SOME CONTRIBUTIONS TO NON LINEAR ESTIMATION AND INTERACTING PARTICLE SYSTEMS

With applications to non linear filtering, optimal control, Feynman-Kac formulae and Schrödinger operators.

Pierre DEL MORAL
Chargé de recherches au C.N.R.S.
Laboratoire de Statistique et Probabilités
de l'Université Paul Sabatier-Toulouse III

Introduction

Filtering/signal processing and optimal control are two basic and central problems in engineering sciences and applied mathematics. Although this thesis goes belong these two subjects they are the starting point and the corner stone of most of my research.

There exist many ways including functional, spectral analysis or geometric approaches to enter into these questions but I found probability theory more appealing and flexible to drive my investigations. Thus, filtering and control will be discussed in a probabilistic framework following the development of the four next directions of research:

1. Feynman-Kac formulae and interacting particle systems.
2. Genetic algorithms and global optimization.
3. Contraction properties of Markov operators and non linear semigroups.
4. Idempotent probability calculus.

This thesis is essentially concerned with Feynman-Kac formulae and branching/interacting particle approximating systems. My principal contribution in this aera was to introduce new continuous or discrete time particle models and to analyze their asymptotic behavior as the size of the systems and/or the time horizon tends to infinity.
In contrast to statistical physics and fluid mechanics the use of particle methods here is not related to a particle microscopic interpretation of a precise physical quantity. These new particle strategies rather belong the class of evolutionary and genetic branching type algorithms. They can also be regarded as a particle simulation technique of conditional distributions or as a stochastic and adaptative grid technique or as a stochastic global search algorithm.

We have studied the performance and the convergence of these particle schemes through differents angles.
According various mathematical and probabilistic technics, we have analyzed weak laws of large numbers and $\mathbb{L}_p$-norm estimates, central limit theorems and large deviations on path space and for the particle density profiles, the convergence of empirical processes indexed by some class of functions and uniform estimates and convergence theorems with respect to the time parameter as well as propagation of chaos for increasing blocks of particles or time horizons, ... Most of these developments are collected in a survey paper in "Le séminaire de Probabilités" [42].

Besides the modeling and the numerical analysis of these schemes one of the main objective of my work was also to throw some new light on interesting links between sometimes too disconnected physical, mathematical and engineering domains. To motivate this synthesis we already present a non exhaustive list of the principal subjects that will be discussed in the forthcoming development:

- *Non linear estimation*: prediction/filtering/smoothing in signal processing, path estimation, sampling from a conditional distribution, optimal control and non linear regulation problems.

- *Genetic algorithms:* evolutionary algorithms, population evolution dynamics, historical processes and genealogical trees.

- *Physics:* generalized Boltzmann equations, distributions of a Markov particle in an environment with obstacles. spectral analysis of Schrödinger operators and Feynman-Kac semigroups.

- *Measure valued processes:* general discrete time models, Moran type interacting particle systems, branching Dawson-Watanabe and Fleming-Viot superprocesses.

- *Stochastic differential equations:* Zakai and Kushner-Stratonovitch equations and robust filtering equations.

The connections between these subjects and the applications of interacting particle systems in non linear estimation problems including the numerical solving of pure jump or McKean-Vlasov equations and discrete time models as well as discrete time models and non linear regulation problems are discussed in a forthcoming textbook [5]. This study also present new results concerning increasing propagation of chaos and large deviations principles for pure jump or McKean-Vlasov mean field particle models.

The second theme is concerned with genetic algorithms and global optimization. Although the resulting genetic models are strongly related to the ones discussed in filtering their asymptotic analysis is not related to Monte-Carlo and mean field interacting particle methods. Here we are concerned with the long time behavior of the systems with fixed population size. We introduce a cooling schedule and we analyze the dynamics of the resulting Markov chain with rare transitions using semigroup technics, relative entropy and log-Sobolev constants. On the other hand we also propose a stability criterium and a test set strategy to analyze the set of extrema of a virtual energy function.

In another research direction we have studied mean cost optimization problems. To this end we have proposed and analyzed a class of generalized simulated annealing algorithms related to stochastic and non homogeneous potentials. Finally we have presented different models including a natural and very flexible genetic scheme with random sequences of mutations and selections.

The third theme of this thesis is concerned with the regularity of Markov operators and more generally non linear Feynman-Kac-Schrödinger semigroups. Although this subject was primarily influenced by investigations on the stability of the filtering equation this theme is based to some extend on Dobrushin's ideas related to the stability of Markov semigroups.

- In the first part we discuss the contraction properties of non linear Feynman-Kac semigroups for the total variation distance using Dobrushin's ergodic coefficient. Our strategy is to analyze the contraction properties of a given Markov semigroup under a new reference probability defined in terms of the underlying Feynman-Kac formulae.

