrix  $I$  is constant in a neighborhood of the set  $M$ , then this set is a differentiable submanifold of codimension  $r$ , (ii) the posterior probability distribution of the parameter converges to a random probability distribution in the limit, supported by the manifold  $M$ , absolutely continuous w.r.t. the Lebesgue measure on  $M$ , with an explicit expression for the density, and (iii) the posterior probability distribution of the suitably normalized difference between the parameter and its projection on the manifold  $M$ , converges to a mixture of Gaussian probability distributions on the normal spaces to the manifold  $M$ , which generalized the usual asymptotic normality property,
- it has been shown [61] that (i) the parameter dependent probability distributions of the observations are locally asymptotically normal (LAN) [58], from which the asymptotic normality of the maximum likelihood estimator follows, with an explicit expression for the asymptotic covariance matrix, i.e. for the Fisher information matrix  $I$ , in terms of the Kalman filter associated with the linear tangent linear Gaussian model, and (ii) the score function (i.e. the derivative of the log-likelihood function w.r.t. the parameter), evaluated at the true value of the parameter and suitably normalized, converges to a Gaussian r.v. with zero mean and covariance matrix  $I$ .

### 3.3. Multilevel splitting for rare event simulation

See 4.2, 5.1, 6.1, and 6.2.

The estimation of the small probability of a rare but critical event, is a crucial issue in industrial areas such as nuclear power plants, food industry, telecommunication networks, finance and insurance industry, air traffic management, etc.

In such complex systems, analytical methods cannot be used, and naive Monte Carlo methods are clearly un-efficient to estimate accurately very small probabilities. Besides importance sampling, an alternate widespread technique consists in multilevel splitting [57], 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 [5], the Feynman–Kac formalism of 3.1 is well suited for the design and analysis of splitting algorithms for rare event simulation.

**Propagation of uncertainty** Multilevel splitting can be used 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 values. This problem is characterized by the lack of an analytical expression for the function, 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, and finally the complexity and / or unavailability of the source code of the computer programme, which makes any modification very difficult or even impossible, for instance to change the model as in importance sampling methods.

The key issue is to learn as fast as possible regions of the input space which contribute most to the computation of the target quantity. 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.

A further remark, is that this conditional probability distribution is precisely the optimal (zero variance) importance distribution needed to compute the probability of reaching the considered intermediate level.

**Rare event simulation** 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).

The proposed splitting method consists in (i) introducing a decreasing sequence of intermediate, more and more critical, regions in the state space, (ii) counting the fraction of trajectories that reach an intermediate region before time  $T$ , given that the previous intermediate region has been reached before time  $T$ , and (iii) regenerating the population at each stage, through redistribution. In addition to the non-intrusive behaviour of the method, the splitting methods make it possible to learn the probability distribution of typical critical trajectories, which reach the critical region before final time  $T$ , an important feature that methods based on importance sampling usually miss. Many variants have been proposed, whether

- the branching rate (number of offsprings allocated to a successful trajectory) is fixed, which allows for depth-first exploration of the branching tree, but raises the issue of controlling the population size,
- the population size is fixed, which requires a breadth-first exploration of the branching tree, with random (multinomial) or deterministic allocation of offsprings, etc.

Just as in the static case, the algorithm learns

- the transition probability between successive levels, hence the probability of reaching each intermediate level,
- and the entrance probability distribution of the Markov process in each intermediate region.

Contributions have been given to

- minimizing the asymptotic variance, obtained through a central limit theorem, with respect to the shape of the intermediate regions (selection of the importance function), to the thresholds (levels), to the population size, etc.
- controlling the probability of extinction (when not even one trajectory reaches the next intermediate level),
- designing and studying variants suited for hybrid state space (resampling per mode, marginalization, mode aggregation),

and in the static case, to

- minimizing the asymptotic variance, obtained through a central limit theorem, with respect to intermediate levels, to the Metropolis kernel introduced in the mutation step, etc.

A related issue is global optimization. Indeed, the difficult problem of finding the set  $M$  of global minima of a real-valued function  $V$  can be replaced by the apparently simpler problem of sampling a population from a probability distribution depending on a small parameter, and asymptotically supported by the set  $M$  as the small parameter goes to zero. The usual approach here is to use the cross-entropy method [65], [33], which

relies on learning the optimal importance distribution within a prescribed parametric family. On the other hand, multilevel splitting methods could provide an alternate nonparametric approach to this problem.

### 3.4. Nearest neighbor estimates

This additional topic was not present in the initial list of objectives, and has emerged only recently.

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 [66] 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.

**Nearest neighbor classification in infinite dimension** In finite dimension, the  $k$ -nearest neighbor classifier is universally consistent, i.e. its probability of error converges to the Bayes risk as  $N$  goes to infinity, whatever the joint probability distribution of the pair, provided that the ratio  $k/N$  goes to zero. Unfortunately, this result is no longer valid in general metric spaces, and the objective is to find out reasonable sufficient conditions for the weak consistency to hold. Even in finite dimension, there are exotic distances such that the nearest neighbor does not even get closer (in the sense of the distance) to the point of interest, and the state space needs to be complete for the metric, which is the first condition. Some regularity on the regression function is required next. Clearly, continuity is too strong because it is not required in finite dimension, and a weaker form of regularity is assumed. The following consistency result has been obtained: if the metric space is separable and if some Besicovich condition holds, then the nearest neighbor classifier is weakly consistent. Note that the Besicovich condition is always fulfilled in finite dimensional vector spaces (this result is called the Besicovich theorem), and that a counterexample [3] can be given in an infinite dimensional space with a Gaussian measure (in this case, the nearest neighbor classifier is clearly nonconsistent). Finally, a simple example has been found which verifies the Besicovich condition with a noncontinuous regression function.

