R INRIA

Project-Team ASPI

Applications of interacting particle systems to statistics

Rennes - Bretagne-Atlantique

Theme : Stochastic Methods and Models

Activity Report

2009

Table of contents

1. Team

Research Scientist

Frédéric Cérou [CR INRIA]

Jean–Pierre Le Cadre [DR CNRS, deceased July 1, 2009, HdR]

François Le Gland [Team leader, DR INRIA]

Faculty Member

Arnaud Guyader [assistant professor, université de Rennes 2, délégation INRIA from September 1, 2007 to August 31, 2009]

Technical Staff

Rui Han [ingénieur–expert, ANR project COHDEQ40, from January 1, 2009 to October 30, 2009]

Liyun He [ingénieur–expert, ANR project FIL, from October 1, 2008 to September 30, 2009]

Cyrille Jégourel [ingénieur–expert, ANR project NEBBIANO, from October 1, 2008 to July 31, 2009]

Jean-Baptiste Le Meur [ACET scientifique, ANR project NEBBIANO, from November 15, 2009 to December 15, 2009]

Nicolas Ramin [ingénieur–expert, DGA, from January 1, 2009 to October 31, 2009]

PhD Student

Sébastien Beyou [université de Rennes 1, ANR project PREVASSEMBLE, jointly with project–team FLUMINANCE, since November 1, 2009]

Paul Bui–Quang [université de Rennes 1, ONERA grant, since October 15, 2009]

Renaud Cariou [université de Rennes 1, ingénieur des études et techniques d'armement, DGA / CELAR Bruz]

Francis Céleste [université de Rennes 1, ingénieur des études et techniques d'armement, DGA / CEP Arcueil]

Nordine El Baraka [université de Rennes 1, CIFRE grant with Thalès Communications]

Adrien Ickowicz [université de Rennes 1, DGA / CNRS grant]

Rudy Pastel [université de Rennes 1, ONERA grant]

Vu–Duc Tran [université de Bretagne Sud, région Bretagne grant, until June 30, 2009]

Post-Doctoral Fellow

Fida El Haje Hussein [from September 1, 2008 to July 31, 2009]

Administrative Assistant

Stéphanie Lemaile [until June 15, 2009]

Fabienne Cuyollaa [since June 15, 2009]

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, 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, detection in sensor networks, risk assessment in air traffic management, protection of digital documents, and credit risk estimation.

Our second research area concerns nearest neighbor estimation. This set of methods belongs to the class of supervised statistical learning algorithms. They aim at predicting some feature of a new object from a learning N–sample in the form of already known object–feature pairs. There are many different approches to this problem, and we decided to focus on the nearest neighbor approach, which only considers the k nearest objects of the new one, with k smaller than N, and for some given metric. Then we predict the unkown feature value by a mean on the features of the neighbors in the case of regression (continuous case), or a majority vote in the case of classification (discrete case). Despite its simplicity, this approach works very well in practice and has led to a large amount of both applied and theoretical literature. Nevertheless, there are still open questions, especially about the convergence of the method when the objects take values in an infinite dimensional space, or in general about the rate of convergence in different settings, or for different choices of k.

3. Scientific Foundations

3.1. Monte Carlo methods

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) [48], 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, a brief mathematical presentation of which is given in 3.2, 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*. The very important case of a sampled continuous–time Markov process, e.g. the solution of a stochastic differential equation driven by a Wiener process or a more general Lévy process, is also covered.

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* [51], *Monte Carlo filter* [55] or *condensation* (conditional density propagation) algorithm [53], 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 [52], [45], visual tracking [53], mobile robotics [46], [72], ubiquitous computing and ambient intelligence, sensor networks, risk evaluation and simulation of rare events [49], genetics, molecular simulation [47], etc.

Other examples of the many applications of particle filtering can be found in the contributed volume [28] 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 [29] can be found, and in the textbook [68] 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 [66].

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.

3.2. Particle approximations of Feynman–Kac distributions

The following abstract point of view, developed and extensively studied by Pierre Del Moral [37], [35], 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 flows. Feynman–Kac distributions are characterized by a Markov chain and by nonnegative potential functions that play the role of selection functions. They naturally arise whenever importance sampling is used: this applies for instance to simulation of rare events, to filtering, i.e. to state estimation in hidden Markov models (HMM), etc. To solve *numerically* the recurrent equation satisfied by the Feynman–Kac distributions, and in view of the basic assumption that it is easy to *simulate* r.v.'s according to the Markov transition kernel, i.e. to mimic the evolution of the Markov chain, and that it is easy to *evaluate* the potential functions, the original idea behind particle methods consists of looking for an approximation in the form of a (possibly weighted) empirical probability distribution associated with a system of particles. The approximation is completely characterized by the set of particle positions and weights, and the algorithm is completely described by the mechanism which builds this set recursively. In practice, in the simplest version of the algorithm, known as the *bootstrap* algorithm, particles

- are selected according to their respective weights (selection step),
- move according to the Markov transition kernel (mutation step),
- are weighted by evaluating the fitness function (weighting step).

The algorithm yields a numerical approximation of the Feynman–Kac distribution as the weighted empirical probability distribution associated with a system of particles, and 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 [73]), see e.g. the survey article [37] or the recent textbook [35], and references therein

convergence in L^p, convergence as empirical processes indexed by classes of functions, uniform convergence in time, see also [8], [9], central limit theorem, see also [59], [63], propagation of chaos, large deviations principle, moderate deviations principle [38], etc.

Beyond the simplest *bootstrap* version of the algorithm, many algorithmic variations have been proposed [40], [31], and are commonly used in practice. For instance (i) in the selection step, sampling with replacement could be replaced with other redistribution schemes so as to reduce the variance (this issue has also been addressed in genetic algorithms), and (ii) to reduce the variance and to save computational effort, it is often a good idea not to redistribute the particles at each time step, but only when the weights are too far from equidistribution. Even with interacting Monte Carlo methods, it could happen that some particles generated in one time step have a negligible weight: if this happens for too many particles in the sample, then computing power has been wasted, and it has been suggested to use importance sampling again in the mutation step, i.e. to let particles explore the state space under the action of an alternate wrong mutation kernel, and to weight the particles according to their likelihood for the true model, so as to compensate for the wrong modeling. More specifically, using an arbitrary importance decomposition results in the following general algorithm, known as the *sampling with importance resampling* (SIR) algorithm, in which particles

- are selected according to their respective weights (selection step),
- move according to the proposed importance Markov transition kernel (mutation step),
- are weighted by evaluating the corresponding importance function, which now depends on the Markov transition (weighting step).

