IN PARTNERSHIP WITH:
CNRS

Université de Bordeaux

Activity Report 2012

Project-Team ALEA

Advanced Learning Evolutionary Algorithms

IN COLLABORATION WITH: Institut de Mathématiques de Bordeaux (IMB)

RESEARCH CENTER
Bordeaux - Sud-Ouest

THEME
Stochastic Methods and Models

Table of contents

Project-Team ALEA

Keywords: Statistical Methods, Machine Learning, Stochastic Methods, Optimization, Inverse Problem

Creation of the Project-Team: March 01, 2009 , *Updated into Project-Team:* January 01, 2010 .

1. Members

Research Scientists

Pierre Del Moral [Inria, Team leader, Senior Researcher (DR), HdR]
François Caron [Inria, Junior Researcher (CR)]

Faculty Members

Bernard Bercu [University Bordeaux 1, Professor (PR), HdR]
Pierrick Legrand [University Bordeaux 2, Associate Professor (MdC)]
Adrien Richou [University Bordeaux 1, Associate Professor (MdC)]

External Collaborators

Arnaud Doucet [University of Oxford]
Josselin Garnier [University Paris 7]
Pierre Minvielle [CEA]

Engineers

Adrien Todeschini [Inria, funding from Inria ADT BiiPS]
Denis Arrivault [Inria, funding from ANR Propagation, Dec. 2010 - Oct. 2012]

PhD Students

Aurélie Le Cain [CEA, CIFRE, Sept. 2008 - Jan. 2012]
François Giraud [CEA, CIFRE, since Nov. 2009]
Peng Hu [University of Bordeaux, funding from French Research Ministry, Oct. 2009 - June 2012]
Philippe Fraysse [University of Bordeaux, funding from ENS Cachan, since Sept. 2010]
Paul Lemaître [EDF, CIFRE, since Sept. 2010]
Frédéric Proia [Inria, funding from Région Aquitaine and Inria, since Dec. 2010]
Nicolas Antunes [University of Bordeaux, funding from French Research Ministry, since Sept. 2011]
Vassili Blandin [University of Bordeaux, funding from ENS Cachan, since Sept. 2011]

Administrative Assistant

Nicolas Jahier [Team Assistant]

2. Overall Objectives

2.1. Overall Objectives

In recent years, a new generation of numerical algorithms have begun to spread through the scientific community. Surprisingly enough, up to a few exceptions, many of these modern ideas do not really come from physics, but from biology and ethology. In a growing number of scientific disciplines, the researchers are now interpreting real world processes and engineering type systems less like purely deterministic and crude clockwork mechanisms, but much more like random and sophisticated biology inspired processes. This new generation of engineering models is based on stochastic ideas and natural principles like : chance, randomness, interactions, reinforcement strategies, exploration rules, biology-inspired adaptation and selection transitions, learning, reproduction, birth and death, ancestral lines and historical processes, genealogical tree evolutions, as well as self-organization principles, and many others.

These biology-inspired stochastic algorithms are often presented as natural heuristic simulation schemes without any mathematical foundations, nor a single performance analysis ensuring their convergence, nor even a single theoretical discussion that clarifies the applicability of these models. An important aspect of our project is to create a concrete bridge between pure and applied probability, statistics, biology, stochastic engineering and computer sciences. This fundamental bridging effort is probably one of the most important key to turn real nature's processes into engineering devices and stochastic algorithms, by learning what can be abstracted, copied or adapted. In the reverse angle, we can mention that these abstracted models adapting nature mechanisms and biological capabilities also provides a better understanding of the real processes.

By essence, *the team-project is not a single application-driven research project* . The reasons are three-fold. Firstly, the same stochastic algorithm is very often used in a variety of application areas. On the other hand every application domain area offers a series of different perspectives that can be used to improve the design and the performances of the algorithms. Last but not least, concrete industrial applications, as well as most of concrete problems arising in biology, physics and chemistry, require a specific attention. In general, we do not use a single class of stochastic algorithm but a broader set of stochastic search algorithms that incorporates facets of nature inspired strategies.

Our research project is centered on two central problems in advanced stochastic engineering: *Bayesian inference and rare event simulation* and more particularly *unsupervised learning, multi-target tracking, data assimilation and forecasting, as well as infection spreads inference.* These important and natural research directions have emerged as logical parts of the team project combined with interdisciplinary approaches well-represented at Bordeaux university campus.

The fundamental and the theoretical aspects of our research project are essentially concerned with the stochastic analysis of the following three classes of biology inspired stochastic algorithms: *branching and interacting particle systems, reinforced random walks and self-interacting processes, random tree based models.* One of our prospective research project is to apply the Bayesian learning methodology and the recent particle filter technology to the design of a new generation of interactive `evolutionary computation` and `stochastic art composition` models.

2.2. Highlights of the Year

The book [33] presents a series of novel results on the design and analysis of new classes of mean field particle models in numerical finance, including particle sensitivity measures, calibration models, and particle option pricing algorithms.

3. Scientific Foundations

3.1. Scientific Foundations

This idea of analyzing nature systems and transferring the underlying principles into stochastic algorithms and technical implementations is one of the central component of the ALEA team project. Adapting nature mechanisms and biological capabilities clearly provides a better understanding of the real processes, and it also improves the performance and the power of engineers devices. Our project is centered on both the understanding of biological processes in terms of mathematical, physical and chemical models, and on the other hand, on the use of these biology inspired stochastic algorithms to solve complex engineering problems.

There is a huge series of virtual interfaces, robotic devices, numerical schemes and stochastic algorithms which were invented mimicking biological processes or simulating natural mechanisms. The terminology *"mimicking or simulating"* doesn't really mean to find an exact copy of natural processes, but *to elaborate the mathematical principles so that they can be abstracted from the original biological or physical model.* In our context, the whole series of evolutionary type principles discussed in previous sections can be abstracted into only three different and natural classes of stochastic algorithms, depending on the nature of the biology-inspired interaction mechanism used in the stochastic evolution model. These three stochastic search models are listed below :

1) *Branching and interacting particle systems (birth and death chains, spatial branching processes, mean-field interaction between generations):*

The first generation of adaptive branching-selection algorithms is very often built on the same genetic type paradigm: When exploring a state space with many particles, we duplicate better fitted individuals at the expense of light particles with poor fitness die. From a computational point of view, we generate a large number of random problem solvers. Each one is then rated according to a fitness or performance function defined by the developer. Mimicking natural selection, an evolutionary algorithm selects the best solvers in each generation and breeds them.