- In the second part we analyze the contraction of integral operators on the set of probabilities with respect to an abstract and general class of ϕ-relative entropies, including classical $\mathbb{L}_p$-norms, the total variation and the Kakutani-Hellinger distance and the Kullback or Havrda-Charvat entropy.
  In contrast to classical studies on the stability of Markov chains here we do not use (whenever it exists) the invariant probability of the underlying Markov transition. We propose an universal estimate of the ϕ-contraction coefficient in terms of Dobrushin's ergodic coefficient. Finally for $\mathbb{L}_p$-norms or Kullback entropies we estimate the contraction in terms of spectral gaps and classical or modified log-Sobolev constants.

The forth and last theme is directly related to my PhD dissertation. It concerns the development of an idempotent probability calculus which gives a natural probabilistic interpretation of deterministic optimal control problems. This novel approach allows to transfer probabilistic intuition to optimal control and it also induces new analysis and approximating technics.
In this context the weak solution analog of the evolution equation of a Markov process distributions provides an alternative approach of the viscosity solution of the Hamilton-Jacobi equation associated to a decision process. This study gives a natural framework to analyze some image erosions and dilations in mathematical morphology.
The probabilistic description of the invariant measures of Bellman transitions leads to explicit representations of the fixed points of non expansive mappings. The analog of Dynkin's formula also allows to estimate the quality of some classes of admissible controls.

In the next development of this synthesis we present our principal contributions in each four themes. We also place with details our studies in the existing literature and we underline in the description of each article our main advances on each subject. To illustrate these descriptions and to motivate this thesis we shed some light on some models and we briefly present a part of the results we obtained.

A part of the material presented in this thesis results from various collaborations with René Carmona, Himilcon Carvalho, Michel Doisy, Leila Kallel, Michael Kouritzin, Michel Ledoux, André Monin, Jean-Charles Noyer, Philip Protter, John Rowe, Gérard Salut and Tim Zajic and more particularly with Dan Crisan, Alice Guionnet, Jean Jacod, Terry Lyons and Laurent Miclo.

Contents

1 Feynman-Kac formulae and particle systems

This thesis is mainly concerned with the analysis and the particle approximations of Feynman-Kac formulae of the form

$$\eta_t(f) = \gamma_t(f)/\gamma_t(1) \qquad \text{with} \qquad \gamma_t(f) = \mathbb{E}(f(X_t)\ Z_t(X)), \qquad t \in I$$

and whether the time parameter is discrete $n \in I = \mathbb{N}$ or continuous $t \in I = \mathbb{R}_+$

$$Z_n(X) = \prod_{p=0}^{n-1} g_n(X_p) \qquad \text{or} \qquad Z_t(X) = \exp \int_0^t V_s(X_s)\ ds.$$

Unless otherwise mentioned g_n and V_s are assumed to be positive functions and X is a Markov process on some measurable state space. Most of the results we obtained are presented in the survey paper [42] and in the more recent articles [40, 45, 44]. Notice that starting from the discrete or continuous formula

$$\gamma_n(f) = \eta_n(f) \times \prod_{p=0}^{n} \eta_p(g_p) \quad \text{ou} \quad \gamma_t(f) = \eta_t(f) \times \exp \int_0^t \eta_s(V_s)\ ds,$$

it is possible to use the approximating measures η_s, $s \leq t$, to compute the positive distributions γ_t. By construction we can prove that the resulting interacting particle approximating models will be un-bias.

1.1 Stochastic particle methods

Our research in this field has started essentially with the two studies [16, 18]. They present a novel interacting particle system approximating model for solving discrete time non linear filtering problems. In this framework we recall that η_t represents the conditional distribution of the state of the signal X_t with respect to a sequence of observations $(Y_s, s \leq t)$. The former articles also give the first rigorous and mathematically well founded proof of the convergence of the heuristic schemes presented in [19, 49] and independently by Gordon, Salmon et Smith in [58] and more recently by Kitagawa in [62].

In this first analysis we propose a weak law of large numbers and $\mathbb{L}_2$-mean error estimates. Furthermore the martingales and the non homogeneous semigroups identified in [16] turn out to be essential to study the long time behavior and the fluctuations of the corresponding interacting particle schemes. Beside these analytic foundations these two articles also propose an original interpretation of the discrete and non linear filtering equation in terms of a genetic algorithm as the size of the population tends to infinity.

These first theoretical works also complete the technical studies [46, 49, 50, 51] in which I was involved during my PhD in the L.A.A.S.-C.N.R.S. of Toulouse. The latter technical reports concern application of these particle methods to radar, sonar or gps ("global positioning system") signal processing. These applications were also discussed in two articles in engineering science literature on signal processing [48] and aerospace science [6].

These interacting particle approximating schemes can be regarded as a stochastic and adaptative grid approximation technic of non linear filtering problems.
Thus they provide a numerically efficient alternative to more classical Monte-Carlo methods based on independent particles weighted by their likelihood functions. The latter were presented and studied in some details in my PhD dissertation and in the corresponding articles [17, 52].
The numerical analysis of these non interacting particle schemes are much more easier. In [17] we introduce a regularization strategy and a first observability condition underwhich a uniform convergence theorem holds with respect to the time parameter. In a second article [22] we use the stability properties of the optimal filter obtained by Ocone and Pardoux in [69] to control uniformly in time the errors in $\mathbb{L}_2$-norms only in terms of the ergodicity of the signal process.
In a recent study [65] Kurtz and Xiong extend these non interacting schemes to solve on every compact interval more general stochastic partial differential equations with cylindric and infinite dimensional noise.