Many of the early convergence results proved in the literature assume that particles are redistributed (i) using sampling with replacement and (ii) at each time step, and move according to the original Markov transition kernel. Systematically studying the impact of the proposed algorithmic variants on the convergence results is still the subject of active research.

3.3. 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 3.2. 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 clasical 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 [39] 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 [36], [9], 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 [58]. 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 [54],

- 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 [6]: 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 [64] that (i) the parameter dependent probability distributions of the observations are locally asymptotically normal (LAN) [61], 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.

3.4. Multilevel splitting for rare event simulation

See 4.2, 6.3, 6.4, 6.5, 6.6, 6.9, 7.1, 7.3 and 7.7.

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 unefficient to estimate accurately very small probabilities. Besides importance sampling, an alternate widespread technique consists in multilevel splitting [60], 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 3.2 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, conditionned 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 [69], [34], 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.

3.5. Nearest neighbor estimates

See 6.1 and 6.2.

In pattern recognition and statistical learning, nearest neighbor algorithms are amongst the most simple available. Nevertheless, they are also very powerful and since the pioneering works by Fix and Hodges [41], [42] they have generated a large amount of literature and developments. Basically, given a training set of data, i.e. an N–sample of i.i.d. object–feature pairs (X_i, Y_i) for $i = 1, \cdots, N$, with real–valued features, we want to be able to generalize, that is to guess the feature Y associated with any new object X, with the same probability distribution as the X_i's. 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 X, for some given metric. From the beginning it was clear that even simple, this method is very powerful.

In general, there is no way to guess exactly the value of Y, and the minimal error that can be done is that of the Bayes estimate $\mathbb{E}[Y \mid X]$, which cannot be computed by lack of knowledge of the distribution of the pair, but the Bayes estimate will help us to characterize the strength of the method. So the best we can wish is that our estimate converges, say when the sample size grows, to the Bayes estimate. This is what has been proved in great generality by Stone [71] for the mean square convergence, provided that X is a d-dimensional vector, Y is square–integrable, and the ratio k/N goes to 0. Nearest neighbor estimate is not the only local averaging estimate having this property, but it is arguably the simplest.

The situation is radically different in general infinite dimensional spaces. In this respect, Cérou and Guyader [3] present counterexamples indicating that the estimate is not consistent, and they argue that restrictions on the state space and the distribution of (X, Y) cannot be dispensed with. First of all, it must be separable for the norm used to compute the neighbors, as already noticed by Cover and Hart [33]. But this is not enough. By working out arguments in Preiss [67], Cérou and Guyader [3] exhibit a random variable X with Gaussian distribution in a separable Hilbert space for which the estimate fails to be consistent. On the positive side, these authors provide a general condition, called the μ–continuity condition, which ensures the consistency of the estimate. Even with this recent results, the situation in infinite dimension is not completely clear, and this is still an interesting field for investigation.

In settings for which the estimate is convergent, there is still the question of the rate of convergence, and how to choose the parameter k in order to achive the best rate of convergence. As noticed by Kulkarni and Posner [57], the rate of convergence of the nearest neighbors is closely related to the notion of entropy, introduced in the late fifties by Kolmogorov and Tikhomirov [56]. These tools are to be used to study cases and algorithm refinements that are not yet to be found in the literature.

4. Application Domains

4.1. Localisation, navigation and tracking

See 5.1, 6.11, 7.2, 7.4 and 7.5.

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 [52], [45], 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 [51], [68], 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 positioning, navigation and tracking, such as geolocalisation and tracking in a wireless network, terrain–aided navigation, see 7.2, and data fusion for indoor localisation, see 7.5.

4.2. Rare event simulation

See 3.4, 6.3, 6.4, 6.5, 6.6, 6.9, 7.1, 7.3 and 7.7.

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 unefficient to estimate accurately very small probabilities. Besides importance sampling, an alternate widespread technique consists in multilevel splitting [60], 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 to several applications of multilevel splitting for rare event simulation, such as risk assessment in air traffic management, see 7.1, detection in sensor networks, see 7.3, protection of digital documents, see 7.7, and credit risk estimation.

5. Software

5.1. Demos

Participant: François Le Gland [corresponding person].

See 4.1, 7.2 and 7.5.

To illustrate that particle filtering algorithms are efficient, easy to implement, and extremely visual and intuitive by nature, for localisation, navigation and tracking problems in complex environments, with geometrical constraints, that would be very difficult to solve with usual Kalman filters. This material has proved very useful in training sessions and seminars that have been organized in response to the demand from industrial partners (SAGEM, CNES and EDF), and also in teaching. At the moment, the following three demos are available

Terrain–aided navigation of an aircraft Inertial position and velocity estimates are known to drift away from their true values, and need to be combined with some external source of information. In this demo, noisy measurements of the terrain height below an aircraft are obtained as the difference between (i) the aircraft altitude above the sea level (provided by a pression sensor) and (ii) the aircraft altitude above the terrain (provided by an altimetric radar), and are compared to the terrain height in any possible point (read on the elevation map). A cloud (swarm) of particles explores various possible trajectories generated from inertial navigation estimates and from a model of inertial navigation errors, and are replicated or discarded depending on whether the terrain height below the particle (i.e. at the same horizontal position) matches or not the available noisy measurement of the terrain height below the aircraft.

Positioning and tracking in the presence of obstacles In this demo, several stations cooperate to locate and track a mobile from noisy angle measurements, in the presence of obstacles (walls, tunnels, etc), which make the mobile temporarily invisible from one or several stations.