2) *Reinforced random walks and self-interacting chains (reinforced learning strategies, interaction processes with respect to the occupation measure of the past visited sites):*

This type of reinforcement is observed frequently in nature and society, where "beneficial" interactions with the past history tend to be repeated. A new class of historical mean field type interpretation models of reinforced processes were developed by the team project leader in a pair of articles [56], [55]. Self interaction gives the opportunity to build new stochastic search algorithms with the ability to, in a sense, re-initialized their exploration from the past, re-starting from some better fitted initial value already met in the past [57], [58].

3) *Random tree based stochastic exploration models (coalescent and genealogical tree search explorations techniques on path space):*

The last generation of stochastic random tree models is concerned with biology-inspired algorithms on paths and excursions spaces. These genealogical adaptive search algorithms coincide with genetic type particle models in excursion spaces. They have been applied with success in generating the excursion distributions of Markov processes evolving in critical and rare event regimes, as well as in path estimation and related smoothing problems arising in advanced signal processing (cf. [53] and references therein). We underline the fact that the complete mathematical analysis of these random tree models, including their long time behavior, their propagations of chaos properties, as well as their combinatorial structures are far from being completed. This class of genealogical tree based models has been introduced in [54] for solving smoothing problems and more generally Feynman-Kac semigroups on path spaces, see also [52], [53], and references therein.

4. Application Domains

4.1. Application Domains

This short section is only concerned with the list of concrete application domains developed by our team project on Bayesian inference and unsupervised learning, nonlinear filtering and rare event analysis. Most of these application areas result from fruitful collaborations with other national institutes or industries.

Three application domains are directly related to evolutionary computing, particle filtering and Bayesian inference. They are currently investigated by our team project:

1. **Multi-object tracking**. Multi-object tracking deals with the task of estimating the states of a set of moving objects from a set of measurements obtained sequentially. These measurements may either arise from one of the targets or from clutter and the measurement-to-target association is generally unknown. This problem can then be recast as a dynamic clustering one where the clusters are the clutter and the different targets. The targets actually move in time, some targets may appear/disappear over time and the number of targets is generally unknown and time-varying. The ALEA team has been involved in the ANR project Propagation (2010-2012) with DCNS, Thalés and Exavision.
2. **Finance**. The Team ALEA is interested in the design and analysis of new advanced particle methods for option pricing, partial observation problems, and sensitivity measures computation. An international workshop has been jointly organized by ALEA and CMAP (Polytechnique) on this topic in October 2012 (organizers: E. Gobet, P. Del Moral, P. Hu).

3. **Epidemiology**. Our team is interested in the development and analysis of particle mean field models for the calibration and uncertainty propagations in complex kinetic population models. The ALEA team is involved in an interdisciplinary exploratory research project with Laboratory Ecologie & Evolution, and co-organized an international workshop on this topic in 2011.

5. Software

5.1. BiiPS software

BiiPS is a general software, developped by Adrien Todeschini, for Bayesian inference with interacting particle systems, a.k.a. sequential Monte Carlo (SMC) methods. It aims at popularizing the use of these methods to non-statistician researchers and students, thanks to its automated “black box" inference engine.

It borrows from the BUGS/JAGS software, widely used in Bayesian statistics, the statistical modeling with graphical models and the language associated with their descriptions.

Unlike MCMC methods used by BUGS/JAGS, SMC methods are more adapted to dynamic problems (tracking, signal filtering, etc).

A beta version of the software can be downloaded from the website of the BiiPS project. This software has been presented at the international workshop BayesComp in Kyoto, the international conference ISBA in Tokyo, the conference on Premières Rencontres R in Bordeaux, and the international workshop on efficient simulation in finance in Paris.

6. New Results

6.1. Bayesian Nonparametric models for ranked data and bipartite graphs.

In [20], the author develops a novel Bayesian nonparametric model for random bipartite graphs. The model is based on the theory of completely random measures and is able to handle a potentially infinite number of nodes. It is shown that the model has appealing properties and in particular it may exhibit a power-law behavior. Posterior characterization, a generative process for network growth, and a simple Gibbs sampler for posterior simulation are derived. The model is shown to be well fitted to several real-world social networks.

In [21], we develop a Bayesian nonparametric extension of the popular Plackett-Luce choice model that can handle an infinite number of choice items. Our framework is based on the theory of random atomic measures, with the prior specified by a gamma process. We derive a posterior characterization and a simple and effective Gibbs sampler for posterior simulation. We develop a time-varying extension of our model, and apply it to the New York Times lists of weekly bestselling books.

6.2. A new model for polychotomous data

Multinomial logistic regression is one of the most popular models for modelling the effect of explanatory variables on a subject choice between a set of specified options. This model has found numerous applications in machine learning, psychology or economy. Bayesian inference in this model is non trivial and requires, either to resort to a Metropolis-Hastings algorithm, or rejection sampling within a Gibbs sampler. In [19], we propose an alternative model to multinomial logistic regression. The model builds on the Plackett-Luce model, a popular model for multiple comparisons. We show that the introduction of a suitable set of auxiliary variables leads to an Expectation-Maximization algorithm to find Maximum A Posteriori estimates of the parameters. We further provide a full Bayesian treatment by deriving a Gibbs sampler, which only requires to sample from highly standard distributions. We also propose a variational approximate inference scheme. All are very simple to implement. One property of our Plackett-Luce regression model is that it learns a sparse set of feature weights. We compare our method to sparse Bayesian multinomial logistic regression and show that it is competitive, especially in presence of polychotomous data.

