



INSTITUT NATIONAL DE RECHERCHE EN INFORMATIQUE ET EN AUTOMATIQUE

Project-Team CQFD

Contrôle de Qualité & Fiabilité Dynamique

Bordeaux - Sud-Ouest

Theme : Stochastic Methods and Models

Activity
R *eport*

2009

Table of contents

1. Team	1
2. Overall Objectives	1
2.1. Presentation	1
2.2. Highlights	2
3. Scientific Foundations	2
3.1. Dynamic Reliability	2
3.1.1. Modeling	2
3.1.2. Estimation	3
3.1.3. Synthesis of maintainability and reconfiguration systems	4
3.2. Statistical Quality Control	4
3.2.1. Modeling	5
3.2.2. Estimation	5
4. Application Domains	6
5. Software	6
6. New Results	6
6.1. Average Continuous Control of Piecewise Deterministic Markov Processes	6
6.2. The Vanishing Discount Approach for the Average Continuous Control of Piecewise Deterministic Markov Processes	7
6.3. The Policy Iteration Algorithm for Average Continuous Control of Piecewise Deterministic Markov Processes	7
6.4. Optimal stopping of Piecewise Deterministic Markov Processes	8
6.5. Modeling of crack propagation by Piecewise Deterministic Markov Processes	8
6.6. Asymptotic behavior of bifurcating autoregressive processes	9
6.7. On the usefulness of persistent excitation in ARX adaptive tracking	9
6.8. A new concept of strong controllability via the Schur complement in adaptive tracking	10
6.9. On the almost sure central limit theorem for vector martingales : convergence of moments and statistical applications	10
6.10. A functional central limit theorem for a class of interacting Markov chain Monte Carlo methods	10
6.11. Almost Sure Stabilization for Feedback Controls of Regime-switching Linear Systems with A Hidden Markov Chain	10
6.12. Dimension reduction via Sliced Inverse Regression	11
6.13. Nonparametric conditional quantile estimation	12
6.14. Semiparametric Mode Regression Estimation	12
6.15. Dimension Reduction for categorical data	12
7. Contracts and Grants with Industry	13
7.1. Electricité de France	13
7.2. Voies Navigables de France	13
7.3. EADS Astrium	13
7.4. National Initiatives	14
7.5. International Initiatives	14
8. Dissemination	14
8.1. Animation of the scientific community	14
8.2. Workshop organisation	15
8.3. Participation to congresses, conferences, invitations	15
8.4. Teaching	16
9. Bibliography	17

CQFD is an INRIA Team joint with University of Bordeaux (UB1, UB2 and UB4) and CNRS (IMB, UMR 5251) and Institut Polytechnique de Bordeaux.

1. Team

Faculty Member

François Dufour [Institut Polytechnique de Bordeaux, Team Leader, Professeur (Pr), HdR]

Bernard Bercu [University Bordeaux 1, Professeur (Pr), HdR]

Marie Chavent [University Bordeaux 2, Assistant Professor (MdB), HdR]

Anne Gégout-Petit [University Bordeaux 2, Assistant Professor (MdB)]

Benoîte de Saporta [University Bordeaux 4, Assistant Professor (MdB)]

Jérôme Saracco [University Bordeaux 4, Professor (Pr), HdR]

Huilong Zhang [University Bordeaux 1, Assistant Professor (MdB)]

Technical Staff

Frédéric Proia [Research Engineer, since Nov. 2009]

PhD Student

Adrien Brandejsky [CORDI-C grant, since Sept. 2008]

Karen Gonzalez [University Bordeaux 1, MENRT grant, since Sep. 2007]

Vanessa Kuentz [University Bordeaux 1, MENRT grant, since Mar. 2007]

Aurélien Le Cain [CEA grant, since Oct. 2008]

Thi Mong Ngoc Nguyen [CNRS, BDI/CNRS grant, since Sep. 2007]

Victor Vasquez [University of Mexico, ECOS/CONACIT grant, since Sep. 2007]

Administrative Assistant

Nathalie Dupuy

Other

Romain Azais [Student internship, from Jun. to Aug. 2009]

2. Overall Objectives

2.1. Presentation

Economic, scientific and military competition leads many industrial sectors to design ever more successful and reliable processes and equipment.

Reliability and quality control, and more generally dependability and safety, have become a crucial area in the field of industrial engineering. The term reliability is an acquisition of the 20th century. Initially, reliability was developed to meet the needs of the electronics industry. This was a consequence of the fact that the first complex systems appeared in this field of engineering. Such systems have a huge number of components which made their global reliability very low in spite of their relatively highly reliable components. This led to a specialized applied mathematical discipline which allowed one to make an a priori evaluation of various reliability indexes at the design stage, to choose an optimal system structure, to improve methods of maintenance, and to estimate the reliability on the basis of special testing or exploitation.

Our objective is to apply probabilistic and statistical tools from estimation and control theory to dependability and safety. We wish to investigate the following fields:

- 1) Design and analysis of realistic and accurate random models for dependability. In particular, we will study parametric models for dynamic reliability and semi- or non-parametric models for quality control.
- 2) Implementation of estimation algorithms in relation with our stochastic models and evaluation of reliability indexes.
- 3) Design of control for maintenance and reconfiguration.

We stress the fact that points 1) and 2) are strongly interlinked. Indeed, designing mathematical models for reliability is an important and basic research field. However, our models will be legitimated and practically validated by point 2). In particular, the feasibility of estimation routines and the quality of the evaluation of reliability indexes will be crucial. Point 3) deals with control through practical issues of maintenance and reconfiguration. This last point legitimates and is based on the first two ones: only after modelling, identifying and evaluating reliability indexes, shall we be able to compute a cost criterion to be optimized.

2.2. Highlights

Our team was just awarded a grant from ARPEGE program of the French National Agency of Research (ANR): project "FAUTOCOES", number ANR-09-SEGI-004 starting in October 2009.

An important contract was signed in June 2009 between EDF R&D team ICAME, and INRIA teams ALEA and CQFD. It deals with new statistical methods for recursive adaptive prediction of the electricity consumption of EDF customers. The contract was divided into five lots and it will last until 2011. The total value of the contract is 140000 euros.

F. Dufour has been nominated as an Associate Editor of the SIAM Journal on Control and Optimization.

The team CQFD has been very involved in the organization of the "41e Journées de Statistique" Annual event of the "Société Française de Statistique" which took place in Bordeaux in 2009 : 450 participants, 20 plenary sessions, 230 contributed talks, budget 100 000 euros. Reliability and Quality Control was one of the main topics of the meeting. More details are available via the link <http://www.sm.u-bordeaux2.fr/JDS2009/>

The team CQFD has organized in June 2009 the "5e Rencontre statistiques Mathématiques Bordeaux-Santander-Toulouse-Valladolid". The goal of this annual workshop is to strengthen the scientific exchanges between our Universities. The universities of Pau and Montpellier have participated in the workshop. More details are available via the link <http://www.math.u-bordeaux1.fr/~bercu/Event/Bosantouval/Bosantouval2009.html>

The teams ALEA and CQFD, as well as INRIA Bordeaux Sud-Ouest, are very involved in the organization of the "Journées MAS 2010". This biannual meeting is the most important event for french probabilists and statisticiens. It will take place in Bordeaux 1 university in September 2010. The topic of this meeting is stochastic algorithms and combinatorics. More details are available via the link <http://www.math.u-bordeaux1.fr/MAS10/>

3. Scientific Foundations

3.1. Dynamic Reliability

Participants: Bernard Bercu, Anne Gégout-Petit, Benoîte de Saporta, François Dufour, Huilong Zhang, Karen Gonzalez.

3.1.1. Modeling

In dependability and safety theory, modeling is a key step to study the properties of the physical processes involved. Nowadays, it appears necessary to take explicitly and realistically the dependencies into account, meaning the dynamic interactions existing between the physical parameters (for example: pressure, temperature, flow rate, level, ...) of the system and the functional and dysfunctional behavior of its components. Classically, the models described in the dependability and safety literature do not take such interactions into account. A first set of methods used in reliability theory is the so-called combinatory approaches (fault trees, event trees, reliability diagrams and networks, ...), which can be used to identify and evaluate the combinations of events leading to the occurrence of other desirable or undesirable events. These powerful methods suffer from the fact that such combinations do not take the order of occurrence into account, in the sense that they eliminate any notion of dependency between events. A second set of methods is described by finite state Markov (or semi-Markov) models. In this context, the system is described by a fixed number of components

which can be in different states. For any component, the set of its possible states is assumed to be finite (generally it contains only two elements: an operational and a failure state). One of the main limitations encountered with such models is their difficulties to model correctly the physical processes involving deterministic behavior. To overcome such difficulties, dynamic reliability was introduced in 1980 as a powerful mathematical framework capable of explicitly handling interactions between components and process variables. Nowadays in the literature, the multi-model approach appears as a natural framework to formulate dynamic reliability problems. The behavior of the physical model is thus described by different modes of operation from nominal to failure states with intermediate dysfunctional regimes. For a large class of industrial processes, the layout of operational or accident sequences generally comes from the occurrence of two types of events:

- 1) The first type of event is directly linked to a deterministic evolution of the physical parameters of the process.
- 2) The second type of event is purely stochastic and usually corresponds to random demands or failures of system components.