**Rates of convergence of the functional  $k$ -nearest neighbor estimator** Motivated by a broad range of potential applications, such as regression on curves, rates of convergence of the  $k$ -nearest neighbor estimator of the regression function, based on  $N$  independent copies of the object–feature pair, have been investigated when the object is in a suitable ball in some functional space. Using compact embedding theory, explicit and general finite sample bounds can be obtained for the expected squared difference between the  $k$ -nearest neighbor estimator and the Bayes regression function, in a very general setting. The results have also been particularized to classical function spaces such as Sobolev spaces, Besov spaces and reproducing kernel Hilbert spaces. The rates obtained are genuine nonparametric convergence rates, and up to our knowledge the first of their kind for  $k$ -nearest neighbor regression.

This emerging topic has produced several theoretical advances [1], [2] in collaboration with Gérard Biau (université Pierre et Marie Curie, ENS Paris and EPI CLASSIC, INRIA Paris—Rocquencourt), and a possible target application domain has been identified in the statistical analysis of recommendation systems, that would be a source of interesting problems.

## 4. Application Domains

### 4.1. Localisation, navigation and tracking

See 6.2, and 6.3.

Among the many application domains of particle methods, or interacting Monte Carlo methods, ASPI has decided to focus on applications in localisation (or positioning), navigation and tracking [49], [43], which already covers a very broad spectrum of application domains. The objective here 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 by the target tracking community [48], [64], 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 (or has contributed recently) to several applications of particle filtering in positioning, navigation and tracking, such as geolocalisation and tracking in a wireless network, terrain-aided navigation, and data fusion for indoor localisation, see 6.3.

### 4.2. Rare event simulation

See 3.3, 5.1, 6.1, and 6.2.

Another application domain of particle methods, or interacting Monte Carlo methods, that ASPI has decided to focus on is the estimation of the small probability of a rare but critical event, in complex dynamical systems. This is a crucial issue in industrial areas such as

nuclear power plants, food industry, telecommunication networks, finance and insurance industry, 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 [57], where trajectories going towards the critical set are given offsprings, thus increasing the number of trajectories that eventually reach the critical set. This approach not only makes it possible to estimate the probability of the rare event, but also provides realizations of the random trajectory, given that it reaches the critical set, i.e. provides realizations of typical critical trajectories, an important feature that methods based on importance sampling usually miss.

ASPI is contributing (or has contributed recently) to several applications of multilevel splitting for rare event simulation, such as risk assessment in air traffic management, see 6.1, detection in sensor networks, see 6.2, and protection of digital documents.

## 5. New Results

### 5.1. Reactive trajectories for molecular dynamics

**Participants:** Frédéric Cérou, Arnaud Guyader, Florent Malrieu.

See 3.3 and 4.2.

This is a collaboration with Tony Lelièvre (CERMICS, Ecole des Ponts ParisTech).

Consider a one-dimensional Brownian motion in a double well potential. It is known that, as its variance goes to zero, the Brownian particle has to wait for a longer and longer time to jump from a well to the other. This metastable behavior is described by the Freidlin–Wentzell theory. We are investigating the length of the paths between the last passage time in the bottom of a well and the hitting time of the other one (reactive trajectory). In the case of an Ornstein–Uhlenbeck process between the wells, we obtained the remarkable result that the time length of the reactive trajectories converges in distribution when the noise intensity goes to zero, to a Gumbel random variable shifted by a deterministic term growing like minus the logarithm of the noise intensity. Our numerical simulations are also in good accordance with this theoretical result. We are also very close to extend this result to more general one dimensional diffusion processes.

### 5.2. Learning the optimal importance distribution

**Participants:** François Le Gland, Rudy Pastel.

See 3.3 and 4.2.

This is a collaboration with Jérôme Morio (ONERA Palaiseau).

As explained in 3.3, multilevel splitting ideas can be useful even to solve some static problems, such as evaluating the (small) probability that a random variable exceeds some (extreme) threshold. Incidentally, it can also be noticed that a population of particles is available at each stage of the algorithm, which is distributed according to the original distribution conditioned on exceeding the current level. Furthermore, this conditional distribution is known to be precisely the optimal importance distribution for evaluating the probability of exceeding the current level. In other words, the optimal importance distribution is learned automatically by the algorithm, as a by-product, and therefore can be used to produce an importance sampling estimate with very low variance. Building on this idea, several other iterative methods have been studied, that learn the optimal importance distribution at each stage of the algorithm, such as nonparametric adaptive importance sampling (NAIS) [69], or the cross-entropy (CE) method [65], [33]. These methods have been applied to a practical example from the aerospace industry, the evaluation of collision probabilities between two satellites, or between a satellite and space debris.

### 5.3. Impact of dimension in particle filtering and the Laplace method

**Participants:** François Le Gland, Paul Bui–Quang.

See 3.1.