Indoor navigation of a mobile robot In this demo, a mobile robot is finding its way inside a building, a digital map of which (including walls, doorways, etc.) is provided. The initial position, velocity and orientation of the robot are unknown, and noisy measurements of its rotation and linear displacement are given by an odometer. In addition, a ring of laser sensors detects with some error the distance from the robot to obstacles in sixteen different directions. A cloud (swarm) of particles explores various possible trajectories generated from odometer navigation estimates and from a model of odometer navigation errors, and are replicated or discarded depending on whether the distance from the particle to obstacles matches or not the available noisy measurement of the distance from the robot to the obstacles, in all sixteen directions, and depending also on whether the generated trajectories are compatible with the presence of obstacles.

6. New Results

6.1. Rate of convergence of the bagged nearest neighbor estimate

Participants: Frédéric Cérou, Arnaud Guyader.

See 3.5

This is a collaboration with Gérard Biau, from université Pierre et Marie Curie, ENS Paris and INRIA Paris Rocquencourt (project–team CLASSIC).

Bagging is a simple way to combine estimates in order to improve their performance. This method, suggested by Breiman in 1996, proceeds by resampling from the original data set, constructing a predictor from each subsample, and decide by combining. By bagging an n-sample, the crude nearest neighbor regression estimate is turned into a consistent weighted nearest neighbor regression estimate, which is amenable to statistical analysis. Letting the resampling size k_n grows with n in such a manner that $k_n \to \infty$ and $k_n/n \to 0$, we have shown that this estimate achieves optimal rate of convergence, independently from the fact that resampling is done with or without replacement. Since the estimate with the optimal rate of convergence depends on the unknown distribution of the observations, adaptation results by data–splitting are also obtained [12], [24].

6.2. Rates of convergence of the functional k–nearest neighbor estimate

Participants: Frédéric Cérou, Arnaud Guyader.

See 3.5

This is a collaboration with Gérard Biau, from université Pierre et Marie Curie, ENS Paris and INRIA Paris Rocquencourt (project–team CLASSIC).

Motivated by a broad range of potential applications, such as regression on curves, we investigate rates of convergence of the k-nearest neighbor estimate of the regression function, based on N independent copies of the pair (X, Y), when X is in a suitable ball in some functional space. Using compact embedding theory, we present explicit and general finite sample bounds on the expected squared difference between k-nearest neighbor estimator and Bayes regression function, in a very general setting. We also particularize our results 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 [13], [25].

6.3. Estimating the minimal length of Tardos code

Participants: Frédéric Cérou, Arnaud Guyader.

See 3.4, 4.2 and 7.7.

This is a collaboration with Luis Pérez–Feire from GRADIANT (Galician R&D Center in Advanced Telecommunications, Vigo, Spain) and with Teddy Furon from INRIA Rennes Bretagne Atlantique (project–team TEMICS).

We are interested in the minimal length of a binary probabilistic traitor tracing code. We consider the code construction proposed by Gábor Tardos in 2003, with the symmetric accusation function as improved by Boris Škorić et al. The length estimation is based on two pillars. First, we consider the worst case attack that a group of c colluders can lead. This attack minimizes the mutual information between the code sequence of a colluder and the pirated sequence. Second, an algorithm pertaining to the field of rare event analysis is constructed in order to estimate the probabilities of error: the probability that an innocent user is framed, and the probabilities that all colluders are missed. Therefore, for a given collusion size, we are able to estimate the minimal length of the code satisfying some error probabilities constraints [21]. This estimation is far lower than the known lower bounds.

6.4. Estimating the probability of false alarm for a zero–bit watermarking technique

Participants: Frédéric Cérou, Arnaud Guyader, Cyrille Jégourel.

See 3.4, 4.2 and 7.7.

This is a collaboration with Teddy Furon from INRIA Rennes Bretagne Atlantique (project–team TEMICS).

Assessing that a probability of false alarm is below a given significance level is a crucial issue in watermarking. We have proposed an iterative and self–adapting algorithm which estimates very low probabilities of error. Some experimental investigations validates its performance for a rare detection scenario where there exists a close form formula of the probability of false alarm. Our algorithm appears to be much quicker and more accurate than a classical Monte Carlo estimator. It even allows the experimental measurement of error exponents [17].

6.5. Credit risk estimation

Participants: Frédéric Cérou, Arnaud Guyader.

See 3.4 and 4.2.

We applied our rare event techniques to the Credit Metrics model in order to estimate the probability of large loss. We found that at least on this Gaussian model, our methods compares well with state of the art importance sampling techniques, as proposed by Glasserman et al. [50]. We are now investigating other models of multifactor portfolio credit risk, as our method is versatile, and can be easily adapted to new models. This is still a work in progress. Preliminary investigations have been made during the internship of Hang Khuc (université de Rennes 2).

6.6. Rare event techniques for integral estimation

Participants: Frédéric Cérou, Arnaud Guyader, Jean-Baptiste Le Meur.

See 3.4 and 4.2.

This is a collaboration with Éric Matzner–Løber, from université de Rennes 2.

We applied our rare event techniques to the estimation of integrals that are too complicated to be computed by either numerical methods, or Monte Carlo. Typically, one can think of the partition function of a Gibbs measure. Very preliminary results indicate that the proposed method may prove very useful for some difficult cases. This is still a work in progress.

6.7. Particle methods with adaptive resampling

Participant: François Le Gland.

This is a collaboration with Élise Arnaud, from université Joseph Fourier and INRIA Grenoble Rhône Alpes (project–team PERCEPTION).

A longstanding problem in particle or sequential Monte Carlo (SMC) methods is to mathematically prove the popular belief that resampling does improve the performance of the estimation (this of course is not always true, and the real question is to clarify classes of problems where resampling helps). A more pragmatic answer to the problem is to use adaptive procedures that have been proposed on the basis of heuristic considerations, where resampling is performed only when it is felt necessary, i.e. when some criterion (effective number of particles, entropy of the sample, etc.) reaches some prescribed threshold. It still remains to mathematically prove the efficiency of such adaptive procedures. Our first contribution has been to consider a design where resampling is performed at some intermediate fixed time instants, and to optimize the asymptotic variance of the estimation error w.r.t. the resampling time instants. The second contribution [15] has been to prove a central limit theorem for particle methods with adaptive resampling, using an interpretation of particle methods where importance weights are interpreted as particles [62], as long as they are not used for resampling purpose, and to minimize the asymptotic variance w.r.t. the threshold.

6.8. Non–parametric learning of the optimal importance distribution

Participant: François Le Gland.