In both cases, these events will induce jumps in the behavior of the system leading to stable or unstable trajectories for the process.

In 1980, M.H.A. Davis [54] introduced in probability theory the Piecewise Deterministic Markov Processes (PDMP) as a general class of models suitable for formulating optimization problems in queuing and inventory systems, maintenance-replacement models, investment scheduling and many other areas of operation research. These models are described by two variables: to the usual Euclidean variable representative of the state of the process, one adds a discrete variable, called regime or mode and takes values in a finite or countable set. In this context, the state variable represents the physical parameters of the system under consideration. For example, it can be the position or the orientation of a satellite or the pressure in a tank. The mode characterizes the regimes of operation of the physical system from nominal to failure regime. From a mathematical point of view, the notion of a piecewise deterministic Markov process is very intuitive and simple to describe. The state space of this system is given, for example, by an open subset E of the set \mathbb{R}^d . Starting from $x \in E$, the process follows a deterministic trajectory, namely a flow indexed by the mode, until the first jump time T_1 , which occurs either spontaneously in a random manner or when the trajectory hits the boundary of E . Between two jumps, the mode is assumed to be constant. In both cases, a new point and a new regime are selected by a random operator and the process restarts from this new point under this new mode. There exist two types of jump :

- 1) The first one is deterministic. From the mathematical point of view, it is given by the fact that the trajectory hits the boundary of E . From the physical point of view, it can be seen as a modification of the mode of operation when a physical parameter reaches a prescribed level (for example when the pressure of a tank reaches the critical value).
- 2) The second one is stochastic. It models the random nature of failures or inputs that modify the mode of operation of the system.

As it has been illustrated above, the key asset of this mathematical model is that it takes naturally into account the two kinds of events previously described. Several examples can be found in [59], [78], [72], [73] and [79]. Most stochastic processes presented in T. Aven and U. Jensen [41] are special cases of piecewise deterministic Markov processes.

In conclusion, it is claimed that piecewise deterministic Markov processes provide a general framework to study dynamic reliability problems. Their dynamical properties allow explicit time dependencies, in contrast with piecewise constant jump Markov processes. Consequently, these processes are really suitable for modeling real phenomena of dynamic reliability.

3.1.2. Estimation

The probabilistic background offers a very suitable framework to evaluate material quality from the dependability and safety point of view. One can classically characterize the performances of a system by several

indicators : availability, reliability, maintainability, safety, etc. Evaluating all these indicators is crucial. It makes it possible to calculate a certain *cost* in order to measure the performances of the system. Hence, the well-known topic in control, called *robustness*, is given emphasis. In this framework, it is necessary to define the concept of subsystem and sensitivity:

- 1) Which are the subsystems of greater impact on the cost?
- 2) What is the evolution of cost sensibility with respect to modifications of one or several components of the system ?

For instance, evaluating the production availability of a factory is a vital concern for the industrial world. This notion complements the more classical notions of instantaneous availability and asymptotic availability. Production availability is a probability measure of the regularity of production. Previously, its calculation was usually based on the naive hypothesis that the production level associated with each mode of regime (operational, damaged, partial breakdown, ...) was instantaneously reached as soon as the system entered that state. Consequently, modeling of production availability was done via a discrete random variable and a typical trajectory of the production availability was piecewise constant. It was shown in [69] on a large set of real cases that this hypothesis was not realistic. In fact, the production level evolves continuously and is influenced by the mode of regime as well as internal variables of the system such as pressure, temperature, etc.

Quite obviously, it is necessary to take into account the naturally continuous dynamic of indicators in Reliability, Availability and Maintainability (RAM). In particular, we shall see that the so-called Piecewise Deterministic Markov Processes are very suitable tools to define and evaluate the indicators in RAM for physical systems.

3.1.3. Synthesis of maintainability and reconfiguration systems

In the domain of safety and dependability, the notions of control and maintainability play a prominent role in the design of reliable systems. This maintainability can be *active* or *passive*.

In this context of vulnerability, the usual way to make the system more tolerant towards failures is to introduce several redundancies. We improve the reliability of a system not only by improving the reliability of its components but also via their redundant organization. This is commonly called *passive maintainability*. However, it is not always possible to introduce direct physical redundancy which clearly restricts the usefulness of this approach. For example, it seems impossible to put motor units or pressure transducers in the same place on certain structures such as oil wells or communication satellites.

A second approach, more realistic and promising, is *active maintainability*. It is organized in the two following steps:

- 1) Detection and identification of failures,
- 2) Reconfiguration of the system.

In this context, PDMP are especially well-adapted for modeling real physical systems. In fact, a natural approach is first to make out a list of possible failures or breakdowns. This will lead to the constitution of a set of regimes or modes for the system described in section 3.1.1. Then, after the detection and identification of all those regimes, the control or maintainability process will be in a position to react and maintain the system in a damaged but acceptable mode. However, this modeling approach is subject to a number of limitations in terms of *efficiency/complexity*. More precisely, if a non-identified breakdown occurs, this approach can fail dramatically. A simple way to rectify this situation out is to include this kind of failure into the list of possible regimes. However, this will increase the complexity of the model. Therefore, a compromise must be sought during the modeling phase.

3.2. Statistical Quality Control

Participants: Bernard Bercu, Marie Chavent, Jérôme Saracco, Huilong Zhang, Vanessa Kuentz, Thi Mong Ngoc Nguyen.

3.2.1. Modeling

A classical aim of reliability is to study censored survival data. In this context, several parametric, semi-parametric, and non-parametric forms of modeling for survival functions estimation have already been proposed. In this project, we focus our attention on another aspect of reliability: Statistical Quality Control (SQC). More precisely, we wish to develop non-parametric and semi-parametric modeling in order to provide tolerance curves and hyper surfaces.

Tolerance curves are used in industry to predict performance of the manufacturing process from external measures such as temperature or pressure. They are particularly useful when the quality control is late (long manufacture time, intermediate storage ...) or results from small size samples. Tolerance curves provide the inspectors with a tool to control whether the evaluated parameters are within the interval required in the specifications and to make the inspection organization more efficient. Because of their graphical representation, they are particularly easy to use.

A tolerance interval differs from the well-known confidence interval. A confidence interval gives information about the position of the mean value of the parameter whereas a tolerance interval gives information about the position of the parameter and the probability for this parameter to be in this interval. Let Y be the random variable which represents this parameter and X the covariate (temperature, pressure,...). To take these covariate into account in the evaluation of the tolerance interval of Y , the conditional distribution of Y given X is studied. When X is in \mathbb{R} , the conditional quantiles of Y given X are used to build tolerance curves and when X is in \mathbb{R}^p , they are used to build tolerance hyper surfaces. Finally when $Y \in \mathbb{R}^q$, several parameters are studied simultaneously and multivariate or spatial conditional quantiles are used to build a tolerance region.

Three types of modeling can be used to define conditional univariate or multivariate quantiles. Parametric modeling has the advantage of giving results easy to interpret but the parametric shape of the conditional distribution of Y given X has to be predetermined. Non-parametric modeling is more flexible because it relaxes the constraint on the conditional distribution. However, in practice the results are difficult to interpret. Semi-parametric modeling is therefore a compromise between these two types of modeling and gives results easy to interpret. Real indices $X'\beta$ are indeed incorporated in order to reduce the dimension of the explicative part of the model and no parametric structure is imposed on the link between Y and $X'\beta$.

The choice of parametric, non-parametric or semi-parametric modeling is thus a key point in the estimation of tolerance curves, hyper surfaces or regions.

3.2.2. Estimation

Non-parametric conditional quantiles estimation is usually based on kernel or local polynomial estimation methods (see for instance [71], [65]) and suffers, like local smoothing methods, from what is called the curse of dimensionality. Indeed, when the dimension p of the covariate X increases, the dispersion of the data increases and the quality of the estimation decreases. Another drawback is that graphical representation is possible only when the dimension of X is equal to 1 or 2: tolerance curves are obtained in 2D when $X \in \mathbb{R}$ and tolerance hyper surfaces are obtained in 3D when $X \in \mathbb{R}^2$.