6.3. Sparsity-Promoting Bayesian Dynamic Linear Models

Sparsity-promoting priors have become increasingly popular over recent years due to an increased number of regression and classification applications involving a large number of predictors. In time series applications where observations are collected over time, it is often unrealistic to assume that the underlying sparsity pattern is fixed. We propose in [37] an original class of flexible Bayesian linear models for dynamic sparsity modelling. The proposed class of models expands upon the existing Bayesian literature on sparse regression using generalized multivariate hyperbolic distributions. The properties of the models are explored through both analytic results and simulation studies. We demonstrate the model on a financial application where it is shown that it accurately represents the patterns seen in the analysis of stock and derivative data, and is able to detect major events by filtering an artificial portfolio of assets.

6.4. Evolutionnary algorithms and genetic programming

In [22], we consider the identification of a nonlinear system modelled by a nonlinear output error (NOE) model when the system output is disturbed by an additive zero-mean white Gaussian noise. In that case, standard on-line or off-line least squares methods may lead to poor results. Here, our approach is based on evolutionary algorithms. Although their computational cost can be higher than the above methods, these algorithms present some advantages, which often lead to an effortless optimisation. Indeed, they do not need an elaborate formalisation of the problem. When their parameters are correctly tuned, they avoid to get stuck at a local optimum. To take into account the influence of the additive noise, we investigate different approaches and we suggest a whole protocol including the selection of a fitness function and a stop rule. Without loss of generality, simulations are provided for two nonlinear systems and various signal-to-noise ratios.

The regularity of a signal can be numerically expressed using Hölder exponents, which characterize the singular structures a signal contains. In particular, within the domains of image processing and image understanding, regularity-based analysis can be used to describe local image shape and appearance. However, estimating the Hölder exponent is not a trivial task, and current methods tend to be computationally slow and complex. The paper [17] presents an approach to automatically synthesize estimators of the pointwise Hölder exponent for digital images. This task is formulated as an optimization problem and Genetic Programming (GP) is used to search for operators that can approximate a traditional estimator, the oscillations method. Experimental results show that GP can generate estimators that achieve a low error and a high correlation with the ground truth estimation. Furthermore, most of the GP estimators are faster than traditional approaches, in some cases their runtime is orders of magnitude smaller. This result allowed us to implement a real-time estimation of the Hölder exponent on a live video signal, the first such implementation in current literature. Moreover, the evolved estimators are used to generate local descriptors of salient image regions, a task for which a stable and robust matching is achieved, comparable with state-of-the-art methods. In conclusion, the evolved estimators produced by GP could help expand the application domain of Hölder regularity within the fields of image analysis and signal processing.

One of the main open problems within Genetic Programming (GP) is to meaningfully characterize the difficulty (or hardness) of a problem. The general goal is to develop predictive tools that can allow us to identify how difficult a problem is for a GP system to solve. In [23] and [24], we identify and compare two main approaches that address this question. We denote the first group of methods as Evolvability Indicators (EI), which are measures that attempt to capture how amendable the fitness landscape is to a GP search. The best examples of current EIs are the Fitness Distance Correlation (FDC) and the Negative Slope Coefficient (NSC). The second, more recent, group of methods are what we call Predictors of Expected Performance (PEP), which are predictive models that take as input a set of descriptive attributes of a particular problem and produce as output the expected performance of a GP system. The experimental work presented here compares an EI, the NSC, and a PEP model for a GP system applied to data classification. Results suggest that the EI fails at measuring problem difficulty expressed by the performance of the GP classifiers, an unexpected result. On the other hand, the PEP models show a very high correlation with the actual performance of the GP system. It appears that while an EI can correctly estimate the difficulty of a given search, as shown by previous research on this topic, it does not necessarily capture the difficulty of the underlying problem that

GP is intended to solve. Conversely, while the PEP models treat the GP system as a computational black-box, they can still provide accurate performance predictions.

In [32], the goal is to predict the alertness of an individual by analyzing the brain activity through electroencephalographic data (EEG) captured with 58 electrodes. Alertness is characterized here as a binary variable that can be in a "normal" or "relaxed" state. We collected data from 44 subjects before and after a relaxation practice, giving a total of 88 records. After a pre-processing step and data validation, we analyzed each record and discriminate the alertness states using our proposed "slope criterion". Afterwards, several common methods for supervised classification (k nearest neighbors, decision trees (CART), random forests, PLS and discriminant sparse PLS) were applied as predictors for the state of alertness of each subject. The proposed "slope criterion" was further refined using a genetic algorithm to select the most important EEG electrodes in terms of classification accuracy. Results shown that the proposed strategy derives accurate predictive models of alertness.

6.5. Moderate Deviations for Mean Field Particle Models

The article [40] is concerned with moderate deviation principles of a general class of mean eld type interacting particle models. We discuss functional moderate deviations of the occupation measures for both the strong - topology on the space of fi nite and bounded measures as well as for the corresponding stochastic processes on some class of functions equipped with the uniform topology. Our approach is based on an original semigroup analysis combined with stochastic perturbation techniques and projective limit large deviation methods.

6.6. Bifurcating autoregressive processes

In [42], we investigate the asymptotic behavior of the least squares estimator of the unknown parameters of random coefficient bifurcating autoregressive processes. Under suitable assumptions on inherited and environmental effects, we establish the almost sure convergence of our estimates. In addition, we also prove a quadratic strong law and central limit theorems. Our approach mainly relies on asymptotic results for vector-valued martingales together with the well-known Rademacher-Menchov theorem.

In [46], we study the asymptotic behavior of the weighted least square estimators of the unknown parameters of random coefficient bifurcating autoregressive processes. Under suitable assumptions on the immigration and the inheritance, we establish the almost sure convergence of our estimators, as well as a quadratic strong law and central limit theorems. Our study mostly relies on limit theorems for vector-valued martingales.