This is a collaboration with Nadia Oudjane, from EDF R&D Clamart.

Evaluating an integral or a mathematical expectation of a nonnegative function can always be seen as computing the normalization constant in a Boltzmann–Gibbs probability distribution. When the probability distribution and the nonnegative function do not agree, i.e. have significant contributions in different parts of the integration space, then the variance of the estimator can be very large, and one should use another importance distribution, ideally the optimal (zero variance) importance distribution μ^* which unfortunately cannot be used since it depends on the desired (but unknown) integral. Alternatively, sequential methods have been designed (under different names, such as annealed sampling, progressive correction, multilevel splitting, etc., depending on the context) which not only provide an expression for the desired integral as the product of intermediate normalization constants, but ultimately provide as well an N–sample approximately distributed according to the optimal importance distribution. From the weighted empirical probability distribution associated with this sample, a regularized probability distribution μ_N can be obtained, using a kernel method or a simple histogram, and can be used as an almost optimal importance distribution to estimate the original integral with a

M–sample distributed according to μ_N. The variance of the resulting estimator depends on the product of the inverse sample size $1/M$ by the χ^2–distance between the almost optimal importance distribution μ_N and the optimal (zero variance) importance distribution μ^*.

Our contribution has been to provide an estimate of this χ^2–distance, under mild assumptions. The impact of dimension on density estimation is a limiting factor here, but the variance reduction is very significant in moderate dimensions.

6.9. Estimation of extreme quantiles and isoquantile curves

Participants: François Le Gland, Rudy Pastel.

See 3.4 and 4.2.

This is a collaboration with Jérôme Morio, from ONERA Châtillon.

Many reliability or hedging problems reduce to quantile estimation, and increasing safety standards creates a need for extreme quantile estimation. Adaptive multilevel splitting methods are studied in this context, which significantly reduce the variance of the relative error. A multilevel definition of quantile is proposed, under the name of isoquantile curves and is compared with the classical definition in terms of circular error probables. Its main advantage is a better use of the angular distribution, as illustrated in an example regarding the impact location of a ballistic launcher.

6.10. Path planning for navigation in a geographic information system

Participants: Jean–Pierre Le Cadre, Francis Céleste.

The problem we considered is the optimization of the navigation of an intelligent mobile in a real world environment, described by a map. The map is composed of features representing natural landmarks in the environment. The vehicle is equipped with sensors which allows it to obtain landmark parameter estimates. These measurements are correlated with the map so as to estimate the mobile position. The optimal trajectory must be designed in order to control a measure of the performance for the filtering algorithm used for mobile navigation. As the mobile state and the measurements are random, a well–suited measure can be a functional of the posterior Cramer–Rao bound (PCRB). In many applications, it is crucial to be able to estimate accurately the state of the mobile during the execution of the plan. So it is necessary to couple the planning and the execution stages.

A classical tool is the constrained Markov decision process (MDP) framework. However, our optimality criterion is based on the posterior Cramer–Rao bound, and the nature of the objective function for path planning makes it impossible to perform complete optimization within the MDP framework. Indeed, the reward in one stage of our MDP depends on all the history of the trajectory. To overcome this problem, the cross–entropy method [69], [34], originally used for rare–event simulation, is a valuable tool. Its principle is to translate a classical optimization method into an associated stochastic problem and then to solve it adaptively as the simulation of rare events. This approach has been tested on various (simple) geographic environments and performs satisfactorily. Our contribution, conducted in the context of Francis Céleste PhD work, has been devoted to the derivation of closed–form approximations of the information we can gain from an elementary motion. Using them, it is possible to immerse the problem within an optimal control framework and to use efficiently the maximum principle.

6.11. Extracting information from a video sensor network

Participants: Jean–Pierre Le Cadre, Adrien Ickowicz.

See 4.1 and 7.4.

Recent trends lead to consider globally networks of (video) sensors. These networks can be relatively large, so we have to face specific problems. How can we use the data at the sensor level, how to represent the information collected at the sensor level, how to fuse ? A first step consists in extracting spatio–temporal informations from video sensors. Of course, these sensors are generally uncalibrated and asynchronous. So we have to consider rather rough informations. Roughness can go up to reduce the information to proximity and to a binary information about object motion (closing or not). A first step has been to consider the use of the estimated closest point of approach (CPA) times for estimating the parameters of the target trajectories. This study is relatively simple but has the great advantage to put in evidence the basic requirements and the limits of this approach. In a second step, we considered the estimation of the CPA times from a sequence of images.

The limits of the above approach are quite obvious. If it is able to exploit a temporal contrast, there is a strong need to use a spatio–temporal contrast at the (binary) sensor network level. Actually, it has been shown that the separation problem we have to solve present strong similarities with the optimization problems we have to solve in a SVM context. The benefits of this approach are multiple: it is well adapted to (robust) tracking and the combinatorial problems which plagued multitarget tracking are fundamentally reduced. For the tracking step, particle filtering is the natural way since it can easily include complex priors, non–linear measurements as well as separation properties, within a hierarchical context. Our contribution this year has been a first step towards multiple target tracking within this framework [18], [20].

7. Contracts and Grants with Industry

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

Participant: François Le Gland.

See 3.4 and 4.2.

INRIA contract ALLOC 2399 — May 2007 to August 2010.

This FP6 project is coordinated by National Aerospace Laboratory (NLR) (The Netherlands). The academic partners are University of Cambridge and University of Leicester (United Kingdom), Politecnico di Milano and Universita dell'Aquila (Italy), University of Twente (The Netherlands), ETH Zürich (Switzerland), University of Tartu (Estonia), National Technical University of Athens (NTUA) and Athens University of Economics and Business (Greece), Direction des Services de la Navigation Aérienne (DSNA), École Nationale de l'Aviation Civile (ENAC), Eurocontrol Experimental Center (EEC) and INRIA Rennes Bretagne Atlantique (France), and the industrial partners are Honeywell (Czech Republic), Isdefe (Spain), Dedale (France), NATS En Route Ltd. (United Kingdom).

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 [30] suited for hybrid state space (resampling per mode, marginalization) are currently investigated.

7.2. Terrain–aided navigation — contract with Thalès Communications

Participants: François Le Gland, Nordine El Baraka.