To avoid these two drawbacks, semi-parametric modeling can be used. The following two-step (one parametric and one non-parametric) methodology for conditional quantiles estimation is proposed. First of all, the Euclidean parameter β used to reduce X to $X'\beta$ is estimated. Next, the functional parameters used in the non-parametric conditional quantile estimation are estimated. More precisely the semi-parametric approach combines the SIR (sliced inverse regression) method and the kernel estimation of the conditional quantile. This methodology allows graphical representation of the tolerance curves in 2D ($X'\beta \in \mathbb{R}$) or surfaces in 3D ($X'\beta \in \mathbb{R}^2$), with the index $X'\beta$ easy to interpret.

We want to focus our attention on dimension-reduction approach for problems of quality control (sliced inverse regression approaches, clustering of variables, ...). We still work on non-parametric and semi-parametric estimation of tolerance curves and hyper-surfaces, via estimation of (multivariate) conditional quantiles. Moreover, another point of interest is the introduction of recursive methods in the estimation process in order

to be able to manage the data stream, and at present we develop recursive methods in semi-parametric model for the estimation of tolerance curves.

4. Application Domains

4.1. Application Domains

The following examples illustrate the importance of dependability and safety in various fields.

A first example concerns oil production in deep water. We have already worked with IFREMER on the reliability of oil rigs and with IFP (French Oil Institute) on risk assessment and control for extraction of hydrocarbon substances from submarine deposits hard to work due to their depth.

A second example in the military field concerns combat aircraft with "relaxed static stability". These aircraft are slightly aerodynamically unstable by design: they will quickly depart from level and controlled flight unless the pilot constantly works to keep it in trim. While this enhances maneuverability, it is very wearing on a pilot relying on a mechanical flight control system. Hence, the aircraft is highly vulnerable to sensors or on-board calculator breakdown.

A third example deals with quality control linked with biomedical and biometric studies led by CERIES (Centre de Recherches et Investigations Epidermiques Sensorielles): the research center of CHANEL on human skin. The knowledge of tolerance curves for numerous skin biophysical parameters is crucial for CERIES in so far as it enables CHANEL chemists to develop new cosmetic products better adapted to the aimed target: elderly women on the Asian market, young Caucasian or African-American women. Thus, knowing the skin features of a person is enough to decide whether or not they fit in the reference limits of the various CHANEL cosmetic products.

A last example concerns the development of air quality control strategies, which is a wide preoccupation for human health. In order to achieve this purpose, air pollution sources have to be accurately identified and quantified. We have already worked in 2007-2008 on a scientific project initiated by the French ministry of Ecology and Sustainable Development. We have also worked on the statistical and quality control parts of a study financed by VNF (Voies Navigables de France) concerning a satisfaction survey of sailors on the "canal des deux mers" in south of France. An other possible application concerns the use of tolerance curves in industrial quality control process. We already had discussions with Danone on this subject.

5. Software

5.1. Software

Most of the statistical methods for dimension reduction and quality control have been implemented in R : variables clustering (Chavent, Kuentz), cluster-SIR (Kuentz, Saracco), bootstrap choice of parameters for SIR related methods (Liquet, Saracco), geometric multivariate (conditional) quantile (Chaouch, Saracco), recursive SIR (Bercu, Nguyen, Saracco), sample selection models (Chavent, Liquet, Saracco), bagging SIR (Kuentz, Liquet, Saracco). Development of an *R*-package is in progress.

6. New Results

6.1. Average Continuous Control of Piecewise Deterministic Markov Processes

Participant: François Dufour.

The long run average continuous control problem of piecewise deterministic Markov processes (PDMP's) taking values in a general Borel space and with compact action space depending on the state variable is investigated. The control variable acts on the jump rate and transition measure of the PDMP, and the running and boundary costs are assumed to be positive but not necessarily bounded. As far as we are aware of, this is the first time that this kind of problem is considered in the literature. Indeed, results are available for the long run average cost problem but for impulse control see O.L.V. Costa [46], D. Gatarek [62] and the book by M.H.A. Davis [54] (see the references therein). On the other hand, the continuous control problem has been studied only for discounted costs by A. Almudevar [40], M.H.A. Davis [53], [54], M.A.H. Dempster and J.J. Ye [56], [57], L. Forwick, M. Schäl, and M. Schmitz [60], M. Schäl [74], A.A. Yushkevich [75], [76].

This work is under revision for SIAM Journal of Control and Optimization [48] The following results have been obtained in collaboration with Oswaldo Luis Do Valle Costa from Escola Politécnica da Universidade de São Paulo, Brazil. In particular,

- It has been obtained an optimality equation for the long run average cost in terms of a discrete-time optimality equation related to the embedded Markov chain given by the post-jump location of the PDMP.
- The existence of a feedback measurable selector for the discrete-time optimality equation has been shown by establishing a connection between this equation and an integro-differential equation
- Sufficient conditions has been derived for the existence of a solution for a discrete-time optimality inequality and an ordinary optimal feedback control for the long run average cost using the so-called vanishing discount approach [66].
- Two examples are presented based on the capacity expansion model, analyzed in [55], [54], and by the authors from the stability point of view in [49], [58]. The first example illustrates the fact that the technical assumptions of the paper are satisfied and so there exists a solution for the optimality inequality. The second example shows that by using our new result (see Proposition 8.7), there exists a solution for the average-cost optimality equation.

6.2. The Vanishing Discount Approach for the Average Continuous Control of Piecewise Deterministic Markov Processes

Participant: François Dufour.

This work is concerned with the existence of an optimal control strategy for the long run average continuous control problem of piecewise deterministic Markov processes (PDMP's). It has been accepted for publication in Journal of Applied Probability [20]. These results have been obtained in collaboration with Oswaldo Luis Do Valle Costa from Escola Politécnica da Universidade de São Paulo, Brazil. The present work can be seen as a continuation of the results derived in [48]. In [48], sufficient conditions were derived to ensure the existence of an optimal control by using the vanishing discount approach. These conditions were mainly expressed in terms of the relative difference of the α -discount value functions. The main goal of this paper is to derive tractable conditions directly related to the primitive data of the PDMP to ensure the existence of an optimal control. Our main assumptions are written in terms of some integro-differential inequalities related to the so-called expected growth condition, and geometric convergence of the post-jump location kernel associated to the PDMP.

6.3. The Policy Iteration Algorithm for Average Continuous Control of Piecewise Deterministic Markov Processes

Participant: François Dufour.

The present work [50] is a continuation of a series of papers by the authors: [48] [20]. Part of this work will be presented at the 48th IEEE Conference on Decision and Control, Shanghai, China, 2009, [32]. These results have been obtained in collaboration with Oswaldo Luis Do Valle Costa from Escola Politécnica da Universidade de São Paulo, Brazil. The main goal of this paper is to apply the so-called policy iteration algorithm (PIA) for the long run average continuous control problem of piecewise deterministic Markov processes taking values in a general Borel space and with compact action space depending on the state variable. The PIA has received considerable attention in the literature and consists of three steps: initialization, policy evaluation, which is related to the Poisson equation (PE) associated to the transition law defining the Markov decision process, and policy improvement. In our context, the policy evaluation step is connected to a kind of PE which we call a pseudo-Poisson equation. This equation is clearly different from a classical PE encountered in the literature of the discrete-time Markov control processes. However, although different, we can show in section that this pseudo-Poisson equation still has the good properties that we might expect to satisfy in order to guarantee the convergence of the policy iteration algorithm. These results are not straightforward to obtain due to the specific structure of this discrete-time optimality equation. Finally, the PIA is studied in details. It is first shown that the convergence of the PIA to a solution satisfying the optimality equation holds under some classical hypotheses. In the sequence it is shown that this optimal solution yields to an optimal control strategy for the average control problem for the continuous-time PDMP in a feedback form.

6.4. Optimal stopping of Piecewise Deterministic Markov Processes

Participants: Benoîte de Saporta, François Dufour, Karen Gonzalez.

We have developed a computational method for optimal stopping of a piecewise deterministic Markov processes by using a quantization technique for Markov chains, see [44]. Optimal stopping problems have been studied for PDMP's in [47], [51], [54], [61], [63], [70]. In [63] the author defines an operator related to the first jump time of the process, which we shall call first jump operator, and shows that the value function of the optimal stopping problem is a fixed point for this operator. Based on a probabilistic interpretation for this jump operator, we designed a numerical scheme to approximate the value function. The originality of our work is two-fold. On the one hand, instead of using a fixed time-discretization grid for our continuous-time process, we first compute the quantization of an underlying discrete-time Markov chain which naturally appears in our problem, and only then derive path-adapted time-discretization grids. On the other hand, some functionals of the underlying Markov chains in our jump operator are not Lipschitz-continuous as they have jumps (they are typically indicatrices of some thresholds). We proved the convergence of our procedure and computed an error-bound. We implemented our procedure on a simple example and obtained very good results. This work will appear in *Annals of Applied Probability* [25]. Parts of this work have been presented at the conferences *Optimal Stopping and their Applications* [35] in Turku, Finland in June 2009 and *International Conference on Analysis and Design of Hybrid Systems* [34] in Zaragoza, Spain in September.