This is a collaboration with Christian Musso (ONERA Palaiseau).

Particle filtering is a widely used Monte Carlo method to approximate the posterior probability distribution in non-linear filtering, with an error scaling as  $1/\sqrt{N}$  in terms of the sample size  $N$ , but otherwise independently of the underlying state dimension. However, it has been observed for a long time in practice that particle filtering can be quite inefficient when the dimension of the system is high. In a simple static linear Gaussian model, it has been possible indeed to check that the error on the estimation of the predicted likelihood, a quantitative indicator of the consistency between the prior distribution and the likelihood function, increases exponentially with the dimension [30]. This preliminary result has been extended to a non-linear / non-Gaussian model, using the Laplace method [23]. The Laplace method, which approximates multidimensional integrals accurately, has also been used to compute the asymptotic variance of the importance weights, as the sample size  $N$  goes to infinity, and to analyze its dependence on the dimension.

## 5.4. Stochastic wind–wave modelling

**Participant:** Valérie Monbet.

This is a collaboration with Pierre Ailliot (UBO) and Christophe Maisondieu (IFREMER)

Directional wave spectra generally exhibit several peaks due to the coexistence of wind sea generated by local wind conditions and swells originating from distant weather systems. This paper proposes a new algorithm for partitioning such spectra and retrieving the various systems which compose a complex sea–state. It is based on a sequential Monte–Carlo algorithm which allows to follow the time evolution of the various systems. The main particularity of the algorithm is that the dimension of the hidden state can change from time to time and a model selection step is included. The proposed methodology is validated on both synthetic and real spectra and the results are compared with a method commonly used in the literature.

## 5.5. Sequential data assimilation: ensemble Kalman filter vs. particle filter

**Participants:** François Le Gland, Valérie Monbet.

See 6.4.

Surprisingly, very little was known about the asymptotic behaviour of the ensemble Kalman filter [38], [39], [40], 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 [25]. In addition, the limit of the empirical distribution of the ensemble elements has been exhibited, which differs from the usual Bayesian filter. The next step 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.

# 6. Contracts and Grants with Industry

## 6.1. Safety, complexity and responsibility based design and validation of highly automated air traffic management (iFLY) — FP6 Aerospace

**Participant:** François Le Gland.

See 3.3 and 4.2.

*INRIA contract ALLOC 2399 — May 2007 to August 2011.*

This FP6 project is coordinated by National Aerospace Laboratory (NLR) (The Netherlands), and ASPI is also collaborating with University of Twente (The Netherlands) and Direction des Services de la Navigation Aérienne (DSNA).

The objective of **iFLY** is 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 proposed research combines expertise in air transport human factors, safety and economics with analytical and Monte Carlo simulation methodologies. The contribution of ASPI to this project concerns 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 [29] suited for hybrid state space (resampling per mode, marginalization) are currently investigated [18].

## 6.2. Optimization of sensors location and activation — contract with DGA / Techniques navales

**Participant:** François Le Gland.

See 3.3 and 4.2

INRIA contract ALLOC 4233 — April 2009 to September 2011.

This is a collaboration with Sébastien Paris (université Paul Cézanne), related with the supervision of the PhD thesis of Mathieu Chouchane.

The objective of this project is 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. Because of the many constraints including timing constraints involved in this optimization problem, a stochastic algorithm is preferred here over a deterministic algorithm. The underlying idea is 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 according to a probability distribution depending on a small parameter, which asymptotically concentrates on the set of global maxima of the cost function, as the small parameter goes to zero. The usual approach here is to use the cross-entropy method [65], [33].

The contribution of ASPI has been to propose a multilevel splitting algorithm, in order to evaluate the probability of detection for a given configuration. When this probability is small, these methods are known to provide a significant reduction in the variance of the relative error.

### 6.3. Information fusion for localisation (FIL) — ANR Télécommunications

**Participant:** François Le Gland.

See 4.1.

INRIA contract ALLOC 2856 — January 2008 to June 2011.

This ANR project is coordinated by Thalès Alenia Space.

The overall objective is to study and demonstrate information fusion algorithms for localisation of pedestrian users in an indoor environment, where GPS solution cannot be used. The sought design combines

- a pedestrian dead-reckoning (PDR) unit, providing noisy estimates of the linear displacement, angular turn, and possibly of the level change through an additional pressure 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 localisation when two users meet and exchange their respective position estimates.

Besides particle methods, which are proposed as the basic information fusion algorithm for the centralized server-based implementation, simpler algorithms such as the extended Kalman filter (EKF) or the unscented Kalman filter (UKF) have been investigated, to be used for the local PDA-based implementation with a map of a smaller part of the building. Constraints could be taken care of automatically with the help of a Voronoi graph [62], but this approach implies heavy pre-computations. A more direct approach, taking care of constraints on the fly, using a simple rejection method, has been preferred. Adapting the sample size using KLD-sampling [42] has also been investigated, which could be useful in the case of a poor initial information, or if the user walks in poorly informative area (open zone, absence of beacons). Collaboration between users has been implemented [41], which allows from a user with a poor localization to benefit from the more accurate localization of another user. In this implementation, the latter user is seen by the former user as a ranging beacon with uncertain position. See [28], [54] for a description of the overall fusion algorithm and an illustration with simulation results.