The article [16] is also the origin of the design of an universal particle strategy to approximate discrete time and non linear dynamical systems in distribution space of the form

$$\eta_n = \Phi_n(\eta_{n-1})$$

where Φ_n is a mapping from the set of all probability into itself. We approximate as $N \to \infty$ the desired distributions η_n with the empirical measures $\eta_n^N = \frac{1}{N}\sum_{i=1}^N \delta_{\xi_n^i}$ associated to a sequence of interacting particle systems. These systems are defined recursively in time according the following rules:

$$\xi_n = (\xi_n^1, \dots, \xi_n^N) \ \text{ i.i.d.}|\xi_{n-1} \ \sim \ \Phi_n(\eta_{n-1}^N).$$

The foundations and the first numerical analysis of such general and abstract particle models including the weak law of large numbers and some local fluctuations and exponential estimates are presented in [21].

When we consider the dynamical system associated to the previously defined Feynman-Kac formulae the above particle approximating model is a genetic algorithm with a two steps mutation/selection transition

$$\xi_n \xrightarrow{\text{Selection}} \widehat{\xi}_n \xrightarrow{\text{Mutation}} \xi_{n+1}.$$

Given the system ξ_n at time n, the intermediate population $\widehat{\xi}_n = (\widehat{\xi}_n^1, \ldots, \widehat{\xi}_n^N)$ consists of N independent and identically distributed random variables with common discrete distribution

$$\sum_{i=1}^{N} \frac{g_n(\xi_n^i)}{\sum_{j=1}^{N} g_n(\xi_n^j)} \; \delta_{\xi_n^i}.$$

During the mutation each particle evolves according the transition of the underlying process X.

In most of our studies the Markov process X takes values in an abstract measurable space. This general framework allows us to treat path-valued Markov processes. In this situation the resulting particle algorithms are based on path selections and the mutations can be regarded as an elementary deformation of the path. This essential observation is partly described in [42]. In [45] it is used to analyze the asymptotic behavior of genealogical trees. In the latter referenced paper we also show that the historical process of a genetic algorithm solves non linear smoothing and path estimation problems.

Even if non linear filtering is one the main application and motivatation of most of my research our articles are context free. We discuss abstract and quite general Feynman-Kac models. Therefore, their use and interpretation may differ from one domain to another.

- In the discrete time case if we suppose the mappings $g_n = 1_{A_n}$ are indicatrices of some measurable sets A_n then with some usual precautions we have that

  $$\eta_n(dx) = \mathbb{P}(X_n \in dx | X_0 \in A_0, \ldots, X_{n-1} \in A_{n-1})$$

  and in the path space interpretation proposed in [45] we also have that

  $$\begin{aligned} &\eta_n(d(x_0, \ldots, x_n)) \\ &= \mathbb{P}((X_0, \ldots, X_n) \in d(x_0, \ldots, x_n) | (X_0, \ldots, X_n) \in (A_0 \times \ldots \times A_{n-1})). \end{aligned}$$

  The studies [32, 33] allows to treat non negative functions g_n which may equal to zero at some point . In the above situation the corresponding interacting particle approximating schemes behave as an acceptation/rejection type algorithm. It is also possible to regularize the functions replacing the indicatrix by some positive approximating functions. To analyze this situation we can use the continuity and robustness results presented in [30].

- The previously defined Feynman-Kac formulae also appear to be useful to describe the distribution laws of a Markov particle evolving in an environment with obstacles. In these settings we rather assume that V_t are non positive functions. Here η_t represents the conditional distribution of the particle X_t

killed at rate $|V_t|$ given the fact that it is still alive at that time. If T denotes its killing time then we have that

$$\eta_t(dx) = \mathbb{P}(X_t \in dx | T > t) \quad \text{and} \quad \mathbb{P}(T > t) = \exp \int_0^t \eta_s(V_s)\, ds.$$

- In the article [44] we show how the above technics can be used to compute the Lyapunov exponent and related spectral quantities of Feynman-Kac-Schrödinger semigroups. To illustrate this study and in continuous time settings we notice that the Lyapunov exponent $\lambda(V)$ of the semigroup

$$P_t^v(f)(x) = \mathbb{E}_x(f(X_t)\ Z_t(X))$$

associated to an homogeneous potential $V_t = V$ coincide with the top of the spectrum of the Schrödinger operator $L^v = L + V$. It can also be rewritten in the form $\lambda(V) = \lim_{t\to\infty} \sup_x \lambda_t^x(V)$ with

$$\lambda_t^x(V) = \frac{1}{t} \log \mathbb{E}_x(\exp[\int_0^t V(X_s) ds]) = \frac{1}{t} \int_0^t \eta_s^{(x)}(V) ds$$

where $\eta_t^{(x)}$ stands for the Feynman-Kac distribution flow with initial condition $\eta_0 = \delta_x$. According to this formulation it is natural to approximate $\lambda(V)$ with the particle Lyapunov exponents

$$\lambda_t^{(x,N)}(V) = \frac{1}{t} \int_0^t \eta_s^{(x,N)}(V)\ ds.$$

where $\eta_t^{(x,N)} = \frac{1}{N} \sum_{i=1}^N \delta_{\xi_t^i}$ denote the N-approximating measures with all particles ξ_t^i starting at $\xi_0^i = x$. When the underlying process X is sufficiently mixing we can find some finite constant $c < \infty$ and some exponent $\beta \in (0,1]$ such that

$$\sup_{t \geq N^{\beta/2}} \mathbb{E}(|\lambda_t^{(x,N)}(V) - \lambda(V)|^2) \leq c/N^\beta.$$