6.5. Modeling of crack propagation by Piecewise Deterministic Markov Processes

Participants: Romain Azais, Anne Gégout-Petit.

We have already shown that Piecewise Deterministic Markov Processes (PDMP) are very suitable for the modeling of reliability problems, see for example [80] and [77].

In this work, we study the relevance of PDMP's in an other context of application: the modeling of crack propagation.

The well known empirical Paris-Erdogan law $\frac{da}{dt} = C(\Delta K(a))^m$ has been adapted in order to take into account the random nature of propagation crack even in a controlled environment. For some authors C and m are random variable, for other there are random processes. We propose here to use PDMP for the modeling of the evolution of a crack length : during a random time T the length a evolves according to a deterministic evolution equation $\frac{da}{dt} = C_1(\Delta K(a))^{m_1}$. At time T new parameter (m_2, C_2) are chosen through a Markov matrix and the new evolution of the length a is $\frac{da}{dt} = C_2(\Delta K(a))^{m_2}$. The parameters of the propagation (m_i, C_i) belong to a finite set. We use the model to predict the evolution of the propagation using only the 10 first measures of the crack. This method consists of reducing the number of possible regimes (m_i, C_i) in keeping the nearest of the beginning of the crack.

We have fitted this model with the Virkler data. 68 replicate constant amplitude crack propagation tests were conducted on 2024-T3 aluminum alloy. The data were recorded starting at a crack length of 9 mm to a final length of 49.8 mm and it produced 164 data points for each replicate. For each replicate, we have chosen by simulated annealing the best vector (m_1, C_1, T, m_2, C_2) such that the theoretical crack is near the experimental one. We use empirical statistics on the 68 vectors (m_1, C_1, T, m_2, C_2) to grade the model.

The results of simulation are very satisfactory. The set of simulated cracks include the experimental one even with only 4 regimes. The results of the prediction are also excellent. 64 of the 68 crack are well predicted with the 10 first measures.

This work is a part of a technical report with EADS Astrium on some mechanical and mathematical aspects of the crack propagation in alloy [43] and is submitted for publication [42].

6.6. Asymptotic behavior of bifurcating autoregressive processes

Participants: Bernard Bercu, Benoîte de Saporta, Anne Gégout-Petit.

Bifurcating autoregressive (BAR) processes are an adaptation of autoregressive (AR) processes to binary tree structured data. They were first introduced by Cowan and Staudte [52] for cell lineage data, where each individual in one generation gives birth to two offspring in the next generation. Cell lineage data typically consist of observations of some quantitative characteristic of the cells over several generations of descendants from an initial cell. BAR processes take into account both inherited and environmental effects to explain the evolution of the quantitative characteristic under study.

There are several results on statistical inference and asymptotic properties of estimators for BAR models in the literature. For maximum likelihood inference on small independent trees, see Huggins and Basawa [67]. For maximum likelihood inference on a single large tree, see Huggins [68] for the original BAR model, Huggins and Basawa [67] for higher order Gaussian BAR models, and Zhou and Basawa [82] for exponential first-order BAR processes. We also refer the reader to Zhou and Basawa [81] for the LS parameter estimation. In all those papers, the BAR process is supposed to be stationary. In Guyon [64], the LS estimator is also investigated, but the process is not stationary, and the author makes intensive use of the tree structure and Markov chain theory. We extended our previous results to p -th order ($p \geq 2$) BAR processes. More precisely, we have carried out a sharp analysis of the asymptotic properties of the least squares (LS) estimators of the unknown parameters of p -th order BAR processes via a martingale approach, based on the generation-wise filtration. More precisely, we have established the almost sure convergence of our LS estimators with a sharp rate of convergence, together with the quadratic strong law and the central limit theorem.

This work is published in the Electronic Journal of Probability [14], and was presented in several seminars in France, as well as at the conference Mathematical models for cell division [37] held in Paris in march 2009.

6.7. On the usefulness of persistent excitation in ARX adaptive tracking

Participant: Bernard Bercu.

The usefulness of persistent excitation is well-known in the control community. Thanks to a persistently excited adaptive tracking control, we show that it is possible to avoid the strong controllability assumption recently proposed by the authors for multivariate ARX models. We establish the almost sure convergence for both least squares and weighted least squares estimators of the unknown parameters. A central limit theorem and a law of iterated logarithm are also provided. All this asymptotical analysis is related to the Schur complement of a suitable limiting matrix.

This work has been recently submitted to International Journal of Control, and it will be presented at the 48^e IEEE Conference on Decision and Control, Shanghai, China, 2009.

6.8. A new concept of strong controllability via the Schur complement in adaptive tracking

Participant: Bernard Bercu.

We propose a new concept of strong controllability related to the Schur complement of a suitable limiting matrix. This new notion allows us to extend the previous convergence results associated with multidimensional ARX models in adaptive tracking. On the one hand, we carry out a sharp analysis of the almost sure convergence for both least squares and weighted least squares algorithms. On the other hand, we also provide a central limit theorem and a law of iterated logarithm for these two algorithms. Our asymptotic results are illustrated by numerical simulations.

This work has been recently submitted to Automatica [45], and it was presented at the 47^e IEEE Conference on Decision and Control, Cancun, Mexico, 2008.

6.9. On the almost sure central limit theorem for vector martingales : convergence of moments and statistical applications

Participant: Bernard Bercu.

The aim of this work, published in [11], is to investigate the almost sure asymptotic properties of vector martingale transforms. Assuming some appropriate regularity conditions both on the increasing process and on the moments of the martingale, we prove that normalized moments of any even order converge in the almost sure central limit theorem for martingales. A conjecture about almost sure upper bounds under wider hypotheses is formulated. The theoretical results are supported by examples borrowed from statistical applications, including linear autoregressive models and branching processes with immigration, for which new asymptotic properties are established on estimation and prediction errors.

6.10. A functional central limit theorem for a class of interacting Markov chain Monte Carlo methods

Participant: Bernard Bercu.

This paper was published in [12] by two authors of ALEA and CQFD Teams. In this work, we present a functional central limit theorem for a new class of interaction Markov chain Monte Carlo algorithm. This stochastic algorithms have been recently introduced to solve non-linear measure-valued equations. We provide an original stochastic analysis based on semigroup techniques on distribution spaces and fluctuation theorems for self-interaction random fields. Additionally, we also present a series of sharp mean error bounds in terms of the semigroup associated with the first order expansion of the limiting measure-valued process. We illustrate our results in the context of Feynman-Kac semigroups.

6.11. Almost Sure Stabilization for Feedback Controls of Regime-switching Linear Systems with A Hidden Markov Chain

Participants: Bernard Bercu, François Dufour.

This work [13] is devoted to the almost sure stabilization of adaptive control systems that involve an unknown Markov chain. The control system displays continuous dynamics represented by differential equations and discrete events given by a hidden Markov chain. In the previous investigation on this class of problems, averaging criteria were used, which provides only the system behavior in some expectation sense. Different from previous work on stabilization of adaptive controlled systems with a hidden Markov chain, where average criteria were considered, this work focuses on the almost sure stabilization or sample path stabilization of the underlying processes. Under simple conditions, it is shown that as long as the feedback controls have linear growth in the continuous component, the resulting process is regular. Moreover, by appropriate choice of the Lyapunov functions, it is shown that the adaptive system is stabilizable almost surely. As a by-product, it is also established that the controlled process is positive recurrent.

6.12. Dimension reduction via Sliced Inverse Regression

Participants: Bernard Bercu, Marie Chavent, Vanessa Kuentz, Thi Mong Ngoc Nguyen, Jérôme Saracco.

Concerning the dimension reduction framework which is a useful tool for the quality control part of the project, some results have been obtained for Sliced Inverse Regression (SIR) and related methods such as Sliced Average Variance Estimation (SAVE).

SIR and related approaches were introduced in order to reduce the dimensionality of regression problems. In general semiparametric regression framework, these methods determine linear combinations of a set of explanatory variables \mathbf{X} related to the response variable Y , without losing information on the conditional distribution of Y given \mathbf{X} . They are based on a “slicing step” in the population and sample versions. They are sensitive to the choice of the number H of slices, and this is particularly true for SIR-II and SAVE methods. At the moment there are no theoretical results nor practical techniques which allows the user to choose an appropriate number of slices. In [24], we propose an approach based on the quality of the estimation of the effective dimension reduction (EDR) space: the square trace correlation between the true EDR space and its estimate can be used as goodness of estimation. We introduce a naïve bootstrap estimation of the square trace correlation criterion to allow selection of an “optimal” number of slices. Moreover, this criterion can also simultaneously select the corresponding suitable dimension K (number of the linear combination of \mathbf{X}). From a practical point of view, the choice of these two parameters H and K is essential. We propose a 3D-graphical tool, implemented in R, which can be useful to select the suitable couple (H, K) . In this article, we focus on the SIR, SIR-II and SAVE methods. We indicate how the proposed criterion can be used in practice. A simulation study is performed to illustrate the behaviour of this approach and the need for selecting properly the number H of slices and the dimension K . The R codes are available.