See 4.1 and 5.1.

INRIA contract ALLOC 2857 — September 2007 to August 2010.

This collaboration with Thalès Communications is supported by DGA (Délégation Générale à l'Armement) and is related with the supervision of the CIFRE thesis of Nordine El Baraka.

The overall objective is to study innovative algorithms for terrain–aided navigation, and to demonstrate these algorithms on four different situations involving different platforms, inertial navigation units, sensors and georeferenced databases. The thesis also considers the special use of image sensors (optical, infra–red, radar, sonar, etc.) for navigation tasks, based on correlation between the observed image sequence and a reference image available on–board in the database.

Marginalized particle filters [70] and regularized particle filters [10] have been implemented, and several propositions have been studied to adapt the sample size, such as KLD–sampling [44], which could be useful in the case of a poor initial information, or if the platform flies over a poorly informative area. Besides particle methods, which are proposed as the basic navigation algorithm, simpler algorithms such as the extended Kalman filter (EKF) or the unscented Kalman filter (UKF) have also been investigated.

7.3. Optimization of sensors location and activation — contract with DGA / CTSN (centre technique des systèmes navals)

Participants: Jean–Pierre Le Cadre, François Le Gland.

See 3.4 and 4.2

INRIA contract ALLOC 4233 — April 2009 to March 2011.

This is a collaboration with Sébastien Paris, from 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 [69], [34].

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.

7.4. Tracking in a binary sensor network — contract with DGA / CEP (centre d'expertise parisien)

Participants: Jean–Pierre Le Cadre, Adrien Ickowicz, Nicolas Ramin.

See 4.1 and 6.11.

INRIA contract ALLOC 2338 — April 2007 to October 2009.

This contract deals with surveillance of large zones via a network of video sensors. Of course, sensor outputs can be treated in a centralized architecture. However, centralized architectures suffer from serious drawbacks. Communication constraints (e.g. bandwidth) are frequently evoked, but still more fundamentally we have to face many problems inherent of this architecture, like:

- sensor calibration, positioning and synchronization,
- false alarms, multiple objects, occlusions, etc.

Overall, there is a strong need for extracting a global picture at the network level. This means that we have to focus on the level of information we can extract at the sensor level and how to fuse them. Work has been done on the first point, using both simulated and real video sequences. The less informative level of information is the binary one. However, there is a fundamental difference between a $\{0, 1\}$ information and $\{-, +\}$ information. A general $\{0, 1\}$ information corresponds to a detection / non–detection information. Such architecture has been widely studied in a distributed detection framework, but is not well suited to our context. However $\{0, 1\}$ information is especially interesting if the detection process includes geographic constraints like proximity, field–of–view, etc. The $\{-, +\}$ information corresponds to a motion information: the object gets closer or is going far away. At the network level, this is a very rich information which can present definite advantages (robustness, multi–target tracking). However, its interest depends on the network density. So, it is also necessary to consider various and complementary decentralized architectures according to the sensing capabilities, the target behaviors and, overall, the combinatorial complexity of the problem. Work has been done for defining processings and architectures adapted to this context.

Work has been devoted to the processing of image sequences for extracting the binary information. After considering the divergence of a local motion model, its estimation and use on real data, we turned toward a temporal analysis of the bearing information. More precisely, bearing rates and bearing rate changes give us an estimate of the (local) target behavior. It is thus possible to derive a local estimate of the ratio $\dot{r}/r$, through purely passive measurements, at the sensor level. Although this analysis is purely local, it performs satisfactorily on simulated sequences.

Concerning the information processing at the sensor level, work has been done in three directions. First, the $\{-, +\}$ informations can be summarized via the times of CPA (closest point approach) on the various sensors. A complete analysis of the target motion analysis has been done in this framework, with various models of target motion. Then, we turned toward the analysis of the spatio–temporal $\{-, +\}$ informations. An original method based on a separation principle allows us to obtain an estimate of the target motion parameters via multiple SVM. Second, we considered random target trajectories and target tracking. To that aim, a specific method based on multiple corrections has been developed.

7.5. Information fusion for localisation (FIL) — ANR Télécommunications

Participants: François Le Gland, Liyun He.

See 4.1 and 5.1.

INRIA contract ALLOC 2856 — January 2008 to December 2010.

This ANR project is 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). Industrial partners are Microtec and Silicom.

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 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 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 [65], 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 [44] 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 [43], 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 [16], [19] for a description of the overall fusion algorithm and an illustration with simulation results.

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

Participants: François Le Gland, Sébastien Beyou, Vu–Duc Tran.

INRIA contract ALLOC 3767 — January 2009 to December 2011.

This ANR project is coordinated by École Normale Supérieure, Paris. The other partner is Météo–France. This is a collaboration with Étienne Mémin from INRIA Rennes Bretagne Atlantique (project–team FLUMINANCE) and with Anne Cuzol and Valérie Monbet from université de Bretagne Sud in Vannes.

The contribution of ASPI to this project is to continue the comparison [11], [23] of sequential data assimilation methods, 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.

7.7. Security and reliability in digital watermarking (NEBBIANO) — ANR Sécurité

Participants: Frédéric Cérou, Arnaud Guyader, Cyrille Jégourel, Fida El Haje Hussein.

See 3.4, 4.2, 6.3 and 6.4.

INRIA contract ALLOC 2229 — January 2007 to December 2009.

Arnaud Guyader is coordinator of this ANR project. This is a collaboration with Teddy Furon from INRIA Rennes Bretagne Atlantique (project–team TEMICS) and Pierre Del Moral from INRIA Bordeaux Sud–Ouest (project–team ALEA). Besides these project–teams, the other partner is LIS–INPG in Grenoble.