## 6.4. Ensemble methods for prediction and data assimilation (PREVASSEMBLE) — ANR Conception et Simulation

**Participants:** François Le Gland, Valérie Monbet.

See 5.5.

*INRIA contract ALLOC 3767 — January 2009 to December 2012.*

This ANR project is coordinated by École Normale Supérieure, Paris. The other partner is Météo-France. This is a collaboration with Étienne Mémin and Anne Cuzol (INRIA Rennes Bretagne Atlantique, project-team FLUMINANCE).

The contribution of ASPI to this project is to continue [25] the comparison of sequential data assimilation methods initiated in [68], such as the ensemble Kalman filter (EnKF) and the weighted ensemble Kalman filter (WEnKF), with particle filters. This comparison will be made on the basis of asymptotic variances, as the ensemble or sample size goes to infinity, and also on the impact of dimension on small sample behavior.

## 6.5. Slightly-multimodal transmission and detection (STRADE) — ANR Réseaux du Futur et Services

**Participant:** Frédéric Cérou.

*INRIA contract ALLOC 4402 — November 2009 to October 2012.*

This ANR project is also coordinated by Alcatel-Lucent Bell Labs France.

The focus of our project is to reduce the impact of nonlinear effect. The objective is twofold: specify, design, realize and evaluate fibres of reduced nonlinear effects by firstly increasing the effective area to unprecedented values and secondly, by splitting optical power along two modes, using bimodal propagation. While the first step is ambitious but primarily relies in the evolution of current fibre technologies, the second is disruptive and requires not only deep changes in fibre technologies but also new advanced transmitter / receiver equipment, preferably based on coherent detection. Naturally, bimodal propagation also brings another key advantage, namely a twofold increase of system capacity.

# 7. Dissemination

## 7.1. Scientific animation

Jointly with the team Processus Stochastiques of IRMAR, ASPI organizes a working group on the **Freidlin–Wentzell theory** and its applications. One of the main goals of these talks is to study the theory of large deviations which describe how a metastable diffusion process evolves. Moreover, several talks are dedicated to simulation algorithms and applications (molecular dynamics, turbulence modelling)

François Le Gland has been a member of the committee for the PhD theses of Xuan-Binh Lam (université de Rennes 1, advisor: Laurent Mevel) and Christophe Avenel (université de Rennes 1, advisor: Étienne Mémin) and for the habilitation thesis of Frédéric Dambreville (université de Bretagne Occidentale).

François Le Gland is a member of the “conseil d’UFR” of the department of mathematics of université de Rennes 1.

Florent Malrieu is a member of the “conseil” of IRMAR (institut de recherche mathématiques de Rennes, UMR 6625).

Valérie Monbet is the director of the master on statistics and econometry at université de Rennes 1.

## 7.2. Teaching

Arnaud Guyader is a member of the committee of “oraux blancs d’agrégation de mathématiques” for ENS Cachan at Ker Lann.



François Le Gland gives a course on **Kalman filtering and hidden Markov models**, at université de Rennes 1, within the master SISEA (signal, image, systèmes embarqués, automatique, école doctorale MATISSE), a 3rd year course on **Bayesian filtering and particle approximation**, at ENSTA (école nationale supérieure de techniques avancées), Paris, within the systems and control module, a 3rd year course on linear and nonlinear filtering, **linear and nonlinear filtering**, at ENSAI (école nationale de la statistique et de l'analyse de l'information), Ker Lann, within the statistical engineering track, and a 3rd year course on **hidden Markov models**, at Télécom Bretagne, Brest.

Florent Malrieu gives a course on Markov models at université de Rennes 1, within the probability and statistics track of the master on mathematics and applications.

Valérie Monbet gives several courses on data analysis, on time series and hidden Markov models, and on mathematical statistics, all at université de Rennes 1 within the master on statistics and econometry.

### 7.3. PhD and Habilitation Theses

Arnaud Guyader has defended his habilitation thesis [10], *Contributions to nonparametric estimation and rare event simulation*, at université de Rennes 2 in December 2011.

François Le Gland is currently supervising four PhD students

- Rudy Pastel, title: *Estimation of rare event probabilities and extreme quantiles. Applications in the aerospace domain*, started in October 2008, expected defense in February 2012, funding: ONERA grant, co-direction: Jérôme Morio (ONERA, Palaiseau).
- Paul Bui-Quang, provisional title: *The Laplace method for particle filtering*, started in October 2009, funding: ONERA grant, co-direction: Christian Musso (ONERA, Palaiseau).
- Alexandre Lepoutre, provisional title: *Monte Carlo methods for dim target tracking*, started in October 2010, funding: ONERA grant, co-direction: Olivier Rabaste (ONERA, Palaiseau).
- Damien Jacquemart, provisional title: *Rare event methods for the estimation of collision risk*, started in October 2011, funding: DGA / ONERA grant, co-direction: Jérôme Morio (ONERA, Palaiseau).

### 7.4. Participation in workshops, seminars, lectures, etc.

In addition to presentations with a publication in the proceedings, and which are listed at the end of the document, members of ASPI have also given the following presentations.