In [47], we study the asymptotic behavior of the weighted least squares estimators of the unknown parameters of bifurcating integer-valued autoregressive processes. Under suitable assumptions on the immigration, we establish the almost sure convergence of our estimators, together with the quadratic strong law and central limit theorems. All our investigation relies on asymptotic results for vector-valued martingales.

6.7. Durbin-Watson statistic and first order autoregressive processes

In [45], we investigate moderate deviations for the Durbin-Watson statistic associated with the stable first-order autoregressive process where the driven noise is also given by a first-order autoregressive process. We first establish a moderate deviation principle for both the least squares estimator of the unknown parameter of the autoregressive process as well as for the serial correlation estimator associated with the driven noise. It enables us to provide a moderate deviation principle for the Durbin-Watson statistic in the easy case where the driven noise is normally distributed and in the more general case where the driven noise satisfies a less restrictive Chen-Ledoux type condition.

In [51], we investigate the asymptotic behavior of the Durbin-Watson statistic for the general stable $p-$order autoregressive process when the driven noise is given by a first-order autoregressive process. We establish the almost sure convergence and the asymptotic normality for both the least squares estimator of the unknown vector parameter of the autoregressive process as well as for the serial correlation estimator associated with the driven noise. In addition, the almost sure rates of convergence of our estimates are also provided. Then, we prove the almost sure convergence and the asymptotic normality for the Durbin-Watson statistic. Finally, we propose a new bilateral statistical procedure for testing the presence of a significative first-order residual autocorrelation and we also explain how our procedure performs better than the commonly used Box-Pierce and Ljung-Box statistical tests for white noise applied to the stable autoregressive process, even on small-sized samples.

6.8. Markovian superquadratic BSDEs

In [Stochastc Process. Appl., 122(9):3173-3208], the author proved the existence and the uniqueness of solutions to Markovian superquadratic BSDEs with an unbounded terminal condition when the generator and the terminal condition are locally Lipschitz. In [50], we prove that the existence result remains true for these BSDEs when the regularity assumptions on the generator and/or the terminal condition are weakened.

6.9. Non-Asymptotic Analysis of Adaptive and Annealed Feynman-Kac Particle Models

Sequential and Quantum Monte Carlo methods, as well as genetic type search algorithms can be interpreted as a mean field and interacting particle approximations of Feynman-Kac models in distribution spaces. The performance of these population Monte Carlo algorithms is strongly related to the stability properties of nonlinear Feynman-Kac semigroups. In [49], we analyze these models in terms of Dobrushin ergodic coefficients of the reference Markov transitions and the oscillations of the potential functions. Sufficient conditions for uniform concentration inequalities w.r.t. time are expressed explicitly in terms of these two quantities. We provide an original perturbation analysis that applies to annealed and adaptive FK models, yielding what seems to be the first results of this kind for these type of models. Special attention is devoted to the particular case of Boltzmann-Gibbs measures' sampling. In this context, we design an explicit way of tuning the number of Markov Chain Monte Carlo iterations with temperature schedule. We also propose and analyze an alternative interacting particle method based on an adaptive strategy to define the temperature increments.

6.10. A Robbins-Monro procedure for a class of models of deformation

The paper [48] deals with the statistical analysis of several data sets as- sociated with shape invariant models with different translation, height and scaling parameters. We propose to estimate these parameters together with the common shape function. Our approach extends the recent work of Bercu and Fraysse to multivariate shape invariant models. We propose a very efficient Robbins-Monro procedure for the estimation of the translation parameters and we use these esti- mates in order to evaluate scale parameters. The main pattern is estimated by a weighted Nadaraya-Watson estimator. We provide almost sure convergence and asymptotic normality for all estimators. Finally, we illustrate the convergence of our estimation procedure on simulated data as well as on real ECG data.

6.11. Individual load curves intraday forecasting

A dynamic coupled modelling is investigated to take temperature into account in the individual energy consumption forecasting. The objective in [44] is both to avoid the inherent complexity of exhaustive SARIMAX models and to take advantage of the usual linear relation between energy consumption and temperature for thermosensitive customers. We first recall some issues related to individual load curves forecasting. Then, we propose and study the properties of a dynamic coupled modelling taking temperature into account as an exogenous contribution and its application to the intraday prediction of energy consumption. Finally, these theoretical results are illustrated on a real individual load curve. The authors discuss the relevance of such an approach and anticipate that it could form a substantial alternative to the commonly used methods for energy consumption forecasting of individual customers.

7. Bilateral Contracts and Grants with Industry

7.1. Bilateral Contracts with Industry

Contract with CEA CESTA. The aim of this contract is to develop several extensions to the software BiiPS and to provide a Bayesian statistical modelling on various problems of interest to CEA.

Contract with Astrium/EADS. The aim of this contract, in collaboration with the EPI AYIN, is to develop automatic object tracking algorithms on a sequence of images taken from a geostationary satellite. P. Del Moral cosupervises with J. Zerubia the PhD thesis of Paula Craciun on this subject.

Contract with Dassault. The aim of this contract, in collaboration with the EPI I4S is to address calibration problems using interacting Kalman filters.

8. Partnerships and Cooperations

8.1. Regional Initiatives

The project PSI (Psychology and sounds interactions), headed by P. Legrand received a grant by the region Aquitaine for a PhD thesis on “Dimension reduction in supervised learning. Application to the study of brain activity".

8.2. National Initiatives

8.2.1. ANR Propagation (2010-2012)

To combat dramatic event such as happened in Bombay last year (coming from the sea, a terrorist commando killed more than 200 peoples in Bombay city), authorities are decided to deploy efficient sea surveillance system to protect coastal zone including sensitive infrastructures often in vicinity of important cities.

Regulation on frequencies allocation and on coastal constructions is strong constraint to be taken into account to install technical capabilities to permanently survey vulnerable littoral zones. For example, new active sensor shall be frequencies compatible within numerous existing ones in inhabited region. In this context to perform coastal surveillance, attractive solution is to deploy passive sensors networks because:

- Not necessarily compatible within existing active sensors network.
- Provide large possibilities to install the passive sensors, because, it is not needed to be on the shoreline, but can be deployed inside the territory. Such as facility offers more potential sites and then, to optimise the deployment for optimal coverage of the sensitive zone.
- Is totally undetectable by external technical means in hand of structured criminal organisations.