1.2 Discrete filtering and particle models

The asymptotic behavior of these new discrete time and interacting particle systems is studied in a chain of articles.

- In [26] we present large deviation principles on path space and for the particle density profile for two type of branching models. In the path space situation the one step mappings of the limiting system Φ_n are chosen so that $\Phi_n(\mu) << \Phi_n(\eta_{n-1})$ for any μ. The main simplification due to this condition is that the distribution law $\mathbb{P}_{[0,n]}^{(N)}$ of the particle systems $\xi_{[0,n]} = (\xi_0, \ldots, \xi_n)$ is absolutely continuous with respect to the tensor product $\eta_{[0,n]}^{\otimes N}$ of the Mckean measures

$$\eta_{[0,n]} = \eta_0 \otimes \ldots \otimes \eta_n$$

To each $x=(x_0,\ldots,x_n)$, $x_p\in E^N$, we associate the empirical measures $m(x)$ and $m(x_p)$ by setting

$$m(x)=\frac{1}{N}\sum_{i=1}^N \delta_{(x_0^i,\ldots,x_n^i)}\quad\text{and}\quad m(x_p)=\frac{1}{N}\sum_{i=1}^N \delta_{x_p^i}$$

and for any probability measure μ on path space we define the mapping

$$\Phi_{[0,n]}(\mu)=\eta_0\otimes\Phi_1(\mu_0)\otimes\ldots\otimes\Phi_n(\mu_{n-1})$$

where μ_n stands for the n-time marginal of μ. With these notations the Radon-Nykodim derivative of $\mathbb{P}^{(N)}_{[0,n]}$ with respect to $\eta^{\otimes N}_{[0,n]}$ is now given by the formula

$$\frac{d\mathbb{P}^{(N)}_{[0,n]}}{d\eta^{\otimes N}_{[0,n]}}(x)=\exp H_n^{(N)}(x)$$

with

$$\begin{aligned} H_n^{(N)}(x) &= N\sum_{p=1}^n \int \log\frac{d\Phi_p(m(x_{p-1}))}{d\Phi_p(\eta_{p-1})}\,dm(x_p)\\ &= N\,F_n(m(x)),\quad\text{and}\quad F_n(\mu)=\int\log\frac{d\Phi_{[0,n]}(\mu)}{d\Phi_{[0,n]}(\eta_{[0,n]})}\,d\mu. \end{aligned}$$

In some regular situations we combine Varadhan's integral lemma with cut off technics and we prove that the distribution laws of the empirical measures on path space

$$\eta^N_{[0,n]}=m(\xi_{[0,n]})=\frac{1}{N}\sum_{i=1}^N\delta_{(\xi_0^i,\ldots,\xi_n^i)}$$

satisfy a large deviations principle with good rate function

$$\mathrm{Ent}(\mu|\Phi_{[0,n]}(\mu))=\mathrm{Ent}(\mu|\eta_{[0,n]})-F_n(\mu).$$

The continuity condition can be relaxed using Laplace's method when studying large deviations for the particle density profiles.

- In the study [27] we extend Shiga and Tanaka article [70] to non linear filtering. For regular signals we obtain a central limit theorem on path space for the interacting particle approximating systems. More precisely we prove that the $\mathbb{L}^2(\eta_{[0,n]})$-random fields defined by

$$W^{(N)}_{[0,n]}=\sqrt{N}\ (\eta^N_{[0,n]}-\eta_{[0,n]})$$

weakly converge, in the sense of finite distributions, to a centered Gaussian field $W_{[0,n]}$ with covariance function given by

$$\mathbb{E}(W_{[0,n]}(\varphi_1)W_{[0,n]}(\varphi_2))=(C_n(\varphi_1),C_n(\varphi_2))_{\mathbb{L}^2(\eta_{[0,n]})}$$

where C_n denoted an integral operator on $\mathbb{L}^2(\eta_{[0,n]})$.
The convergence analysis and the precise description of operators C_n are in part based on the same changes of probability presented in our study of large deviations. Their main advantage is to deal and work with independent particles. With above notations we can write the characteristic function of the random field $W^{(N)}_{[0,n]}$ as follows

$$\mathbb{E}_{\mathbb{P}^{(N)}_{[0,n]}}(e^{i\ W^{(N)}_{[0,n]}(\varphi)}) = \mathbb{E}_{\eta^{\otimes N}_{[0,n]}}(e^{i\ W^{(N)}_{[0,n]}(\varphi)+H^{(N)}_n(\xi_0,\dots,\xi_n)}).$$