Arnaud Guyader has given a talk on multilevel Monte Carlo for extreme quantiles and probabilities at the SIAM Conference on Computational Science and Engineering, held in Reno in March 2011. He has also been invited by Nicolas Hengartner to visit Los Alamos National Laboratories in April 2011.

Frédéric Cerou has been invited to give a talk on smoothed splitting methods for counting at the 16th INFORMS Applied Probability Conference, held in Stockholm in July 2011.

Florient Malrieu has been invited to give seminar talks on non-uniqueness of equilibrium states for McKean-Vlasov equations in Toulouse in March 2011, on long time behavior of McKean-Vlasov equations in Pau in March 2011, on Markov modulated ordinary differential equations in Bordeaux in March 2011, on functional inequalities for mixtures in Grenoble in March 2011, on functional inequalities and concentration inequalities for mixtures in Vannes in June 2011, on ergodicity of piecewise deterministic Markov processes in Paris in June 2011. He has also been invited to give a three hours lecture on uniform (in time) propagation of chaos for a class of McKean-Vlasov equations, at the workshop on Mean Field Limit, held at IHP (institut Henri Poincaré) Paris in March 2011.

François Le Gland has been invited to give a talk on marginalization for rare event simulation in switching diffusions, at the workshop on Numerical Methods for Filtering and for Parabolic PDE's, held at Imperial College in September 2011. He has also presented a poster on information fusion for indoor navigation, at the DGA seminar on Information Fusion and Planning for Surveillance and Intelligence, held at ENSTA ParisTech in June 2011.

Paul Bui–Quang has been invited to give seminar talks on importance sampling in high–dimension via the Laplace method at the seminar of the **BigMC** ANR project in April 2011 and on the Laplace method for particle filtering at the ENSAI seminar for PhD students in May 2011.

Rudy Pastel has given talks on splitting methods for the evaluation of satellite vs. debris collision probabilities at the 3rd IEEE International Conference on Computer Modeling and Simulation (ICCMS), held in Mumbai in January 2011 and at the European Conference for Aero–Space Sciences (EUCASS), held in St. Petersburg in July 2011.

## 8. Bibliography

### Major publications by the team in recent years

- [1] G. BIAU, F. CÉROU, A. GUYADER. *On the rate of convergence of the bagged nearest neighbor estimate*, in "Journal of Machine Learning Research", February 2010, vol. 11, p. 687–712, <http://jmlr.csail.mit.edu/papers/v11/biau10a.html>.
- [2] G. BIAU, F. CÉROU, A. GUYADER. *On the rate of convergence of the functional  $k$ –nearest neighbor estimates*, in "IEEE Transactions on Information Theory", April 2010, vol. IT–56, n<sup>o</sup> 4, p. 2034–2040, <http://dx.doi.org/10.1109/TIT.2010.2040857>.
- [3] F. CÉROU, A. GUYADER. *Nearest neighbor classification in infinite dimension*, in "ESAIM : Probability and Statistics", 2006, vol. 10, p. 340–355, <http://dx.doi.org/10.1051/ps:2006014>.
- [4] F. CÉROU, A. GUYADER. *Adaptive multilevel splitting for rare event analysis*, in "Stochastic Analysis and Applications", March 2007, vol. 25, n<sup>o</sup> 2, p. 417–443, <http://dx.doi.org/10.1080/07362990601139628>.
- [5] F. CÉROU, P. DEL MORAL, F. LE GLAND, P. LEZAUD. *Genetic genealogical models in rare event analysis*, in "ALEA, Latin American Journal of Probability and Mathematical Statistics", 2006, vol. 1, p. 181–203 (electronic), Paper 01–08.
- [6] T. FURON, A. GUYADER, F. CÉROU. *On the design and optimization of Tardos probabilistic fingerprinting codes*, in "10th International Workshop on Information Hiding, Santa Barbara", Berlin, K. SOLANKI, K. SULLIVAN, U. MADHOW (editors), Lecture Notes in Computer Science, Springer, Berlin, May 2008, vol. 5284, p. 341–356, [http://dx.doi.org/10.1007/978-3-540-88961-8\\_24](http://dx.doi.org/10.1007/978-3-540-88961-8_24).
- [7] F. LE GLAND, L. MEVEL. *Exponential forgetting and geometric ergodicity in hidden Markov models*, in "Mathematics of Control, Signals, and Systems", 2000, vol. 13, n<sup>o</sup> 1, p. 63–93, <http://dx.doi.org/10.1007/PL00009861>.
- [8] F. LE GLAND, N. OUDJANE. *A sequential algorithm that keeps the particle system alive*, in "Stochastic Hybrid Systems : Theory and Safety Critical Applications", Berlin, H. A. P. BLOM, J. LYGEROS (editors), Lecture Notes in Control and Information Sciences, Springer–Verlag, Berlin, 2006, n<sup>o</sup> 337, p. 351–389, [http://dx.doi.org/10.1007/11587392\\_11](http://dx.doi.org/10.1007/11587392_11).
- [9] C. MUSSO, N. OUDJANE, F. LE GLAND. *Improving regularized particle filters*, in "Sequential Monte Carlo Methods in Practice", New York, A. DOUCET, N. DE FREITAS, N. J. GORDON (editors), Statistics for Engineering and Information Science, Springer–Verlag, New York, 2001, chap. 12, p. 247–271.