For these objectives, the PROPAGATION project study, develop and experiment a demonstrator to carry out maritime traffic picture from a set of passive sensors: passive radar, AIS and optronic cameras deployed over a coastal site. This is a joint ANR project with DCNS, Thalés and Exavision.

8.2.2. Project PEPII

This is an interdisciplinary exploratory research project, between Institut de Mathématiques de Bordeaux and Laboratory Ecologie & Evolution, UMR 7625 CNRS-UMPC-ENS (responsible: B. Cazelles). The objective of this project on the dynamics of epidemic diseases characterized by multiple strains of pathogens, is to use the competencies of the ALEA team to get efficient Bayesian optimization techniques. An opening workshop on stochastic models and bayesian inference in epidemiology has been organized in Bordeaux in November 2011.

8.3. European Initiatives

8.3.1. Collaborations with Major European Organizations

Partner 1: Oxford University, Department of Statistics (UK)

Interacting particle systems

Bayesian Nonparametrics

Partner 2: Imperial College (UK)

Interacting Particle Systems

8.4. International Research Visitors

8.4.1. Visits of International Scientists

The following researchers visited the Team ALEA during 2012: A. Doucet (Univ. Oxford), C. Holmes (Oxford), N. Whiteley (Univ. Bristol), R. Xu (Univ. of Tech. Sydney), G. Peters (University College London), Pavel V. Shevchenko (CSIRO).

9. Dissemination

9.1. Scientific Animation

9.1.1. Editorial Board

P. Del Moral is currently associate editor/editor for the following journals

- Chief editor : American Journal of Algorithms and Computing, since 2012.
- Associate editor : Applied Mathematics and Optimization, since 2009.
- Associate editor Revista de Matematica: Teoria y Aplicaciones , since 2009.
- Associate editor : Stochastic Analysis and Applications, since 2001.

9.1.2. Senior Program Committee

- International conference on Uncertainty in Artificial Intelligence (UAI'2012): F. Caron
- 1eres rencontres R: P. Legrand

9.1.3. Responsibilities

B. Bercu is responsible of the thematic group MAS (Modélisation Aléatoire et Statistique) at SMAI.

B. Bercu is an assistant director of the Institut de Mathématiques de Bordeaux (IMB). He is also a member of the IMB council and the UFR council of the University of Bordeaux. He is a member of the CNU section 26.

B. Bercu is co-responsible of the specialty "Modélisation Statistique et Stochastique" of the Master MIMSE.

P. Legrand is a member of “bureau de l'association Evolution artificielle".

P. Legrand is in charge of the learning management system MOODLE of the UFR sciences et modélisation (University of Bordeaux II).

9.1.4. Organization of Conferences

- International Conference Evolve 2012: P. Legrand, P. Del Moral (with A.A. Tantar, E. Tantar, P. Bouvry, O. Schütze)
- International workshop on Sequential Monte Carlo methods and Efficient simulation in Finance: P. Del Moral (with E. Gobe, P. Hu)
- Workshop EDF/Inria, New stochastic forecasting methods for individual temporal series in energy context : B. Bercu, F. Proïa (with S. Bercu, P. Lé)

9.1.5. Reviewing

- Journals: Annals of Statistics, IEEE TPAMI, Journal of the Royal Statistical Society B, Computational Statistics and Data Analysis, Statistics and Computing, Journal de la Société Francaise de Statistiques
- Conferences: UAI, NIPS, ICML, AISTATS

9.2. Teaching - Supervision - Juries

9.2.1. Teaching

Licence :

- B. Bercu, Mathématiques générales, Analyse et Algèbre SVE, 36h, L1, University of Bordeaux, France
- P. Legrand, Espaces Euclidiens, 54h, L2, University of Bordeaux, France
- P. Legrand, Traitement du Signal, 18h, L3, University of Bordeaux, France
- P. Legrand, Informatique pour les mathématiques, 36h, L1, University of Bordeaux, France
- P. Legrand, Algèbre, 72h, L1, University of Bordeaux, France
- P. Legrand, Technologies de l'information, de la communication pour l'éducation, 42h, University of Bordeaux, France
- A. Richou, Probabilités et statistiques, 32h, L3, University of Bordeaux, France
- A. Richou, Probabilités et statistiques, 32h, L1, University of Bordeaux, France

Master :

- A. Richou, Probabilité, 32h, M1, University of Bordeaux, France
- B. Bercu, Séries chronologiques, 48h, M2, University of Bordeaux, France
- B. Bercu, Processus aléatoires à temps discret, 30h, M1, University of Bordeaux, France
- B. Bercu, Probabilités, 30h, L3, University of Bordeaux, France
- F. Caron, Bayesian Methods, 33h, M2, University Bordeaux II, France
- F. Caron, Statistical Methods in Robotics, 25h, M2, IPB, France
- F. Caron, Advanced estimation tools in signal and image processing, 30h, M2, University Bordeaux I, France
- P. Legrand, Traitement du signal, 15h, M2, IPB, France

Other:

- P. Del Moral, Professeur chargé de cours (1/3 temps), Monte Carlo methods and Stochastic models, and introduction to probability calculus, Ecole Polytechnique, France.
- P. Del Moral, Mean field particle simulation for Monte Carlo integration, 10h, Lectures INLN-CNRS of the University of Nice Sophia Antipolis.
- P. Legrand, Course on Matlab, 42H

9.2.2. Supervision

PhD:

- Aurélie Le Cain, Caractéristiques spatiales et temporelles d'une tache focale LMJ, university Bordeaux I, Jan. 2012, P. Del Moral and B. Bercu
- Peng Hu, Méthodes particulaires et applications en finance , university Bordeaux I, June 2012, P. Del Moral

PhD in progress :