Using uniform integrability arguments their convergence, as $N\to\infty$, is reduced to the weak convergence of $H^{(N)}_n(\xi_0,\dots,\xi_n)$ under the corresponding tensor product measure.
An asymptotic expansion of $H^{(N)}_n$ combined with Dynkin-Mandelbaum's theorem [56] on symmetric statistics and a formula of Shiga-Tanaka allows to identify C_n and to conclude that

$$\lim_{N\to\infty}\mathbb{E}_{\mathbb{P}^{(N)}_{[0,n]}}(e^{i\ W^{(N)}_{[0,n]}(\varphi)}) = e^{-\frac{1}{2}\|C_n(\varphi)\|^2_{\mathbb{L}^2(\eta_{[0,n]})}}.$$

- The long time behavior and the first uniform estimates with respect to the time parameter are presented in the articles [28, 29]. The original and key idea is to connect the stability properties of the limiting system with the long time behavior of the particle systems. We prove that if the limiting system is sufficiently stable then

  $$\lim_{N\to\infty}\ \sup_{n\geq 0}\mathbb{E}(|\eta^N_n(f)-\eta_n(f)|)=0.$$

  For exponentially stable dynamics we conclude that the order of convergence is $\mathrm{O}(N^{\alpha/2})$ for some explicitly defined $\alpha\in(0,1]$. For sufficiently mixing signals we show in [42] that we can chose $\alpha=1$.

  In filtering settings the semigroup technics described in the previously referenced papers also allows to connect the ergodicity properties of the signal with the asymptotic stability properties of the optimal filter. We will discuss this connection in more details in the second part of this thesis related to the contraction properties of Markov operators and non linear Feynman-Kac semigroups.

- In [36] we extend a part of our results to empirical processes indexed by a class of functions $\mathcal{F}$ and for Zolotarev type seminorms

  $$\|\mu-\nu\|_{\mathcal{F}} = \sup\{\ |\mu(f)-\nu(f)|\ ;\ f\in\mathcal{F}\}.$$

We prove the analog for interacting systems of Glivenko-Cantelli and Donsker's theorem as well as Talagrand's exponential estimates for polynomial Vapnik-Cervonenkis' classes of functions. To throw some light on this work we obtain for any regular class of function $\mathcal{F}$

$$\lim_{N\to\infty} \|\eta_n^N - \eta_n\|_{\mathcal{F}} = 0 \quad \text{et} \quad \sup_{n\geq 0} \mathbb{E}(\|\eta_n^N - \eta_n\|_{\mathcal{F}}) = \mathrm{O}(N^{\alpha/2}).$$

The last uniform estimate with respect to the time parameter holds when the limiting system satisfies the exponential stability condition studied in [28, 29]. Finally by inverting explicitly an integral operator on some particular class of tensor product functions we describe explicitely in [36] the covariance function associated to the fluctuations on path space in terms of the central semigroups exhibited in [16].
We will use in [42] this analytic representation to identify the key martingales which will lead to the fluctuations of the particle density profiles without any mixing condition.

- A part of the work [45] is concerned with propagations of chaos for increasing particle blocks sizes and time horizon parameters. We propose several estimates with respect to the total variation distance, the relative entropy or $\mathbb{L}_p$-norms. These results allows to calibrate the independence between the particles paths. For instance, if we denote by $\mathbb{P}^{(N,q)}_{[0,n]}$ the law of the first q paths $(\xi_0^i,\ldots,\xi_n^i)_{1\leq i\leq q}$ then, under some regularity conditions on Φ_n we have that

$$\lim_{N\to\infty} q(N)n(N)/N = 0 \implies \lim_{N\to\infty} \mathrm{Ent}(\ \mathbb{P}^{(N,q(N))}_{[0,n(N)]} \mid \eta^{\otimes q(N)}_{[0,n(N)]}\ ) = 0.$$

- In [31] we discuss and compare, in a particular linear Gaussian situation, interacting approximating models and Monte-Carlo methods with independent particles and unregularized weights. By an explicit computation of the covariance function associated to the fluctuations we prove the superiority of interacting approximating models.

At the same period we have proposed and studied several particle models including new branching rules or performing path-particles. On one hand these studies are intended to speed up the time execution of the algorithms on a computer. On the other hand the use of path-particles also allows to solve important non linear smoothing and path-estimation problems.