In [22], we propose bagging versions of SIR which consist in using bootstrap replications of the data set and in aggregating the corresponding estimators. We give the asymptotic distribution of the Bagging-SIR estimator. A simulation study is used to compare the numerical performance of the proposed alternative bagging versions of SIR with the classical SIR approach. The benefits of these methods are significant for noisy models and when the sample size is small. The R codes are available.

In the theory of sufficient dimension reduction, SIR is a famous technique that enables to reduce the dimensionality of regression problems by determining linear combinations of a p -dimensional explanatory variable \mathbf{X} related to a response variable Y . However it is based on a crucial condition on the marginal distribution of the predictor \mathbf{X} , often called the linearity condition. From a theoretical and practical point of view, this condition appears to be a limitation. Using an idea of Li et al. (2004) in the Ordinary Least Squares framework, we propose in [23] to cluster the predictor space so that the linearity condition approximately holds in the different partitions. Then we apply SIR in each cluster and finally estimate the dimension reduction subspace by combining these individual estimates. We give asymptotic properties of the corresponding estimator. We show with a simulation study that the proposed approach, referred as cluster-based SIR, improves the estimation of the e.d.r. basis. We also propose an iterative implementation of cluster-based SIR and show in simulations that it increases the quality of the estimator. Finally the methodology is applied on the horse mussel data and the comparison of the prediction reached on test samples shows the superiority of cluster-based SIR over SIR.

Most of the common estimation methods for sample selection models rely heavily on parametric and normality assumptions. We consider in [17] a multivariate semiparametric sample selection model and develop a geometric approach to the estimation of the slope vectors in the outcome equation and in the selection equation. Contrary to most existing methods, we deal symmetrically with both slope vectors. Moreover, the estimation method is link-free and distribution-free. It works in two main steps: a multivariate sliced inverse regression step, and a canonical analysis step. We establish \sqrt{n} -consistency and asymptotic normality of the estimates. We describe how to estimate the observation and selection link functions. The theory is illustrated with a simulation study.

In [27], we propose a recursive estimator for the matrices of interest in SIR approach. Moreover, when the number H of slices is equal to two, we obtain a recursive estimator of the direction of the parameter θ in the semiparametric regression model $Y = f(X'\theta, \varepsilon)$ which does not require the estimation of the link function f . We give asymptotic results for this estimator. A simulation study illustrates the good numerical behaviour of this estimator for moderate sample size even if the dimension of the covariate X is important.

6.13. Nonparametric conditional quantile estimation

Participant: Jérôme Saracco.

In a multidimensional setting, lack of objective basis for ordering multivariate observations is a major problem in extending the notion of quantiles. Conditional quantiles are required in various biomedical or industrial problems. Numerous alternative definitions of (conditional) quantile for multidimensional variables, have been proposed in statistical literature. In [26] we focus on the notion of geometric quantile and conditional geometric quantile, based on the minimization of a loss function. Asymptotic results has been obtained and an implementation in R allowed us to show the good numerical performances of the proposed estimators in practice on simulated and real datasets.

6.14. Semiparametric Mode Regression Estimation

Participant: Jérôme Saracco.

It has been found that, for a variety of probability distributions, there is a surprising linear relation between mode, mean and median. In this paper, the relation between mode, mean and median regression functions is assumed to follow a simple parametric model. In [21], we propose a semiparametric conditional mode (mode regression) estimation for an unknown (unimodal) conditional distribution function in the context of regression model, so that any m -step-ahead mean and median forecasts can then be substituted into the resultant model to deliver m -step-ahead mode prediction. In the semiparametric model, Least Squared Estimator (LSEs) for the model parameters and the simultaneous estimation of the unknown mean and median regression functions by the local linear kernel method are combined to infer about the parametric and nonparametric components of the proposed model. The asymptotic normality of these estimators is derived, and the asymptotic distribution of the parameter estimates is also given and is shown to follow usual parametric rates in spite of the presence of the nonparametric component in the model. These results are applied to obtain a data-based test for the dependence of mode regression over mean and median regression under a regression model.

6.15. Dimension Reduction for categorical data

Participants: Marie Chavent, Vanessa Kuentz, Jérôme Saracco.

Clustering of variables.

Clustering of variables is studied as a way to arrange variables into homogeneous clusters, thereby organizing data into meaningful structures. Once the variables are clustered into groups such that variables are similar to the other variables belonging to their cluster, the selection of a subset of variables is possible. Several specific methods have been developed for the clustering of numerical variables. However concerning categorical variables, much less methods have been proposed. In [15], we extend the criterion used by Vigneau and Qannari (2003) in their Clustering around Latent Variables approach for numerical variables to the case of categorical data. The homogeneity criterion of a cluster of categorical variables is defined as the sum of the correlation ratio between the categorical variables and a latent variable, which is in this case a numerical variable. We show that the latent variable maximizing the homogeneity of a cluster can be obtained with Multiple Correspondence Analysis. Different algorithms for the clustering of categorical variables are proposed: iterative relocation algorithm, ascendant and divisive hierarchical clustering. The proposed methodology is illustrated by a real data application to satisfaction of pleasure craft operators.

Rotation in Multiple Correspondence Analysis

Multiple Correspondence Analysis (MCA) is a well-known multivariate method for statistical description of categorical data. Similarly to what is done in Principal Component Analysis (PCA) and Factor Analysis, the MCA solution can be rotated to increase the components simplicity. The idea behind a rotation is to find subsets of variables which coincide more clearly with the rotated components. This implies that maximizing components simplicity can help in factor interpretation and in variables clustering. In [16], we propose a two-dimensional analytic solution for rotation in MCA. Similarly to what is done by Kaiser (1958) for PCA, this planar solution is computed in a practical algorithm applying successive pairwise planar rotations for optimizing the rotation criterion. This criterion is a varimax-based one relying on the correlation ratio between the categorical variables and the MCA components. A simulation study is used to illustrate the proposed solution. An application on a real data set shows the possible benefits of using rotation in MCA.

7. Contracts and Grants with Industry

7.1. Electricité de France

Participants: Bernard Bercu, Marie Chavent, François Dufour, Anne Gégout-Petit, Frédéric Proia, Jérôme Saracco.

An important contract was signed in June 2009 between EDF R&D team ICAME, and INRIA teams ALEA and CQFD. It deals with new statistical methods for recursive adaptive prediction of the electricity consumption of EDF customers. The contract was divided into five lots and it will last until 2011. The first lot was delivered in November 2009. The total value of the contract is 140000 euros.

7.2. Voies Navigables de France

Participants: Marie Chavent, Vanessa Kuentz, Jérôme Saracco.

We are working on the statistical and quality control parts (sampling and data collection, survey weights, sampling errors) of a study financed by VNF (Voies Navigables de France) concerning a satisfaction survey of sailors on the “canal des deux mers” in south of France (total value of the contract: 6000 euros).

7.3. EADS Astrium

Participants: Adrien Brandejsky, Benoîte de Saporta, François Dufour, Anne Gégout-Petit, Huilong Zhang.

The goal of this project is to propose and study an approach to evaluate the probability of occurrence of events defined by the crossing of a threshold. For this we have started collaborating with Marie Touzet of the LMP (Laboratoire de Mécanique Physique). Together, we have investigated the different laws adapted to the propagation of a crack in alloys. We have proposed the modeling of the evolution of a crack length by Piecewise Deterministic Markov Processes. This work will give rise to the technical report [43].

7.4. National Initiatives

J erome Saracco is the leader of a research project financed by the R egion Aquitaine for three years (2007-2009), named *Estimation recursive pour des mod eles semiparam etriques en Statistique* with a total amount of 120 000 euros including the PhD-grant of Thi Mong Ngoc Nguyen.

The goal of the project "FAUTOCOES" (number ANR-09-SEGI-004) of the ARPEGE program of the French National Agency of Research (ANR) can be described as follows. Today, complex technological processes must maintain an acceptable behavior in the event of random structural perturbations, such as failures or component degradation. Aerospace engineering provides numerous examples of such situations: an aircraft has to pursue its mission even if some gyroscopes are out of order, a space shuttle has to succeed in its re-entry trip with a failed on-board computer. Failed or degraded operating modes are parts of an embedded system history and should therefore be accounted for during the control synthesis.