There are mainly two strategic axes in NEBBIANO: watermarking and independent component analysis, and watermarking and rare event simulations. To protect copyright owners, user identifiers are embedded in purchased content such as music or movie. This is basically what we mean by watermarking. This watermarking is to be "invisible" to the standard user, and as difficult to find as possible. When content is found in an illegal place (e.g. a P2P network), the right holders decode the hidden message, find a serial number, and thus they can trace the traitor, i.e. the client who has illegally broadcast their copy. However, the task is not that simple as dishonest users might collude. For security reasons, anti–collusion codes have to be employed. Yet, these solutions (also called weak traceability codes) have a non–zero probability of error defined as the probability of accusing an innocent. This probability should be, of course, extremely low, but it is also a very sensitive parameter: anti–collusion codes get longer (in terms of the number of bits to be hidden in content) as the probability of error decreases. Fingerprint designers have to strike a trade–off, which is hard to conceive when only rough estimation of the probability of error is known. The major issue for fingerprinting algorithms is the fact that embedding large sequences implies also assessing reliability on a huge amount of data which may be practically unachievable without using rare event analysis. Our task within this project is to adapt our

methods for estimating rare event probabilities to this framework, and provide watermarking designers with much more accurate false detection probabilities than the bounds currently found in the literature. We have already applied these ideas to some randomized watermarking schemes and obtained much sharper estimates of the probability of accusing an innocent.

A patent *"Validation de schémas de verrous numériques en watermarking et fingerprinting"* has been submitted by INRIA and by université de Rennes 2.

7.8. Coherent detection for QPSK 40Gb systems (COHDEQ40) — ANR Télécommunications

Participant: Frédéric Cérou.

INRIA contract ALLOC 2205 — December 2006 to November 2009.

This ANR project is coordinated by Alcatel–Lucent. The other partners are Alcatel Thales III–V Lab, INT Évry, INRIA Rennes Bretagne Atlantique (project–teams ASPI and TEMICS), Kylia, Photline and XLIM (université de Limoges).

The project COHDEQ40 intends to demonstrate the potential of coherent detection associated with digital signal processing for the next generation high density 40Gb/s WDM systems optimized for transparency and flexibility. Key integrated optoelectronics components and specific algorithms will be developed and system evaluation performed. The INRIA task is to develop these signal processing algorithms needed to recover the message on the decoder side. This makes full use of our knowledge of equalization and synchronization techniques involved in digital communications.

A patent *"A decision directed algorithm for adjusting a polarization demultiplexer in a coherent detection optical receivers"* has been jointly submitted in September by Alcatel Lucent and by INRIA.

7.9. Real–time adaptive coherent heterodyne detector for future optical netowrks (TCHATER) — ANR Télécommunications

Participant: Frédéric Cérou.

INRIA contract ALLOC 2801 — January 2008 to December 2010.

This ANR project is coordinated by Alcatel–Lucent. The other partners are E2V, TELECOM ParisTech, LIP (ENS Lyon).

The primary goal of the TCHATER project is to demonstrate a coherent terminal operating at 40Gb/s using *real–time* digital signal processing and efficient polarization division multiplexing. The terminal will benefit to next-generation high information-spectral density optical networks, while offering straightforward compatibility with current 10Gbit/s networks. It will require that advanced high–speed electronic components, especially analog–to–digital converters, are designed within the project. Specific algorithms for polarisation demultiplexing and forward error correction with soft decoding will also have to be developed.

7.10. 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 coordinated by Alcatel–Lucent Bell Labs France. The other partners are Draka, Kylia, Université Lille 1, Institut Télécom / Télécom & Management Sud Paris

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.

8. Dissemination

8.1. Scientific animation

Arnaud Guyader and Frédéric Cérou have organized, with Éric Matzner–Løber from université de Rennes 2, the workshop JSTAR (Journées de Statistique de Rennes) in October 2009.

François Le Gland has reported on the PhD thesis of Benoît Landelle (université Paris Sud, advisor: Élisabeth Gassiat). He was also a member of the committee and co–advisor for the PhD thesis of Vu–Duc Tran (université de Bretagne Sud, Vannes, co–advisor: Valérie Monbet).

Arnaud Guyader is a member of the "comité de sélection" in applied mathematics (section 26) of université d'Angers. François Le Gland is a member of the "comité de sélection" in mathematics (sections 25–26) of INSA (institut national de sciences appliquées) Rennes, and he is a member of the "conseil d'UFR" of the department of mathematics of université de Rennes 1.

8.2. Teaching

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, 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.

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

8.3. 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 at École Polytechnique in the working group on rare events, in March 2009. He has also given a talk on the rate of convergence for the bagged nearest neighbor estimators during the annual meeting of the SFdS (Société Française de Statistique) in Bordeaux in May 2009.

François Le Gland has given a talk on the multilevel splitting approach to rare event simulation at the workshop on rare events held at ONERA Châtillon in December 2009. He has also given an introductory talk on applications of particle methods at a meeting held at EDF R&D in Clamart in December 2009.

Vu–Duc Tran has defended her PhD thesis on the asymptotic properties of the ensemble Kalman filter, in Vannes in June 2009.

9. Bibliography

Major publications by the team in recent years

[1] F. CAMPILLO, Y. A. KUTOYANTS, F. LE GLAND. *Small noise asymptotics of the GLR test for off–line change detection in misspecified diffusion processes*, in "Stochastics and Stochastics Reports", vol. 70, n⁰ 1–2, 2000, p. 109–129.

[2] F. CÉROU. *Long time behaviour for some dynamical noise–free nonlinear filtering problems*, in "SIAM Journal on Control and Optimization", vol. 38, n⁰ 4, 2000, p. 1086–1101.

[3] F. CÉROU, A. GUYADER. *Nearest neighbor classification in infinite dimension*, in "ESAIM : Probability and Statistics", vol. 10, 2006, p. 340–355.

[4] F. CÉROU, A. GUYADER. *Adaptive multilevel splitting for rare event analysis*, in "Stochastic Analysis and Applications", vol. 25, n⁰ 2, March 2007, p. 417–443.

[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", vol. 1, 2006, p. 181–203 (electronic), Paper 01–08.

[6] M. JOANNIDES, F. LE GLAND. *Small noise asymptotics of the Bayesian estimator in nonidentifiable models*, in "Statistical Inference for Stochastic Processes", vol. 5, n⁰ 1, 2002, p. 95–130.

[7] F. LE GLAND, L. MEVEL. *Exponential forgetting and geometric ergodicity in hidden Markov models*, in "Mathematics of Control, Signals, and Systems", vol. 13, n⁰ 1, 2000, p. 63–93.

[8] 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", vol. 106, n⁰ 2, August 2003, p. 279-316.