## Publications of the year

### Doctoral Dissertations and Habilitation Theses

- [10] A. GUYADER. *Contributions à l'estimation non paramétrique et à la simulation d'évènements rares*, Université de Rennes 2, December 2011, Habilitation à Diriger des Recherches.

### Articles in International Peer-Reviewed Journal

- [11] P. AILLIOT, A. BAXEVANI, A. CUZOL, V. MONBET, N. RAILLARD. *Space-time models for moving fields. Application to significant wave height*, in "Environmetrics", May 2011, vol. 22, n<sup>o</sup> 3, p. 354–369, <http://dx.doi.org/10.1002/env.1061>.
- [12] P. AILLIOT, V. MONBET. *Markov switching autoregressive models for wind time series*, in "Environmental Modelling & Software", 2012, to appear, <http://dx.doi.org/10.1016/j.envsoft.2011.10.011>.
- [13] F. CÉROU, P. DEL MORAL, A. GUYADER. *A nonasymptotic variance theorem for unnormalized Feynman-Kac particle models*, in "Annales de l'Institut Henri Poincaré, Probabilités et Statistiques", 2011, vol. 47, n<sup>o</sup> 3, p. 629–649, <http://dx.doi.org/10.1214/10-AIHP358>.
- [14] F. CÉROU, P. DEL MORAL, A. GUYADER, T. FURON. *Sequential Monte Carlo for rare event estimation*, in "Statistics and Computing", 2012, to appear, <http://dx.doi.org/10.1007/s11222-011-9231-6>.
- [15] F. CÉROU, A. GUYADER, T. LELIÈVRE, D. POMMIER. *A multiple replica approach to simulate reactive trajectories*, in "Journal of Chemical Physics", February 2011, vol. 134, n<sup>o</sup> 5, <http://dx.doi.org/10.1063/1.3518708>.
- [16] F. CÉROU, A. GUYADER, R. Y. RUBINSTEIN, R. VAISMAN. *Smoothed splitting method for counting*, in "Stochastic Models", 2011, vol. 27, n<sup>o</sup> 4, p. 629–650, <http://dx.doi.org/10.1080/15326349.2011.614188>.
- [17] A. GUYADER, N. W. HENGARTNER, E. MATZNER-LØBER. *Simulation and estimation of extreme quantiles and extreme probabilities*, in "Applied Mathematics & Optimization", October 2011, vol. 64, n<sup>o</sup> 2, p. 171–196, <http://dx.doi.org/10.1007/s00245-011-9135-z>.
- [18] J. KRYSZTUL, F. LE GLAND, P. LEZAUD. *Sampling per mode for rare event simulation in switching diffusions*, in "Stochastic Processes and their Applications", 2012, to appear.
- [19] J. MORIO, R. PASTEL. *Plug-in estimation of d-dimensional density minimum volume set of a rare event in a complex system*, in "Proceedings of the IMechE, Part O, Journal of Risk and Reliability", 2012, to appear, <http://dx.doi.org/10.1177/1748006X11426973>.
- [20] J. MORIO, R. PASTEL, F. LE GLAND. *Estimation de probabilités et de quantiles rares pour la caractérisation d'une zone de retombée d'un engin*, in "Journal de la Société Française de Statistique", 2011, vol. 152, n<sup>o</sup> 4.
- [21] J. MORIO, R. PASTEL, F. LE GLAND. *Missile target accuracy estimation with importance splitting*, in "Aerospace Science and Technology", 2012, to appear, <http://dx.doi.org/10.1016/j.ast.2011.12.006>.

### International Conferences with Proceedings

- [22] M. CHOUCANE, S. PARIS, F. LE GLAND, C. MUSSO, D.-T. 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, ISIF, July 2011, p. 99–101.
- [23] C. MUSSO, P. BUI-QUANG, F. LE GLAND. *Introducing the Laplace approximation in particle filtering*, in "Proceedings of the 14th International Conference on Information Fusion", Chicago, ISIF, July 2011, p. 290–297.

### National Conferences with Proceeding

- [24] M. BOCQUEL, A. LEPOUTRE, O. RABASTE, F. 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", September 2011.

### Scientific Books (or Scientific Book chapters)

- [25] F. LE GLAND, V. MONBET, V.-D. TRAN. *Large sample asymptotics for the ensemble Kalman filter*, in "The Oxford Handbook of Nonlinear Filtering", Oxford, D. O. CRISAN, B. L. ROZOVSKII (editors), Oxford University Press, Oxford, 2011, chap. 22, p. 598–631.