These few basic examples show that complex systems like embedded systems are inherently vulnerable to failure of components and their reliability has to be improved through fault-tolerant control. Embedded systems require mathematical representations which are in essence dynamic, multi-model and stochastic. This increasing complexity poses a genuine scientific challenge:

- to model explicitly and realistically the dynamical interactions existing between the physical state variables defining the system: pressure, temperature, flow rate, intensity, etc, and the functional and dysfunctional behavior of its components;
- to estimate the performance of the system through the evaluation of reliability indexes such as availability, quality, and safety;
- to optimize the control to prevent system failures, as well as to maintain the system function when a failure has occurred.

Our aim is to meet the previously mentioned challenge by using the framework of piecewise deterministic Markov processes (PDMP's in short) with an emphasis on probabilistic and deterministic numerical methods. More precisely, our objectives are

- to use the framework of piecewise deterministic Markov processes to model complex physical systems and phenomena;
- to compute expectations of functionals of the process in order to evaluate the performance of the system;
- to develop theoretical and numerical control tools for PDMP's to optimize the performance and/or to maintain system function when a failure has occurred.

7.5. International Initiatives

Bernard Bercu is the leader of the ECOS Action between France and Mexico, entitled *Probl emes de rupture et le contr ole adaptatif pour les mod eles de r egression*. The Universities of Bordeaux 1, Toulouse 3, Paris 5, Mexico, and Puebla are involved in this action for four years.

8. Dissemination

8.1. Animation of the scientific community

B. Bercu belongs to the MAS thematic group of the SMAI.

B. Bercu is an assistant director of the Institute of Mathematics of Bordeaux. He is also a member of the IMB council as well as of the UFR council of the University of Bordeaux

B. Bercu is the director of applied mathematic department of the University of Bordeaux.

M. Chavent is webmaster of the web site of the SFDS (Société Française de Statistique).

M. Chavent is a member of the scientific committee of the following conferences EGC' 09.

M. Chavent is vice-secretary and member of the administration council of the SFDS (Société Française de Statistique).

B. de Saporta belongs to the board of SMAI-MAS group. She is webmaster of their website.

B. de Saporta is a member of the organizing committee of enigmath 2008 and 2009, a free online mathematical quizz. <http://enigmath.org>

B. de Saporta presented an animation at Cap Science for the *fête de la science* 2009.

F. Dufour is associate editor of the journal : SIAM Journal of Control and Optimization since 2009.

F. Dufour is member of the IFAC Technical Committee TC 1.4 Stochastic Systems, term Period 2008-2011.

F. Dufour is member of the scientific council of the engineering school ENSEIRB-MATMECA.

A. Gégout-Petit is in charge to promote the "Licence MASS" (Applied mathematics degree) of the University of Bordeaux 2 to the secondary school pupils.

A. Gégout-Petit is an elected member of the administration council of the SFdS.

A. Gégout-Petit is member of the administration team of the web domain emath.fr

A. Gégout-Petit is an elected member of the CEVU (Conseil des Etudes et de la Vie Universitaire) of the Bordeaux 2 University

J. Saracco is a member of the administration council of the University of Bordeaux 4.

J. Saracco is a reviewer for Computational Statistics and Data Analysis, Journal of Multivariate Analysis, The Annals of Statistics, Statistica Sinica, Biometrika, Journal of the American Statistical Association.

8.2. Workshop organisation

The team CQFD was very involved in the organization of the "**41e Journées de Statistique**" Annual event of the "Société Française de Statistique" which took place in Bordeaux in 2009.

The teams ALEA and CQFD, as well as INRIA Bordeaux Sud-Ouest, are very involved in the organization of the "8e Journées MAS de la SMAI", <http://www.math.u-bordeaux1.fr/MAS10/> It is the most important bi-annual workshop for french probabilists and statisticians. It will take place in Bordeaux in september 2010. The subject of this workshop will be stochastic algorithms and combinatorics.

B. Bercu was the organizer of the "**Fifth Meeting of Mathematical Statistics**" between the Universities of Bordeaux, Montpellier, Santander, Valladolid, Pau, and Toulouse in June 2009.

B. de Saporta co-organized with Christian Paroissin a workshop on applied mathematics in dependability and safety

8.3. Participation to congresses, conferences, invitations

B. Bercu gave contributed talks at

- Orsay University, February 2009
- University of l'Aquila, Italia, March 2009
- Bordeaux University, June 2009
- Journées Algorithmes stochastiques, November 2009
- Rennes University, December 2009
- 48^e IEEE CDC Conference, Shanghai, December 2009

M. Chavent gave contributed talks at

- IFCS 2009 (11th Conference of the International Federation of Classification Societies), Dresden, Germany, March 2009.
- 41èmes journées de Statistique, Bordeaux, May 2009.
- XIIIth Applied Stochastic Models and Data Analysis (ASMDA2009) International Conference, Vilnius, Lithuania, July 2009.
- SFC 2009 (XVIèmes Rencontres de la Société Francophone de Classification), Grenoble, France, September 2009.

B. de Saporta gave contributed talks at

- the seminar of probability and statistics of Lille 1 university
- the seminar of applied mathematics of IMB
- the workshop Mathematical models for cell division at IHP Paris
- the symposium on Optimal stopping with applications in Turku, Finland
- the 3rd IFAC Conference on Analysis and Design of Hybrid Systems, Zaragoza, Spain.

F. Dufour spent two weeks at Escola Politecnica, Universidade de Sao Paulo, Brasil (June, 2009) working with O.L.V Costa.

F. Dufour spent two weeks at University of Wisconsin, Milwaukee, USA in February and October 2009 working with R. Stockbridge.

F. Dufour gave a contributed talk at 48th IEEE CDC Conference, Shanghai, December 2009.

F. Dufour chairs the 48th CDC/28th CCC session "Stochastic Control I".

A. Gégout-Petit gave contributed talks at

- 41èmes journées de Statistique, Bordeaux, May 2009.
- Causal modeling workshop, Oslo, september 2009.

M. Saracco gave contributed talks at

- IFCS 2009 (11th Conference of the International Federation of Classification Societies), Dresden, Germany, March 2009.
- 41èmes journées de Statistique, Bordeaux, May 2009.
- XIIIth Applied Stochastic Models and Data Analysis (ASMDA2009) International Conference, Vilnius, Lithuania, July 2009.
- Troisièmes Rencontres des Jeunes Statisticiens (sous l'égide de la SFDS), Aussois, France, September 2009.

8.4. Teaching

B. Bercu, A. Brandejsky, M. Chavent, B. de Saporta, F. Dufour, A. Gégout-Petit, J. Saracco, and H. Zhang teach graduate probability and Statistics and in the cursus "Statistique et Fiabilité" (Statistic and Reliability) of the Master "Ingénierie Mathématique Statistique et Economique" at the Universities of Bordeaux 1, 2, 4, and Institut Polytechnique de Bordeaux.

M. Chavent and A. Gégout-Petit teach Statistic in licence MASS of University of Bordeaux 2.

A. Gégout-Petit is an academic tutor for students work experiences in company or Research Organism.

B. de Saporta teaches undergraduate mathematics and postgraduate probability and finance at the university of economics Montesquieu Bordeaux 4.

J. Saracco teaches statistics in Magistère d'Économie et de Finance Internationale (MAGEFI), University Bordeaux 4 and linear algebra and Statistics in Licence of Economics, University Bordeaux 4.

9. Bibliography

Major publications by the team in recent years

- [1] B. BERCU. *On the convergence of moments in the almost sure central limit theorem for martingales with statistical applications*, in "Stochastic Process. Appl.", vol. 111, n^o 1, 2004, p. 157–173.
- [2] B. BERCU, F. DUFOUR, G. YIN. *Almost Sure Stabilization for Feedback Controls of Regime-switching Linear Systems with A Hidden Markov Chain*, in "IEEE Trans. on Automatic Control", vol. 54, n^o 9, 2009, p. 2114–2125 US .
- [3] B. BERCU, B. PORTIER. *Kernel density estimation and goodness-of-fit test in adaptive tracking*, in "SIAM J. on Control and Optimization", vol. 47, n^o 5, 2008, p. 2440–2457.
- [4] B. BERCU, B. DE SAPORTA, A. GÉGOUT-PETIT. *Asymptotic analysis for bifurcating autoregressive processes via a martingale approach*, in "Electronic Journal of Probability", vol. 14, 2009, p. 2492–2526.
- [5] M. CHAVENT, B. LIQUET, J. SARACCO. *A semiparametric approach for multivariate sample selection model*, in "Statistica Sinica", 2009, To appear.
- [6] O. L. D. V. COSTA, F. DUFOUR. *Stability and ergodicity of piecewise deterministic Markov processes*, in "SIAM J. Control Optim.", vol. 47, n^o 2, 2008, p. 1053–1077.
- [7] O. L. D. V. COSTA, F. DUFOUR. *The Vanishing Discount Approach for the Average Continuous Control of Piecewise Deterministic Markov Processes*, in "Journal of Applied Probability", vol. 46, n^o 4, 2009, <http://arxiv.org/abs/0812.0820>, To appear BR .
- [8] A. GANNOUN, J. SARACCO, A. YUAN, G. E. BONNEY. *Non-parametric quantile regression with censored data*, in "Scand. J. Statist.", vol. 32, n^o 4, 2005, p. 527–550.
- [9] J. SARACCO. *Asymptotics for pooled marginal slicing estimator based on SIR_α approach*, in "J. Multivariate Anal.", vol. 96, n^o 1, 2005, p. 117–135.
- [10] B. DE SAPORTA, F. DUFOUR, K. GONZALEZ. *Numerical method for optimal stopping of piecewise deterministic Markov processes*, in "Annals of Applied Probability", 2009, To appear.