- Vassili Blandin, Processus autorégressifs à bifurcation, Sept. 2011, B. Bercu
- Frédéric Proia, Processus autorégressifs stables, Dec. 2010, B. Bercu, P. Del Moral
- Philippe Fraysse, Algorithmes stochastiques pour la régression semi-paramétrique, Sept. 2010, B. Bercu
- Paul Lemaitre, Analyse de sensibilité et analyse de risques, Sept. 2010, P. Del Moral
- François Giraud, Méthodes particulaires adaptatives pour l'estimation non linéaire, Nov. 2009, P. Del Moral
- Nicolas Antunès, Etude du modèle GARP pour la prédiction de niches écologiques, Sept. 2011, P. Del Moral and P. Legrand
- Laurent Vézard, Réduction de dimension en apprentissage supervisé. Application à l'étude de l'activité cérébrale, Sept. 2010, P. Legrand
- Antoine Campi, Filtrage particulaire de fluides turbulents, 2012, P. Del Moral
- Christelle Vergé, Méthodes particulaires pour la propagation d'incertitudes dans des codes numériques, 2012, P. Del Moral
- Paula Craciun, Méthodes de filtrage multi-objets en analyse d'image, 2012, P. Del Moral

9.2.3. Juries

- Peng Hu, PhD, University of Bordeaux: P. Del Moral
- Cyrille Dubarry, PhD, TelecomParisSud: P. Del Moral
- Sébastien Gadat, HDR, University of Toulouse: P. Del Moral
- Jérémie Bureau, PhD, University of Toulouse : B. Bercu
- Valère Bitseki Penda, PhD, University Blaise Pascal : B. Bercu
- Pascal Szacherski, PhD, University of Bordeaux: F. Caron

10. Bibliography

Publications of the year

Doctoral Dissertations and Habilitation Theses

[1] P. HU. *Méthodes particulaires et applications en finance*, University of Bordeaux I, 2012, http://www.theses.fr/162166761.

[2] A. LE CAIN. *Caractéristiques spatiales et temporelles d'une tache focale LMJ*, University Bordeaux I, 2012, http://www.theses.fr/2012BOR14475.

Articles in International Peer-Reviewed Journals

[3] B. BERCU, J.-F. BONY, V. BRUNEAU. *Large deviations for Gaussian stationary processes and semi-classical analysis*, in "Séminaire de Probabilités", 2012, vol. 44, p. 409-428 [*DOI :* 10.1007/978-3-642-27461-9], http://hal.inria.fr/hal-00642658.

[4] B. BERCU, L. COUTIN, N. SAVY. *Sharp large deviations for the non-stationary Ornstein-Uhlenbeck process*, in "Stochastic Processes and their Applications", 2012, vol. 122, p. 3393-3424 [*DOI :* 10.1016/J.SPA.2012.06.006], http://hal.inria.fr/hal-00645074.

[5] B. BERCU, P. DEL MORAL, A. DOUCET. *Fluctuations of Interacting Markov Chain Monte Carlo Models*, in "Stochastic Processes and their Applications", 2012, vol. 122, n^o 4, p. 1304-1331, http://hal.inria.fr/inria-00227536.

[6] B. BERCU, P. FRAYSSE. *A Robbins-Monro procedure for estimation in semiparametric regression models*, in "Annals of Statistics", 2012, vol. 40, n^o 2, p. 666-693 [*DOI :* 10.1214/12-AOS969], http://hal.inria.fr/hal-00551832.

[7] B. BERCU, T. MONG NGOC NGUYEN, J. SARACCO. *A new approach on recursive and non-recursive SIR methods*, in "Journal of the Korean Statistical Society", 2012, vol. 41, p. 17-36 [*DOI :* 10.1016/J.JKSS.2011.05.005], http://hal.inria.fr/hal-00642653.

[8] B. BERCU, F. PROÏA. *A sharp analysis on the asymptotic behavior of the Durbin-Watson statistic for the first-order autoregressive process*, in "ESAIM: Probability and Statistics", 2012, vol. 16, p. 1-29 [*DOI :* 10.1051/PS/2012005], http://hal.inria.fr/hal-00642634.

[9] A. BUDHIRAJA, P. DEL MORAL, S. RUBENTHALER. *Discrete Time Markovian Agents Interacting Through a Potential*, in "Journal ESAIM Probability & Statistics", 2012, 22 [*DOI :* 10.1051/PS/2012014], http://hal.inria.fr/hal-00601151.

[10] F. CARON, A. DOUCET. *Efficient Bayesian Inference for Generalized Bradley-Terry Models*, in "Journal of Computational and Graphical Statistics", 2012, vol. 21, n^o 1, p. 174-196, http://hal.inria.fr/inria-00533638.

[11] F. CARON, A. DOUCET, R. GOTTARDO. *On-line changepoint detection and parameter estimation with application to genomic data*, in "Statistics and Computing", 2012, vol. 22, n^o 2, p. 579-595 [*DOI :* 10.1007/S11222-011-9248-X], http://hal.inria.fr/inria-00577217.

[12] F. CÉROU, P. DEL MORAL, T. FURON, A. GUYADER. *Sequential Monte Carlo for rare event estimation*, in "Statistics and Computing", 2012, vol. 22, n^o 3, p. 795-908 [*DOI :* 10.1007/S11222-011-9231-6], http://hal.inria.fr/inria-00584352.

[13] P. DEL MORAL, A. DOUCET, A. JASRA. *On Adaptive Resampling Procedures for Sequential Monte Carlo Methods*, in "Bernoulli", 2012, vol. 18, n^o 1, p. 252-278, http://hal.inria.fr/inria-00332436.

[14] P. DEL MORAL, P. HU, N. OUDJANE. *Snell envelope with small probability criteria*, in "Applied Mathematics and Optimization", 2012, vol. 66, n^o 3, p. 309-330, http://hal.inria.fr/hal-00758905.