[9] 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", vol. 14, n⁰ 1, February 2004, p. 144–187.

[10] 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, chap. 12, Springer–Verlag, 2001, p. 247–271.

Year Publications

Doctoral Dissertations and Habilitation Theses

[11] 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/, Ph. D. Thesis.

Articles in International Peer-Reviewed Journal

[12] G. BIAU, F. CÉROU, A. GUYADER. *On the rate of convergence of the bagged nearest neighbor estimate*, in "Journal of Machine Learning Research", 2010, to appear.

[13] G. BIAU, F. CÉROU, A. GUYADER. *Rates of convergence of the functional k–nearest neighbor estimator*, in "IEEE Transactions on Information Theory", 2010, to appear.

[14] 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", 2010, to appear.

International Peer-Reviewed Conference/Proceedings

[15] É. ARNAUD, F. LE GLAND. *SMC with adaptive resampling : Large sample asymptotics*, in "Proceedings of the 2009 Workshop on Statistical Signal Processing, Cardiff 2009", IEEE–SPS, September 2009, p. 481–484.

[16] 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.

[17] F. CÉROU, T. FURON, A. GUYADER, C. JÉGOUREL. *Estimating the probability of false alarm for a zero–bit watermarking technique*, in "16th International Conference on Digital Signal Processing, Santorin 2009", July 2009.

[18] A. ICKOWICZ, J.–P. LE CADRE. *Bi–target tracking within a binary sensor network*, in "Proceedings of the 12th International Conference on Information Fusion, Seattle 2009", ISIF, July 2009.

[19] 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.

National Peer-Reviewed Conference/Proceedings

[20] A. ICKOWICZ, J.–P. LE CADRE. *Suivi de cibles à l'aide d'un réseau de capteurs*, in "Actes du Colloque GRETSI, Dijon 2009", September 2009.

Scientific Books (or Scientific Book chapters)

[21] T. FURON, L. PÉREZ–FREIRE, A. GUYADER, F. CÉROU. *Estimating the minimal length of Tardos codes*, in "11th International Workshop on Information Hiding, Darmstadt 2009", S. KATZENBEISSER, A.–R. SADEGHI (editors), Lecture Notes in Computer Science, n$^{\text{o}}$ 5806, Springer, June 2009, p. 176–190.

[22] P. L'ÉCUYER, F. LE GLAND, P. LEZAUD, B. TUFFIN. *Splitting methods*, in "Monte Carlo Methods for Rare Event Analysis, Chichester", G. RUBINO, B. TUFFIN (editors), chap. 3, John Wiley & Sons, 2009, p. 39–61.

[23] F. LE GLAND, V. MONBET, V.–D. TRAN. *Large sample asymptotics for the ensemble Kalman filter*, in "Handbook on Nonlinear Filtering, Oxford", D. O. CRISAN, B. L. ROZOVSKII (editors), Oxford University Press, 2010, to appear.

Research Reports

[24] G. BIAU, F. CÉROU, A. GUYADER. *On the rate of convergence of the bagged nearest neighbor estimate*, n° 6860, INRIA, February 2009, http://hal.inria.fr/inria-00363875/en/, Rapport de Recherche.

[25] G. BIAU, F. CÉROU, A. GUYADER. *On the rate of convergence of the functional k-NN estimates*, n° 6861, INRIA, February 2009, http://hal.inria.fr/inria-00364555/en/, Rapport de Recherche.

[26] F. CÉROU, P. DEL MORAL, T. FURON, A. GUYADER. *Rare event simulation for a static distribution*, n° 6792, INRIA, January 2009, http://hal.inria.fr/inria-00350762/en/, Rapport de Recherche.

[27] F. LE GLAND, V. MONBET, V.–D. TRAN. *Large sample asymptotics for the ensemble Kalman filter*, n° 7014, INRIA, August 2009, http://hal.inria.fr/inria-00409060/en/, Rapport de Recherche.

References in notes

[28] 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.

[29] 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", vol. SP–50, n° 2 (Special issue on Monte Carlo Methods for Statistical Signal Processing), February 2002, p. 174–188.

[30] H. A. P. BLOM, B. BAKKER, J. KRYSTUL. *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), chap. 9, John Wiley & Sons, 2009, p. 193–214.

[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", vol. 95, n° 5, May 2007, 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] T. M. COVER, P. E. HART. *Nearest neighbor pattern classification*, in "IEEE Transactions on Information Theory", vol. IT–13, n° 1, January 1995, p. 21–27.

[34] P.–T. DE BOER, D. P. KROESE, S. MANNOR, R. Y. RUBINSTEIN. *A tutorial on the cross–entropy method*, in "Annals of Operations Research", vol. 134 (Special issue on the Cross-Entropy Method for Combinatorial Optimization, Rare Event Simulation and Neural Computation), n° 1, January 2005, p. 19–67.

[35] P. DEL MORAL. *Feynman–Kac Formulae. Genealogical and Interacting Particle Systems with Applications*, Probability and its Applications, Springer–Verlag, New York, 2004.

[36] 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", vol. 37, n° 2, 2001, p. 155–194.

[37] 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, vol. 1729, Springer–Verlag, 2000, p. 1–145.

[38] R. DOUC, A. GUILLIN, J. NAJIM. *Moderate deviations for particle filtering*, in "The Annals of Applied Probability", vol. 15, n° 1B, February 2005, p. 587–614.

[39] R. DOUC, C. MATIAS. *Asymptotics of the maximum likelihood estimator for general hidden Markov models*, in "Bernoulli", vol. 7, n° 3, June 2001, p. 381–420.

[40] A. DOUCET, S. J. GODSILL, C. ANDRIEU. *On sequential Monte Carlo sampling methods for Bayesian filtering*, in "Statistics and Computing", vol. 10, n° 3, July 2000, p. 197–208.

[41] E. FIX, J. L. HODGES. *Discriminatory analysis. Nonparametric discrimination : Consistency properties*, n° 4, USAF School of Aviation Medicine, Randolph Field, February 1951, Project Number 21–49–004, Technical Report.

[42] E. FIX, J. L. HODGES. *Discriminatory analysis : Small sample performance*, n° 11, USAF School of Aviation Medicine, Randolph Field, August 1952, Project Number 21–49–004, Technical Report.