- In a first study [13] we introduce and analyze a general and abstract interacting particle approximating model which includes branching rules of several types: multinomial, Poisson, Bernoulli, binomial...
To throw some light on these models we point out here that they are all based on the same idea. Namely, they are intended to approximate a discrete probability measure $\sum_{i=1}^N\ p_i\ \delta_{x^i}$ by another discrete measure $\sum_{i=1}^N\ (M_i/N)\ \delta_{x^i}$

with weights multiples of $1/N$.
In genetics interpretation the random numbers M_i correspond to the number of particles which have selected the site x^i. In terms of branchings M_i can also be regarded as the number of offsprings created or the number of individuals borned at that site.
In this connection, the selection stage of the previously defined genetic algorithm correspond to the branching rule:

$$(M_1, \ldots, M_N) = \text{Multinomial}(N; p_1, \ldots, p_N).$$

We can also choose the random numbers independently one each other according the Poisson or Bernoulli distributions:

$$\begin{aligned} \mathbb{P}(M_i = m) &= e^{-Np_i} \frac{(Np_i)^m}{m!} \\ \mathbb{P}(M_i = [Np_i] + 1) &= 1 - \mathbb{P}(M_i = [Np_i]) = \{Np_i\}. \end{aligned}$$

It is useful to notice that if we use the former two branching rules then the resulting particle scheme is no more conservative and the size of the population becomes an integer valued martingale. Other more performing variants including periodic selection, regularized mutations, or Baker's type conservative branching rules are also presented in [42].

In [13] we also underline some connections between these particle approximating schemes and the branching Dawson-Watanabe and Fleming-Viot superprocesses. We extend the results of Etheridge and March [57] to our discrete time and non linear models. We prove that the genetic algorithm can also be viewed as an particle model with Poisson branching conditioned to fixed population size. We also present a weak law of large numbers and the first proof of the consistency of the Bernoulli type branching particle schemes introduced by Crisan and Lyons in a series of papers [9, 10, 11] for filtering non linear diffusions.

- The path-space models introduced in the previously referenced work [45] also allow to describe in a very natural and simple way the complete line of ancestors of each particle and the genealogical tree of genetic populations. More precisely we show that the corresponding historical process is defined in term of path-valued genetic particles of the same type as before. This study also allows to transfer without further work the asymptotic results presented in several articles to historical models and related path-valued branching schemes. The key idea is to consider Feynman-Kac formulae associated to a path-valued Markov process of the following form

$$X_n = (X'_0, \ldots, X'_n)$$

where X' denotes an auxiliary Markov process. The corresponding genetic approximating model now consists of path-particles $\xi_n^i = (\xi_n'^i, \ldots, \xi_n'^i)$. In the selection stage we choose randomly the paths $\widehat{\xi}_n^i = (\widehat{\xi}_{0,n}'^i, \ldots, \widehat{\xi}_{n,n}'^i)$ according to the discrete distribution

$$\sum_{i=1}^N \frac{g_n(\xi_n^i)}{\sum_{j=1}^N g_n(\xi_n^i)}\ \delta_{\xi_n^i} = \sum_{i=1}^N \frac{g_n(\xi_{0,n}'^i, \ldots, \xi_{n,n}'^i)}{\sum_{j=1}^N g_n(\xi_{0,n}'^i, \ldots, \xi_{n,n}'^i)}\ \delta_{(\xi_{0,n}'^i, \ldots, \xi_{n,n}'^i)}.$$

The mutation transition simply consists in extending each path with an elementary move according to the transitions of X'.

In filtering settings the main contribution and impact of this paper is to prove that the empirical measures of genealogical genetic trees converge as the size of the population tends to infinity to the conditional distribution of the whole signal path with respect to the corresponding line of observations. If we consider the traditional filtering problem associated to the signal/observation pair process (X', Y') then with several forms of estimates we proved that

$$\lim_{N\to\infty} \frac{1}{N} \sum_{i=1}^N \delta_{(\widehat{\xi}_{0,n}'^i, \ldots, \widehat{\xi}_{n,n}'^i)} = \mathrm{Loi}(X_0', \ldots, X_n' | Y_0', \ldots, Y_n').$$

- In a more recent study [35] we introduce a natural and common framework to study discrete time genetics and a class of mean field interacting particle systems arising in physics as the two-velocity McKean model of gaz proposed in [70]. More precisely we analyze non linear dynamical systems in distribution space of the following form

$$\eta_{n+1} = \eta_n \mathcal{K}_{n+1,\eta_n}$$

where $\mathcal{K}_{n+1,\eta_n}$ is a non homogeneous Markov transition which depends on the probability measure η_n. The corresponding particle approximating scheme is here defined as the Markov chain $\xi_n = (\xi^1, \ldots, \xi_n^N)$ on the product space E^N and evolving according to the transitions

$$\mathbb{P}(\xi_{n+1} \in dy | \xi_n = x) = \prod_{p=1}^N \ \mathcal{K}_{n+1, \frac{1}{N}\sum_{i=1}^N \delta_{x^i}}(x^p, dy^p).$$

where $dy = dy^1 \times \ldots \times dy^N$ stands for an infinitesimal neighborhood of the point $y = (y^1, \ldots, y^N)$ and $x = (x^1, \ldots, x^N)$. Under a Lipschitz-type regularity condition on $\mathcal{K}_{n,\eta}$ we estimate the $\mathbb{L}_p$-mean norms between the empirical path measure $\eta_{[0,n]}^N$ and the corresponding McKean measure

$$\eta_{[0,n]}(dx_0, \ldots, dx_n) = \eta_0(dx_0)\ \mathcal{K}_{1,\eta_0}(x_0, dx_1) \ldots \mathcal{K}_{n,\eta_{n-1}}(x_{n-1}, dx_n).$$

For bounded test functions $f : E^{n+1} \to \mathbb{R}$ we prove that

$$\mathbb{E}(|\eta_{[0,n]}^N(f) - \eta_{[0,n]}(f)|^p)^{\frac{1}{p}} \leq \frac{c_p(n)}{\sqrt{N}} \; \|f\|.$$

We propose uniform estimates with respect to the time parameter for particle density profiles. We also compare the covariance functions associated to the fluctuations of two type particle approximating models.