[15] A. LEE, F. CARON, A. DOUCET, C. HOLMES. *Bayesian Sparsity-Path-Analysis of Genetic Association Signal using Generalized t Priors*, in "Statistical Applications in Genetics and Molecular Biology", 2012, vol. 11, n^o 2 [*DOI :* 10.2202/1544-6115.1712], http://hal.inria.fr/inria-00598490.

[16] A. RICHOU. *Markovian quadratic and superquadratic BSDEs with an unbounded terminal condition*, in "Stochastic Processes and their Applications", September 2012, vol. 122, n^o 9, p. 3173–3208 [*DOI :* 10.1016/J.SPA.2012.05.015], http://hal.inria.fr/hal-00643198.

[17] L. TRUJILLO, P. LEGRAND, G. OLAGUE, J. LÉVY-VEHEL. *Evolving Estimators of the Pointwise Holder Exponent with Genetic Programming*, in "Information Sciences", 2012, vol. 209, p. 61-79 [*DOI :* 10.1016/J.INS.2012.04.043], http://hal.inria.fr/hal-00643387.

Articles in National Peer-Reviewed Journals

[18] L. VEZARD, P. LEGRAND, M. CHAVENT, F. FAITA-AINSEBA, J. CLAUZEL. *Classification de données EEG par algorithme évolutionnaire pour l'étude d'états de vigilance*, in "Revue des nouvelles technologies de l'information", 2012, http://hal.inria.fr/hal-00643438.

International Conferences with Proceedings

[19] C. ARCHAMBEAU, F. CARON. *Plackett-Luce regression: A new Bayesian model for polychotomous data*, in "Conference on Uncertainty in Artificial Intelligence (UAI'2012)", Catalina Island, United States, August 2012, http://hal.inria.fr/hal-00708441.

[20] F. CARON. *Bayesian nonparametric models for bipartite graphs*, in "NIPS - Neural Information Processing Systems", Lake Tahoe, United States, December 2012, http://hal.inria.fr/hal-00730789.

[21] F. CARON, Y. W. TEH. *Bayesian nonparametric models for ranked data*, in "NIPS - Neural Information Processing Systems", Lake Tahoe, United States, MIT Press, December 2012, http://hal.inria.fr/hal-00730792.

[22] Z. SIGRIST, P. LEGRAND, E. GRIVEL, B. ALCOVERRO. *Considering the measurement noise for a nonlinear system identification with evolutionary algorithms.*, in "EUSIPCO 2012", Bucharest, Romania, August 2012, http://hal.inria.fr/hal-00757231.

[23] L. TRUJILLO, Y. MARTINEZ, E. GALVAN-LOPEZ, P. LEGRAND. *A Comparative Study of an Evolvability Indicator and a Predictor of Expected Performance for Genetic Programming*, in "GECCO", Philadelphie, United States, July 2012, http://hal.inria.fr/hal-00757266.

[24] L. TRUJILLO, Y. MARTINEZ, E. GALVAN-LOPEZ, P. LEGRAND. *A comparison of predictive measures of problem difficulty for classification with Genetic Programming*, in "ERA 2012", Tijuana, Mexico, November 2012, http://hal.inria.fr/hal-00757363.

Conferences without Proceedings

[25] A. TODESCHINI, F. CARON, P. LEGRAND, P. DEL MORAL. *BiiPS : un logiciel pour l'inférence bayésienne dans les modèles graphiques utilisant des méthodes de Monte Carlo séquentielles*, in "1ères Rencontres R", Bordeaux, France, July 2012, http://hal.inria.fr/hal-00717547.

[26] L. VEZARD, P. LEGRAND, M. CHAVENT, F. FAITA-AINSEBA, J. CLAUZEL. *Classification of EEG signals by an evolutionary algorithm*, in "COMPSTAT 2012", Limassol, Cyprus, August 2012, http://hal.inria.fr/hal-00757270.

[27] L. VEZARD, P. LEGRAND, M. CHAVENT, F. FAITA-AINSEBA. *Sélection de variables par algorithme génétique pour l'étude d'états de vigilance*, in "Journée Évolutionnaire Thématique, 23éme édition", Paris, France, November 2012, http://hal.inria.fr/hal-00759420.

Scientific Books (or Scientific Book chapters)

[28] R. CARMONA, P. DEL MORAL, P. HU, N. OUDJANE. *An Introduction to Particle Methods with Financial Applications*, in "Numerical methods in finance", R. CARMONA, P. DEL MORAL, P. HU, N. OUDJANE (editors), Proceedings in mathematics, Springer, 2012, vol. 12, p. 3-49, http://hal.inria.fr/hal-00641132.

[29] P. DEL MORAL, P. HU, L. WU. *On the Concentration Properties of Interacting Particle Processes*, Foundations and Trends in Machine Learning, NEW PUBL INC, 2012, vol. 3, 167 [*DOI :* 10.1561/2200000026], http://hal.inria.fr/hal-00758901.

[30] P. DEL MORAL, B. REMILLARD, S. RUBENTHALER. *Monte Carlo approximations of American options that preserve monotonicity and convexity*, in "Numerical methods in finance", R. CARMONA, P. DEL MORAL, P. HU, N. OUDJANE (editors), Proceedings in mathematics, Springer, 2012, vol. 12, p. 115-143, http://hal.inria.fr/hal-00641133.

[31] P. DEL MORAL, E. TANTAR, A.-A. TANTAR. *On the foundations and the applications of evolutionary computing*, in "Evolve : A bridge between Probability, Set Oriented Numerics and Evolutionary Computation", E. TANTAR, A.-A. TANTAR, P. BOUVRY, P. DEL MORAL, P. LEGRAND, C. COELLO COELLO, O. SCHUTZE (editors), Springer, 2013, p. 3-89, http://hal.inria.fr/hal-00756117.

[32] L. VEZARD, P. LEGRAND, M. CHAVENT, F. FAITA-AINSEBA, J. CLAUZEL, L. TRUJILLO. *Classification of EEG signals by evolutionary algorithm*, in "Advances in Knowledge Discovery and Management", Springer, 2013, vol. 4, http://hal.inria.fr/hal-00759439.