### References in notes

- [26] A. DOUCET, N. DE FREITAS, N. J. GORDON (editors). *Sequential Monte Carlo methods in practice*, Statistics for Engineering and Information Science, Springer-Verlag, New York, 2001.
- [27] M. S. ARULAMPALAM, S. MAKSELL, N. J. GORDON, T. CLAPP. *A tutorial on particle filters for online nonlinear / non-Gaussian Bayesian tracking*, in "IEEE Transactions on Signal Processing", February 2002, vol. SP-50, n° 2 (Special issue on Monte Carlo Methods for Statistical Signal Processing), p. 174–188.
- [28] P. BLANCHART, L. HE, F. LE GLAND. *Information fusion for indoor localization*, in "Proceedings of the 12th International Conference on Information Fusion, Seattle 2009", ISIF, July 2009, p. 2083–2090.
- [29] H. A. P. BLOM, B. BAKKER, J. KRISTUL. *Rare event estimation for a large scale stochastic hybrid system with air traffic application*, in "Monte Carlo Methods for Rare Event Analysis", Chichester, G. RUBINO, B. TUFFIN (editors), John Wiley & Sons, Chichester, 2009, chap. 9, p. 193–214.
- [30] P. BUI-QUANG, C. MUSSO, F. LE GLAND. *An insight into the issue of dimensionality in particle filtering*, in "Proceedings of the 13th International Conference on Information Fusion, Edinburgh 2010", ISIF, July 2010, Paper We2.1.2.
- [31] O. CAPPÉ, S. J. GODSILL, É. MOULINES. *An overview of existing methods and recent advances in sequential Monte Carlo*, in "Proceedings of the IEEE", May 2007, vol. 95, n° 5, p. 899–924.
- [32] O. CAPPÉ, É. MOULINES, T. RYDÉN. *Inference in hidden Markov models*, Springer Series in Statistics, Springer-Verlag, New York, 2005.
- [33] P.-T. DE BOER, D. P. KROESE, S. MANNOR, R. Y. RUBINSTEIN. *A tutorial on the cross-entropy method*, in "Annals of Operations Research", January 2005, vol. 134 (Special issue on the Cross-Entropy Method for Combinatorial Optimization, Rare Event Simulation and Neural Computation), n° 1, p. 19–67.

- 
- [34] P. DEL MORAL. *Feynman–Kac formulae. Genealogical and interacting particle systems with applications*, Probability and its Applications, Springer–Verlag, New York, 2004.
- [35] P. DEL MORAL, A. GUIONNET. *On the stability of interacting processes with applications to filtering and genetic algorithms*, in "Annales de l'Institut Henri Poincaré, Probabilités et Statistiques", 2001, vol. 37, n<sup>o</sup> 2, p. 155–194.
- [36] P. DEL MORAL, L. MICLO. *Branching and interacting particle systems approximations of Feynman–Kac formulae with applications to nonlinear filtering*, in "Séminaire de Probabilités XXXIV", Berlin, J. AZÉMA, M. ÉMERY, M. LEDOUX, M. YOR (editors), Lecture Notes in Mathematics, Springer–Verlag, Berlin, 2000, vol. 1729, p. 1–145.
- [37] R. DOUC, C. MATIAS. *Asymptotics of the maximum likelihood estimator for general hidden Markov models*, in "Bernoulli", June 2001, vol. 7, n<sup>o</sup> 3, p. 381–420.
- [38] G. EVENSEN. *Ensemble Kalman filter : theoretical formulation and practical implementations*, in "Ocean Dynamics", 2003, vol. 53, p. 343–367.
- [39] G. EVENSEN. *Sampling strategies and square root analysis schemes for the EnKF*, in "Ocean Dynamics", 2004, vol. 54, p. 539–560.
- [40] G. EVENSEN. *Data assimilation. The ensemble Kalman filter*, Springer–Verlag, Berlin, 2006.
- [41] D. FOX, W. BURGARD, H. KRUPPA, S. THRUN. *A probabilistic approach to collaborative multi-robot localization*, in "Autonomous Robots", June 2000, vol. 8, n<sup>o</sup> 3 (Special issue on Heterogeneous Multi-Robot Systems), p. 325–344.
- [42] D. FOX. *Adapting the sample size in particle filters through KLD-sampling*, in "International Journal of Robotics Research", December 2003, vol. 22, n<sup>o</sup> 12, p. 985–1004.
- [43] D. FOX, J. HIGHTOWER, L. LIAO, D. SCHULZ, G. BORRIELLO. *Bayesian filtering for location estimation*, in "IEEE Pervasive Computing", July/September 2003, vol. 2, n<sup>o</sup> 3, p. 24–33.
- [44] D. FOX, S. THRUN, W. BURGARD, F. DELLAERT. *Particle filters for mobile robot localization*, in "Sequential Monte Carlo Methods in Practice", New York, A. DOUCET, N. DE FREITAS, N. J. GORDON (editors), Statistics for Engineering and Information Science, Springer–Verlag, New York, 2001, chap. 19, p. 401–428.
- [45] D. FRENKEL, B. SMIT. *Understanding molecular simulation. From algorithms to applications*, Computational Science Series, 2nd, Academic Press, San Diego, 2002, vol. 1.
- [46] P. GLASSERMAN. *Monte Carlo methods in financial engineering*, Applications of Mathematics, Springer–Verlag, New York, 2004, vol. 53.
- [47] P. GLASSERMAN, P. HEIDELBERGER, P. SHAHABUDDIN, T. ZAJIC. *Multilevel splitting for estimating rare event probabilities*, in "Operations Research", July–August 1999, vol. 47, n<sup>o</sup> 4, p. 585–600.
- [48] N. J. GORDON, D. J. SALMOND, A. F. M. SMITH. *Novel approach to nonlinear / non-Gaussian Bayesian state estimation*, in "IEE Proceedings, Part F", April 1993, vol. 140, n<sup>o</sup> 2, p. 107–113.