1.3 Continuous time filtering problems

The branching particle strategies recently introduced by Crisan and Lyons in a chain of papers [9, 10, 11] partly rely on the fact that the stochastic integrals arising in Girsanov exponentials and in the branching transitions are readily accessible without further work. We have tackle this numerical problem in three different ways.

1. In [12] we present a classical approach. In this paper we extend to continuous time filtering problems the discrete time particle approximating schemes presented in our earlier studies. The idea is to combine traditional time discretization strategies with discrete time genetic approximating models.
   This combination allows to transfer in a natural way discrete time analysis to continuous time models. It also leads to the first $\mathbb{L}_1$-mean error estimates for interacting particle approximating models of the Kushner-Stratonovitch stochastic partial differential equation. If we consider a traditional non linear filtering problem of diffusions

   $$dS_t = a(S_t)dt + b(S_t)dW_t \qquad dY_t = h(S_t)dt + dV_t \qquad t \in [0,1]$$

   then the flow of conditional distributions π_t of S_t with respect to the observations $(Y_s\,, s \leq t)$ satisfies, for regular test functions f in the domain of the signal generator L, the Kushner-Stratonovich equation

   $$d\pi_t(f) \;=\; \pi_t(L(f))\;dt \;+\; (\pi_t(h^\star f) - \pi_t(h^\star)\pi_t(f))\;(dY_t - \pi_t(h)\;dt).$$

   A careful analysis of the associated Kallianpur-Striebel formulae induces a way to discretize in time the above equation in terms of the semigroup of the signal. Pour each time step $\Delta t = 1/M$ the resulting system is defined on the mesh $t_n = n/M$ in terms of a measure valued process having the same form as the one we used in discrete time filtering

   $$\pi_{t_n}^{(M)} = \Phi_{t_n}^{(M)}(\Delta Y_{t_n}, \pi_{t_{n-1}}^{(M)}).$$

   The elementary transitions $\Phi_{t_n}^{(M)}(\Delta Y_{t_n}, .)$ are again defined as a two step transformation. One depends on the discrete approximating semigroup of the signal between the dates t_{n-1} and t_n. The second one is related to the likelihood function of the observed increment $\Delta Y_{t_n} = Y_{t_n} - Y_{t_{n-1}}$. In the previously defined situation it can be chosen of the form

$$g_{t_n}^{(M)}(\Delta Y_{t_n}, x) = \exp\left(-|h(x)|^2/(2M) + h^\star(x)\Delta Y_{t_n}\right)$$

The resulting particle approximating model of the Kushner-Stratonovitch equation is now the genetic particle approximating model associated to such discrete time models. If N denotes the number of particles then we show that the rate of convergence in $\mathbb{L}_1$-norm is of order $\mathrm{O}(\sqrt{M} + \sqrt{N/M})$.
The question of finding sharp orders of convergence and the corresponding study of fluctuations remain open.

2. In [33] we look at continuous time filtering problems from a different angle. More precisely, we examine $\mathbb{R}^p \times \mathbb{R}^q$-valued and continuous signal/observation processes (S_t, Y_t), $t \in \mathbb{R}_+$, but the observation and estimation times are discrete $n \in \mathbb{N}$. Here the objective is to estimate the conditional distributions of the state signal S_n given the discrete sequence of observations $(Y_k\,, k = 0, 1, \ldots, n)$.
It seems that this natural question has not been covered in filtering literature. Here again the knowledge of some continuous time semigroups makes necessary the use of discrete time approximating models (between the observations). In this work we examine several sensor models including the degenerate situation in which the observation noise is correlated to the signal.
In this last situation we introduce an novel approximating strategy based on pair signal/observation particles $\xi_n^i = (S_n^i, Y_n^i)$. Roughly speaking S_n^i represents a likely signal component and Y_n^i its sampled noisy observation.
During the mutation of the corresponding genetic algorithm each pair particle (S_n^i, Y_n^i) evolves according to the one step transition of the pair process (S, Y) (for diffusions we use an Euler-type scheme with time step Δt). When the current observation is received $Y_n = y_n$ we select N-pair particles with the discrete distribution

$$\sum_{i=1}^N \frac{g^{(\epsilon)}(Y_n^i - y_n)}{\sum_{j=1}^N g^{(\epsilon)}(Y_n^j - y_n)}\, \delta_{(S_n^i, Y_n^i)} \qquad \text{with} \qquad g^{(\epsilon)}(y) = g(y/\epsilon)/\epsilon^q,$$

where g denotes some regular and positive probability density and ϵ is a regularization parameter which tends to zero.
Yet we notice that the $\mathbb{L}_1$-mean error rate depends on the dimension q of the observation process. For time step $\Delta t = 1/(1 + [N^{1/(2+q)}])$ and regularization order $\epsilon = 1/N^{1/(2+q)}$ we obtain in [33] a convergence order $\mathrm{O}(N^{1/(2+q)})$.