Year Publications

Articles in International Peer-Reviewed Journal

- [11] B. BERCU, P. CÉNAC, G. FAYOLLE. *On the almost sure central limit theorem for vector martingales: convergence of moments and statistical applications*, in "J. Appl. Probab.", vol. 46, n^o 1, 2009, p. 151–169, <http://dx.doi.org/10.1239/jap/1238592122>.

- [12] B. BERCU, P. DEL MORAL, A. DOUCET. *A functional central limit theorem for a class of interacting Markov chain Monte Carlo methods*, in "Electron. J. Probab.", vol. 14, n^o 73, 2009, p. 2130–2155 (electronic).
- [13] B. BERCU, F. DUFOUR, G. YIN. *Almost Sure Stabilization for Feedback Controls of Regime-switching Linear Systems with A Hidden Markov Chain*, in "IEEE Trans. on Automatic Control", vol. 54, n^o 9, 2009, p. 2114–2125 US .
- [14] B. BERCU, B. DE SAPORTA, A. GÉGOUT-PETIT. *Asymptotic analysis for bifurcating autoregressive processes via a martingale approach*, in "Electronic Journal of Probability", vol. 14, 2009, p. 2492–2526.
- [15] M. CHAVENT, V. KUENTZ, J. SARACCO. *Clustering of categorical variables around latent variables*, 2009, Submitted paper.
- [16] M. CHAVENT, V. KUENTZ, J. SARACCO. *Rotation in Multiple Correspondence Analysis: a planar rotation iterative procedure*, 2009, Submitted paper.
- [17] M. CHAVENT, B. LIQUET, J. SARACCO. *A semiparametric approach for multivariate sample selection model*, in "Statistica Sinica", 2009, To appear.
- [18] D. COMMENGES, A. GÉGOUT-PETIT. *A general definition of influence between stochastic processes*, in "LIDA", 2009, To appear.
- [19] D. COMMENGES, A. GÉGOUT-PETIT. *A general dynamical statistical model with possible causal interpretation*, in "JRSS B", vol. 71, 2009, p. 719-736.
- [20] O. L. D. V. COSTA, F. DUFOUR. *The Vanishing Discount Approach for the Average Continuous Control of Piecewise Deterministic Markov Processes*, in "Journal of Applied Probability", vol. 46, n^o 4, 2009, <http://arxiv.org/abs/0812.0820>, To appear BR .
- [21] A. GANNOUN, J. SARACCO, Y. KEMING. *On Semiparametric Mode Regression Estimation*, in "Communications in Statistics - Theory and Methods", 2009, To appear GB .
- [22] V. KUENTZ, B. LIQUET, J. SARACCO. *Bagging versions of Sliced Inverse Regression*, in "Communications in Statistics - Theory and Methods", 2009, To appear.
- [23] V. KUENTZ, J. SARACCO. *Cluster-based Sliced Inverse Regression*, in "Journal of the Korean Statistical Society", 2009, To appear.
- [24] B. LIQUET, J. SARACCO. *A graphical tool for selecting the number of slices and the dimension of the model in SIR and SAVE approaches*, in "Computational Statistics", 2009, In revision.
- [25] B. DE SAPORTA, F. DUFOUR, K. GONZALEZ. *Numerical method for optimal stopping of piecewise deterministic Markov processes*, in "Annals of Applied Probability", 2009, To appear.

Articles in National Peer-Reviewed Journal

- [26] M. CHAOUCH, A. GANNOUN, J. SARACCO. *Estimation des quantiles géométriques conditionnels et non conditionnels*, in "Journal de la Société Française de Statistique", vol. 150, n^o 2, 2009, p. 1–27.

- [27] T. M. N. NGUYEN, J. SARACCO. *Estimation récursive en régression inverse par tranches (sliced inverse regression)*, in "Journal de la Société Française de Statistique", 2009, In revision.

International Peer-Reviewed Conference/Proceedings

- [28] M. CHAVENT, V. KUENTZ, J. SARACCO. *A partitioning method for the clustering of categorical variables*, in "Proceedings of the IFCS'2009: Classification as a Tool for Research, Dresde", H. LOCAREK-JUNGE, C. WEIHS (editors), Springer, 2009.
- [29] M. CHAVENT, V. KUENTZ, J. SARACCO. *A two-dimensional analytic solution for rotation in Multiple Correspondence Analysis*, in "ASMDA 2009, Vilnius", 2009.
- [30] M. CHAVENT, V. KUENTZ, J. SARACCO. *Une méthode de partitionnement pour la classification de variables qualitatives*, in "SFC 2009, Grenoble", 2009.
- [31] M. CHAVENT, B. LIQUET, J. SARACCO. *Données manquantes en ACM : l'algorithme NIPALS*, in "SFC 2009, Grenoble", 2009.
- [32] O. L. D. V. COSTA, F. DUFOUR. *The Vanishing Approach for the Average Continuous Control of Piecewise Deterministic Markov Processes*, in "Proceedings of the Conference on Decision and Control, Cancun, Mexico", 2009 BR .
- [33] B. LIQUET, J. SARACCO. *A criterion for selecting the number of slices and the dimension of the model in SIR and SAVE approaches*, in "ASMDA 2009, Vilnius", 2009.
- [34] B. DE SAPORTA, F. DUFOUR, K. GONZALEZ. *Numerical method for optimal stopping of hybrid processes*, in "Preprints of the 3rd IFAC International Conference on Analysis and Design of Hybrid Systems 3rd IFAC International Conference on Analysis and Design of Hybrid Systems, Zaragoza Espagne", 2009, p. 114-119.
- [35] B. DE SAPORTA, F. DUFOUR, K. GONZALEZ. *Numerical method for optimal stopping of piecewise deterministic Markov processes*, in "OPTIMAL STOPPING WITH APPLICATIONS, Turku Finlande", 06 2009.

National Peer-Reviewed Conference/Proceedings

- [36] M. CHAVENT, V. KUENTZ, J. SARACCO. *Une solution analytique pour la rotation planaire en Analyse Factorielle des Correspondances Multiples*, in "SFdS 2009, Bordeaux", 2009.

Workshops without Proceedings

- [37] B. DE SAPORTA, B. BERCU, A. GÉGOUT-PETIT. *Asymptotic behavior of bifurcating autoregressive processes*, in "Mathematical models for cell division, Paris France", 03 2009.
- [38] B. DE SAPORTA, F. DUFOUR. *Numerical method for optimal stopping of piecewise deterministic Markov Processes*, in "Cinquième Rencontre de Statistiques Mathématiques Bordeaux-Santander-Toulouse-Valladolid, Le Teich France", 06 2009.

Scientific Books (or Scientific Book chapters)

- [39] C. BLANCHET-SCALLIET, R. GIBSON BRANDON, B. DE SAPORTA, D. TALAY, E. TANRÉ. *Viscosity Solutions to Optimal Portfolio Allocation Problems in Models with Random Time Changes and Transaction*

Costs, in "Advanced Financial Modelling", H. ALBRECHER, W. J. RUNGALDIER, W. SCHACHERMAYER (editors), Radon Series on Computational and Applied Mathematics, vol. 8, de Gruyter, 2009 CH .