[43] D. FOX, W. BURGARD, H. KRUPPA, S. THRUN. *A probabilistic approach to collaborative multi–robot localization*, in "Autonomous Robots", vol. 8, n° 3 (Special issue on Heterogeneous Multi–Robot Systems), June 2000, p. 325–344.

[44] D. FOX. *Adapting the sample size in particle filters through KLD–sampling*, in "International Journal of Robotics Research", vol. 22, n° 12, December 2003, p. 985–1004.

[45] D. FOX, J. HIGHTOWER, L. LIAO, D. SCHULZ, G. BORRIELLO. *Bayesian filtering for location estimation*, in "IEEE Pervasive Computing", vol. 2, n° 3, July/September 2003, p. 24–33.

[46] 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, chap. 19, Springer–Verlag, 2001, p. 401–428.

[47] D. FRENKEL, B. SMIT. *Understanding Molecular Simulation. From Algorithms to Applications*, Computational Science Series, 2nd, vol. 1, Academic Press, San Diego, 2002.

[48] P. GLASSERMAN. *Monte Carlo Methods in Financial Engineering*, Applications of Mathematics, vol. 53, Springer–Verlag, New York, 2004.

[49] P. GLASSERMAN, P. HEIDELBERGER, P. SHAHABUDDIN, T. ZAJIC. *Multilevel splitting for estimating rare event probabilities*, in "Operations Research", vol. 47, n° 4, July–August 1999, p. 585–600.

[50] P. GLASSERMAN, W. KANG, P. SHAHABUDDIN. *Fast simulation of multifactor portfolio credit risk*, in "Operations Research", vol. 56, n° 5, September–October 2008, p. 1200–1217.

[51] N. J. GORDON, D. J. SALMOND, A. F. M. SMITH. *Novel approach to nonlinear / non–Gaussian Bayesian state estimation*, in "IEE Proceedings, Part F", vol. 140, n° 2, April 1993, p. 107–113.

[52] F. GUSTAFSSON, F. GUNNARSSON, N. BERGMAN, U. FORSSELL, J. JANSSON, R. KARLSSON, P.–J. NORDLUND. *Particle filters for positioning, navigation, and tracking*, in "IEEE Transactions on Signal Processing", vol. SP–50, nº 2 (Special issue on Monte Carlo Methods for Statistical Signal Processing), February 2002, p. 425–437.

[53] M. ISARD, A. BLAKE. CONDENSATION — *Conditional density propagation for visual tracking*, in "International Journal of Computer Vision", vol. 29, nº 1, August 1998, p. 5–28.

[54] M. R. JAMES, F. LE GLAND. *Consistent parameter estimation for partially observed diffusions with small noise*, in "Applied Mathematics & Optimization", vol. 32, nº 1, July/August 1995, p. 47–72.

[55] G. KITAGAWA. *Monte Carlo filter and smoother for non–Gaussian nonlinear state space models*, in "Journal of Computational and Graphical Statistics", vol. 5, nº 1, 1996, p. 1–25.

[56] A. N. KOLMOGOROV, V. M. TIKHOMIROV. ε–*entropy and* ε–*capacity of sets in functional spaces*, in "Twelve papers on Algebra and Real Functions, Providence", AMS Translations — Series 2, vol. 17, AMS, 1961, p. 277–364.

[57] S. R. KULKARNI, S. E. POSNER. *Rates of convergence of nearest neighbor estimation under arbitrary sampling*, in "IEEE Transactions on Information Theory", vol. IT–41, nº 4, July 1995, p. 1028–1039.

[58] Y. A. KUTOYANTS. *Identification of Dynamical Systems with Small Noise*, Mathematics and its Applications, vol. 300, Kluwer Academic Publisher, Dordrecht, 1994.

[59] H. R. KÜNSCH. *Recursive Monte Carlo filters : Algorithms and theoretical analysis*, in "The Annals of Statistics", vol. 33, nº 5, October 2005, p. 1983–2021.

[60] P. L'ÉCUYER, V. DEMERS, B. TUFFIN. *Rare events, splitting, and quasi–Monte Carlo*, in "ACM Transactions on Modeling and Computer Simulation", vol. 17, nº 2 (Special issue honoring Perwez Shahabuddin), April 2007, Article 9.

[61] L. LE CAM. *Asymptotic Methods in Statistical Decision Theory*, Springer Series in Statistics, Springer–Verlag, New York, 1986.

[62] F. LE GLAND. *Combined use of importance weights and resampling weights in sequential Monte Carlo methods*, in "ESAIM : Proceedings", vol. 19, 2007, p. 85–100 (electronic).

[63] 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, nº 337, Springer–Verlag, 2006, p. 351–389.

[64] 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, nº 280, Springer–Verlag, 2002, p. 267–282.

[65] 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.

[66] J. S. LIU. *Monte Carlo Strategies in Scientific Computing*, Springer Series in Statistics, Springer–Verlag, New York, 2001.

[67] D. PREISS. *Gaussian measures and the density theorem*, in "Commentationes MathematicæUniversitatis Carolinæ", vol. 22, n$^{\text{o}}$ 1, 1981, p. 181–193.

[68] B. RISTIĆ, M. S. ARULAMPALAM, N. J. GORDON. *Beyond the Kalman Filter : Particle Filters for Tracking Applications*, Artech House, Norwood, MA, 2004.

[69] 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.

[70] T. SCHÖN, F. GUSTAFSSON, P.–J. NORDLUND. *Marginalized particle filters for mixed linear / nonlinear state–space models*, in "IEEE Transactions on Signal Processing", vol. SP–53, n$^{\text{o}}$ 7, July 2005, p. 2279–2289.

[71] C. J. STONE. *Consistent nonparametric regression (with discussion)*, in "The Annals of Statistics", vol. 5, n$^{\text{o}}$ 4, July 1977, p. 595–645.

[72] S. THRUN, W. BURGARD, D. FOX. *Probabilistic Robotics*, Intelligent Robotics and Autonomous Agents, The MIT Press, Cambridge, MA, 2005.

[73] A. W. VAN DER VAART, J. A. WELLNER. *Weak Convergence and Empirical Processes*, Springer Series in Statistics, Springer–Verlag, Berlin, 1996.