Books or Proceedings Editing

[33] R. CARMONA, P. DEL MORAL, P. HU, N. OUDJANE (editors). *Numerical methods in finance*, Proceedings in mathematics, Springer, 2012, vol. 12, 471, http://hal.inria.fr/hal-00641465.

[34] J.-K. HAO, P. LEGRAND, P. COLLET, N. MONMARCHÉ, E. LUTTON, M. SCHOENAUER (editors). *Artificial Evolution 2011*, Springer, November 2012, 508, to appear, http://hal.inria.fr/hal-00643404.

[35] O. SCHUTZE, C. COELLO COELLO, A.-A. TANTAR, E. TANTAR, P. BOUVRY, P. DEL MORAL, P. LEGRAND (editors). *Evolve - A bridge between probability, set oriented numerics and evolutionary computation II*, Advances in Intelligent Systems and Computing, Springer, 2013, vol. 175, 508, http://hal.inria.fr/hal-00756101.

[36] E. TANTAR, A.-A. TANTAR, P. BOUVRY, P. DEL MORAL, P. LEGRAND, C. COELLO COELLO, O. SCHUTZE (editors). *Evolve : A bridge between probability, set oriented numerics and evolutionary computation*, Studies in Computational Intelligence, Springer, 2013, vol. 447, 414, http://hal.inria.fr/hal-00641466.

Research Reports

[37] F. CARON, L. BORNN, A. DOUCET. *Sparsity-Promoting Bayesian Dynamic Linear Models*, Inria, February 2012, n^o RR-7895, 23, http://hal.inria.fr/hal-00675274.

[38] F. CARON, C. HOLMES, E. RIO. *On the sampling distribution of an ℓ^2 distance between Empirical Distribution Functions with applications to nonparametric testing*, Inria, April 2012, n^o RR-7931, http://hal.inria.fr/hal-00688141.

[39] F. CARON, Y. W. TEH, T. B. MURPHY. *Bayesian nonparametric Plackett-Luce models for the analysis of clustered ranked data*, Inria, November 2012, n^o RR-8143, http://hal.inria.fr/hal-00755478.

[40] P. DEL MORAL, S. HU, L. WU. *Moderate Deviations for Mean Field Particle Models*, Inria, April 2012, 37, http://hal.inria.fr/hal-00687827.

[41] P. LEMAÎTRE, E. SERGIENKO, A. ARNAUD, N. BOUSQUET, F. GAMBOA, B. IOOSS. *Density modification based reliability sensitivity analysis*, Inria, October 2012, http://hal.inria.fr/hal-00737978.

Other Publications

[42] B. BERCU, V. BLANDIN. *A Rademacher-Menchov approach for random coefficient bifurcating autoregressive processes*, 2012, http://hal.inria.fr/hal-00745634.

[43] B. BERCU, T. MONG NGOC NGUYEN, J. SARACCO. *On the asymptotic behavior of the Nadaraya-Watson estimator associated with the recursive SIR method*, 2012, http://hal.inria.fr/hal-00673832.

[44] S. BERCU, F. PROÏA. *A SARIMAX coupled modelling applied to individual load curves intraday forecasting*, 2012, 17 pages, 18 figures, 2 tables, http://hal.inria.fr/hal-00732844.

[45] V. BITSEKI PENDA, H. DJELLOUT, F. PROÏA. *Moderate deviations for the Durbin-Watson statistic related to the first-order autoregressive process*, 2012, 24 pages., http://hal.inria.fr/hal-00661075.

[46] V. BLANDIN. *Asymptotic results for bifurcating random coefficient autoregressive processes*, 2012, http://hal.inria.fr/hal-00696295.

[47] V. BLANDIN. *Limit theorems for bifurcating integer-valued autoregressive processes*, 2012, http://hal.inria.fr/hal-00696297.

[48] P. FRAYSSE. *A Robbins-Monro procedure for a class of models of deformation*, 2012, 29 pages, 12 figures, http://hal.inria.fr/hal-00696050.

[49] F. GIRAUD, P. DEL MORAL. *Non-Asymptotic Analysis of Adaptive and Annealed Feynman-Kac Particle Models*, 2012, 36 pages, http://hal.inria.fr/hal-00735690.

[50] F. MASIERO, A. RICHOU. *On the existence of Markovian superquadratic BSDEs with an unbounded terminal condition*, 2012, http://hal.inria.fr/hal-00713123.

[51] F. PROÏA. *A new statistical procedure for testing the presence of a significative correlation in the residuals of stable autoregressive processes*, 2012, 34 pages, 2 figures., http://hal.inria.fr/hal-00677600.

References in notes

[52] N. BARTOLI, P. DEL MORAL. *Simulation and Stochastic Algorithms (An introduction with applications)*, Cépaduès Editions, 2001.

[53] P. DEL MORAL. *Feynman-Kac formulae. Genealogical and interacting particle systems, with applications*, Springer Verlag New York, Series: Probability and its Applications, 2004.

[54] P. DEL MORAL, L. MICLO. *Genealogies and Increasing Propagations of Chaos for Feynman-Kac and Genetic Models*, in "Annals of Applied Probability", 2001, vol. 11, n^{o} 4, p. 1166-1198.

[55] P. DEL MORAL, L. MICLO. *On convergence of chains with time empirical self-interactions*, in "Proc. Royal Soc. Lond. A.", 2003, vol. 460, p. 325–346.

[56] P. DEL MORAL, L. MICLO. *Self Interacting Markov chains*, in "Stochastic Analysis and Applications", 2006, vol. 24, n^o 3, p. 615–660.

[57] A. JASRA, C. ANDRIEU, A. DOUCET, P. DEL MORAL. *Non-Linear Markov chain Monte Carlo via Self Interacting Approximations*, 2007, Univ. Bristol Preprints.

[58] A. JASRA, C. ANDRIEU, A. DOUCET, P. DEL MORAL. *A Note on Convergence of the Equi-energy Sampler*, in "Journal Stochastic Analysis and its Applications", 2008.