- 
- [49] F. GUSTAFSSON, F. GUNNARSSON, N. BERGMAN, U. FORSELL, J. JANSSON, R. KARLSSON, P.-J. NORDLUND. *Particle filters for positioning, navigation, and tracking*, in "IEEE Transactions on Signal Processing", February 2002, vol. SP-50, n<sup>o</sup> 2 (Special issue on Monte Carlo Methods for Statistical Signal Processing), p. 425–437.
- [50] M. ISARD, A. BLAKE. CONDENSATION — *Conditional density propagation for visual tracking*, in "International Journal of Computer Vision", August 1998, vol. 29, n<sup>o</sup> 1, p. 5–28.
- [51] M. R. JAMES, F. LE GLAND. *Consistent parameter estimation for partially observed diffusions with small noise*, in "Applied Mathematics & Optimization", July/August 1995, vol. 32, n<sup>o</sup> 1, p. 47–72.
- [52] M. JOANNIDES, F. LE GLAND. *Small noise asymptotics of the Bayesian estimator in nonidentifiable models*, in "Statistical Inference for Stochastic Processes", 2002, vol. 5, n<sup>o</sup> 1, p. 95–130.
- [53] G. KITAGAWA. *Monte Carlo filter and smoother for non-Gaussian nonlinear state space models*, in "Journal of Computational and Graphical Statistics", 1996, vol. 5, n<sup>o</sup> 1, p. 1–25.
- [54] D. KUBRAK, F. LE GLAND, L. HE, Y. OSTER. *Multi-sensor fusion for localization. Concept and simulation results*, in "Proceedings of the 2009 ION Conference on Global Navigation Satellite Systems, Savannah 2009", ION, September 2009.
- [55] Y. A. KUTOYANTS. *Identification of dynamical systems with small noise*, Mathematics and its Applications, Kluwer Academic Publisher, Dordrecht, 1994, vol. 300.
- [56] H. R. KÜNSCH. *Recursive Monte Carlo filters : Algorithms and theoretical analysis*, in "The Annals of Statistics", October 2005, vol. 33, n<sup>o</sup> 5, p. 1983–2021.
- [57] P. L'ÉCUYER, V. DEMERS, B. TUFFIN. *Rare events, splitting, and quasi-Monte Carlo*, in "ACM Transactions on Modeling and Computer Simulation", April 2007, vol. 17, n<sup>o</sup> 2 (Special issue honoring Perwez Shahabuddin), Article 9.
- [58] L. LE CAM. *Asymptotic methods in statistical decision theory*, Springer Series in Statistics, Springer-Verlag, New York, 1986.
- [59] F. LE GLAND, N. OUDJANE. *A robustification approach to stability and to uniform particle approximation of nonlinear filters : the example of pseudo-mixing signals*, in "Stochastic Processes and their Applications", August 2003, vol. 106, n<sup>o</sup> 2, p. 279-316.
- [60] F. LE GLAND, N. OUDJANE. *Stability and uniform approximation of nonlinear filters using the Hilbert metric, and application to particle filters*, in "The Annals of Applied Probability", February 2004, vol. 14, n<sup>o</sup> 1, p. 144–187.
- [61] F. LE GLAND, B. WANG. *Asymptotic normality in partially observed diffusions with small noise : application to FDI*, in "Workshop on Stochastic Theory and Control, University of Kansas 2001. In honor of Tyrone E. Duncan on the occasion of his 60th birthday", Berlin, B. PASIK-DUNCAN (editor), Lecture Notes in Control and Information Sciences, Springer-Verlag, Berlin, 2002, n<sup>o</sup> 280, p. 267–282.

- 
- [62] L. LIAO, D. FOX, J. HIGHTOWER, H. KAUTZ, D. SCHULZ. *Voronoi tracking : Location estimation using sparse and noisy sensor data*, in "Proceedings of the IEEE / RSJ International Conference on Intelligent Robots and Systems, Las Vegas 2003", October 2003, p. 723–728.
- [63] J. S. LIU. *Monte Carlo strategies in scientific computing*, Springer Series in Statistics, Springer–Verlag, New York, 2001.
- [64] B. RISTIĆ, M. S. ARULAMPALAM, N. J. GORDON. *Beyond the Kalman Filter : Particle Filters for Tracking Applications*, Artech House, Norwood, MA, 2004.
- [65] R. Y. RUBINSTEIN, D. 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.
- [66] C. J. STONE. *Consistent nonparametric regression (with discussion)*, in "The Annals of Statistics", July 1977, vol. 5, n<sup>o</sup> 4, p. 595–645.
- [67] S. THRUN, W. BURGARD, D. FOX. *Probabilistic robotics*, Intelligent Robotics and Autonomous Agents, The MIT Press, Cambridge, MA, 2005.
- [68] V.–D. TRAN. *Assimilation de données : les propriétés asymptotiques du filtre de Kalman d'ensemble*, Université de Bretagne Sud, Vannes, June 2009, <http://tel.archives-ouvertes.fr/tel-00412447/fr/>.
- [69] P. ZHANG. *Nonparametric importance sampling*, in "Journal of the American Statistical Association", September 1996, vol. 91, n<sup>o</sup> 434, p. 1245–1253.
- [70] A. W. VAN DER VAART, J. A. WELLNER. *Weak convergence and empirical processes*, Springer Series in Statistics, Springer–Verlag, Berlin, 1996.