References in notes

- [40] A. ALMUDEVAR. *A Dynamic Programming Algorithm for the Optimal Control of Piecewise Deterministic Markov Processes*, in "SIAM J. of Control and Optim.", vol. 40, n^o 2, 2001, p. 525-539.
- [41] T. AVEN, U. JENSEN. *Stochastic Models in Reliability*, Springer Verlag, New York, 1999.
- [42] R. AZAIS, A. GÉGOUT-PETIT, M. TOUZET. *Estimation, simulation et prévision d'un modèle de propagation de fissures par des processus markoviens déterministes par morceaux*, 2010, Submitted.
- [43] R. AZAIS, A. GÉGOUT-PETIT, M. TOUZET, B. DE SAPORTA, F. DUFOUR. *Aspects aléatoires de la propagation de fissures*, Equipe Projet Contrôle de Qualité et Fiabilité Dynamique (CQFD), INRIA Bordeaux, EADS Astrium, 2009, Technical report.
- [44] V. BALLY, G. PAGÈS. *A quantization algorithm for solving multi-dimensional discrete-time optimal stopping problems*, in "Bernoulli", vol. 9, n^o 6, 2003, p. 1003–1049.
- [45] B. BERCU, V. VAZQUEZ. *A new concept of strong controllability via the Schur complement in adaptive tracking*, in "Submitted to Automatica", 2008.
- [46] O. L. D. V. COSTA. *Average impulse control of piecewise deterministic processes*, in "IMA J. Math. Control Inform.", vol. 6, n^o 4, 1989, p. 375–397.
- [47] O. L. D. V. COSTA, M. DAVIS. *Approximations for optimal stopping of a piecewise-deterministic process*, in "Math. Control Signals Systems", vol. 1, n^o 2, 1988, p. 123–146.
- [48] O. L. D. V. COSTA, F. DUFOUR. *Average Control of Piecewise Deterministic Markov Processes*, in "Under review in SIAM Journal on Control and Optimization", 2008, <http://arxiv.org/abs/0809.0477>.
- [49] O. L. D. V. COSTA, F. DUFOUR. *Stability and ergodicity of piecewise deterministic Markov processes*, in "SIAM J. Control Optim.", vol. 47, n^o 2, 2008, p. 1053-1077.
- [50] O. L. D. V. COSTA, F. DUFOUR. *The Policy Iteration Algorithm for Average Continuous Control of Piecewise Deterministic Markov Processes*, 2009, <http://arxiv.org/abs/0902.2673>.
- [51] O. L. D. V. COSTA, C. A. B. RAYMUNDO, F. DUFOUR. *Optimal stopping with continuous control of piecewise deterministic Markov processes*, in "Stochastics Stochastics Rep.", vol. 70, n^o 1-2, 2000, p. 41–73.
- [52] R. COWAN, R. G. STAUDTE. *The bifurcating autoregressive model in cell lineage studies*, in "Biometrics", vol. 42, 1986, p. 769–783.
- [53] M. DAVIS. *Control of piecewise-deterministic processes via discrete-time dynamic programming*, in "Stochastic differential systems (Bad Honnef, 1985), Berlin", Lecture Notes in Control and Inform. Sci., vol. 78, Springer, 1986, p. 140–150.

- [54] M. DAVIS. *Markov Models and Optimization*, Chapman and Hall, London, 1993.
- [55] M. DAVIS, M. DEMPSTER, S. P. SETHI, D. VERMES. *Optimal capacity expansion under uncertainty*, in "Adv. in Appl. Probab.", vol. 19, n^o 1, 1987, p. 156-176.
- [56] M. DEMPSTER, J. YE. *Necessary and Sufficient Optimality Conditions for Control of Piecewise Deterministic Processes*, in "Stochastic and Stochastics Reports", vol. 40, 1992, p. 125-145.
- [57] M. DEMPSTER, J. YE. *Generalized Bellman-Hamilton-Jacob Optimality Conditions for a Control Problem with Boundary Conditions*, in "Appl. Math. Optimization", vol. 33, 1996, p. 211-225.
- [58] F. DUFOUR, O. L. D. V. COSTA. *Stability of piecewise-deterministic Markov processes*, in "SIAM J. Control Optim.", vol. 37, n^o 5, 1999, p. 1483–1502.
- [59] F. DUFOUR, Y. DUTUIT. *Dynamic Reliability: A new model*, in "Proceedings of $\lambda\mu$ 13-ESREL2002, Lyon, France, Mars", 2002.
- [60] L. FORWICK, M. SCHÄL, M. SCHMITZ. *Piecewise deterministic Markov control processes with feedback controls and unbounded costs*, in "Acta Appl. Math.", vol. 82, n^o 3, 2004, p. 239–267.
- [61] D. GATAREK. *On first-order quasi-variational inequalities with integral terms*, in "Appl. Math. Optim.", vol. 24, n^o 1, 1991, p. 85–98.
- [62] D. GATAREK. *Impulsive control of piecewise-deterministic processes with long run average cost*, in "Stochastics Stochastics Rep.", vol. 45, n^o 3-4, 1993, p. 127–143.
- [63] U. S. GUGERLI. *Optimal stopping of a piecewise-deterministic Markov process*, in "Stochastics", vol. 19, n^o 4, 1986, p. 221–236.
- [64] J. GUYON. *Limit theorems for bifurcating Markov chains. Application to the detection of cellular aging*, in "Ann. Appl. Probab.", vol. 17, n^o 5-6, 2007, p. 1538–1569.
- [65] W. HENDRICKS, R. KOENKER. *Hierarchical spline models for conditional quantiles and the demand for electricity*, in "Journal of the American Statistical Association", vol. 99, 1992, p. 58-68.
- [66] O. HERNÁNDEZ-LERMA, J. LASSERRE. *Discrete-time Markov control processes: Basic optimality criteria*, Applications of Mathematics, vol. 30, Springer-Verlag, New York, 1996.
- [67] R. M. HUGGINS, I. V. BASAWA. *Extensions of the bifurcating autoregressive model for cell lineage studies*, in "J. Appl. Probab.", vol. 36, n^o 4, 1999, p. 1225–1233.
- [68] R. M. HUGGINS. *Robust inference for variance components models for single trees of cell lineage data*, in "Ann. Statist.", vol. 24, n^o 3, 1996, p. 1145–1160.
- [69] P.-E. LABEAU, Y. DUTUIT. *Fiabilité dynamique et disponibilité de production : un cas illustratif*, in "Proceedings of $\lambda\mu$ 14, Bourges, France, October", 2004, p. 431-436.

-
- [70] S. LENHART, Y. C. LIAO. *Integro-differential equations associated with optimal stopping time of a piecewise-deterministic process*, in "Stochastics", vol. 15, n^o 3, 1985, p. 183–207.
- [71] L. MAGEE, J. BURBIDGE, A. ROBB. *Computing kernel-smoothed Conditional Quantiles from Many Observations*, in "Journal of the American Statistical Association", vol. 86, 1991, p. 673–677.
- [72] J. MALATRONAS, J. PERKINS, R. SMITH. *The availability of a system of two unreliable machines connected by an intermediate storage tank*, in "IIE Transactions", vol. 15, n^o 3, 1983, p. 195–201.
- [73] A. SCHORNAGEL. *Optimum preventive maintenance for complex systems with interdependent equipment units*, in "Amsterdam, The Netherlands: Shell Res. Lab.", 1987.
- [74] M. SCHÄL. *On piecewise deterministic Markov control processes: control of jumps and of risk processes in insurance*, in "Insurance Math. Econom.", vol. 22, n^o 1, 1998, p. 75–91.
- [75] A. YUSHKEVICH. *Verification theorems for Markov decision processes with controlled deterministic drift and gradual and impulsive controls.*, in "Theory Probab. Appl.", vol. 34, n^o 3, 1989, p. 474–496.
- [76] A. YUSHKEVICH. *The compactness of a policy space in dynamic programming via an extension theorem for Carathéodory functions*, in "Math. Oper. Res.", vol. 22, n^o 2, 1997, p. 458–467.
- [77] H. ZHANG, F. DUFOUR, Y. DUTUIT, C. ELEGBEDE. *Application des processus déterministes par morceaux à un système de production pétrolière offshore*, in "Proceedings of $\lambda\mu$, Avignon, France", 2008.
- [78] H. ZHANG, F. DUFOUR, I. FARES, Y. DUTUIT. *Fiabilité dynamique: Outils analytiques et numériques*, in "Proceedings of $\lambda\mu$ 15, Lille, France, October", 2006.
- [79] H. ZHANG, K. GONZALEZ, F. DUFOUR, Y. DUTUIT. *Piecewise Deterministic Markov Processes and Dynamic Reliability*, in "Proceedings of Mathematical Methods in Reliability, Glasgow, Ecosse, July", 2007.
- [80] H. ZHANG, K. GONZALEZ, F. DUFOUR, Y. DUTUIT. *Piecewise Deterministic Markov Processes and Dynamic Reliability*, in "The Journal of Risk and Reliability", vol. 222, n^o 4, 2008, p. 545–551.
- [81] J. ZHOU, I. V. BASAWA. *Least-squares estimation for bifurcating autoregressive processes*, in "Statist. Probab. Lett.", vol. 74, n^o 1, 2005, p. 77–88.
- [82] J. ZHOU, I. V. BASAWA. *Maximum likelihood estimation for a first-order bifurcating autoregressive process with exponential errors*, in "J. Time Ser. Anal.", vol. 26, n^o 6, 2005, p. 825–842.