In the paper [33] we also examine the situation in which the discrete observations depend on the whole path of the signal on each time interval. We propose several $\mathbb{L}_1$-estimates as well as exponential rates in each situation.

The regularity and robustness of such particle numerical schemes is partly described in [30] and the fluctuations are studied in [32].
The paper [30] also examine the situation in which the initial distribution of the signal, its probability transitions and the statistics of the observation noise are unknown.
In these studies we also observe that the use of discrete time approximating schemes induces a bias in the fluctuations of the corresponding particle models. For the particle model with pair signal/observation components we use the above result to improve previous estimates. For finest Euler schemes with time step $\Delta t = 1/(1+[N^{1/(2+q/2)}])$ and coarsest regularizations with rate $\epsilon = 1/N^{1/(4+q)}$ we obtain an $\mathbb{L}_1$-mean estimate of order $\mathrm{O}(N^{1/(2+q/2)})$.

3. The paper [43] is context free but it can be used to study robust filtering equations. In contrast to Kallianpur-Striebel or Kushner-Stratonovitch formulations the robust description of the filter does not involves any stochastic integrals with respect to the observation process. Thus, this formulation holds for any observation flow and not only on a set of full probability. These optimal robust filters can also be written in terms of non homogeneous Feynman-Kac formulae. For regular test functions f, their evolution in time is given by equations of the form

$$\frac{d}{dt}\eta_t(f) = \eta_t(L_t(f)) + \eta_t(fV_t) - \eta_t(V_t)\eta_t(f)$$

where L_t is a non homogeneous generator. In the filtering situation we also indicate that L_t now depends on the current observation data at time t.
In this paper we present an genetic type interacting particle system approach to solve numerically these robust non linear and quadratic equations. This particle approximating scheme is defined using the above evolution formula. As traditionally we notice that they can be rewritten in the form

$$\frac{d}{dt}\eta_t(f) = \eta_t(L_{t,\eta_t}(f))$$

with a generator depending on a mutation L_t and related to a jump selection generator

$$L_{t,\eta_t}(f)(x) = L_t(f)(x) + \int (f(y) - f(x))\ V_t(y)\ \eta_t(dy).$$

From this description we see that the approximating particles evolve according to L_t-motions. Then, at some exponential random time, they interact with the whole system by jumping to more likely sites according to the potential V_t. Besides the novelty of this particle model we propose a semigroup approach to analyze their asymptotic behavior.

The choice of the generator $L_{t,\eta}$ and the corresponding particle approximating model is not unique. In [44] we examine this non uniqueness of the corresponding non linear martingale problems and associated McKean measures. We compare the fluctuations of the above particle model with the one associated to the other choice of generator

$$L_{t,\eta_t}(f)(x) = L_t(f)(x) + V_t'(x) \int (f(y) - f(x))\ \eta_t(dy)$$

Here the function V_t' denotes a non negative potential such that $V_t' + V_t = \mathrm{Cte}$ (we can chose for instance $V_t' = 2\|V_t\| - V_t$).

This choice induces a novel particle interpretation of Feynman-Kac formulae: In the latter the particles evolve according to L_t-motions in an environment with "soft" obstacles and absorbtion rates related to the potential V_t'. When a particle is killed then at the same time another particle is splitted into two individuals.

Interestingly, the rate of convergence of such continuous particle schemes improve the ones discussed in [12] combining the continuous time Kushner-Stratonovitch formulation with an auxiliary time discretization scheme. If N is the number of particles we obtain in [43] $\mathbb{L}_2$-mean errors of order $\mathrm{O}(\sqrt{N})$.

1.4 Continuous time particle systems

The asymptotic analysis of the continuous time models introduced in the previously referenced paper [43] is discussed in several studies. Before to come into more details it is useful to connect these particle models with two related subjects:

1. As we did originally, these continuous time particle models can be regarded as a re-normalization of a Moran-type genetic algorithm. This interpretation is clear since these two evolutionary algorithms are based on similar mutation and jump selection mechanisms. The Moran particle scheme contains much more frequent interacting jumps. As a result the limiting system now contains genetic noise and it converges to the stochastic Fleming-Viot superprocess.

2. We can also relate the quadratic nature of the "gain" term in the robust filtering equation and the Feynman-Kac flow η_t with the one of the collision term in the generalized Boltzmann equation introduced by Méléard in [67] and by Graham-Méléard in [59]. In this connection the resulting particle scheme can also be regarded as a the corresponding Nanbu type particle approximating model.

The asymptotic behavior as the size of the systems or the time horizon increases is discussed in the following studies:

- In [42] we study central limit theorems and exponential estimates. Our analysis is based on judicious errors decompositions and on the identification of some key martingales.
The uniform estimates with respect to the time parameter we obtained in this paper are partly based on the stability properties of the non linear Feynman-Kac semigroups presented in [38].

- In a more recent paper [40] we propose an abstract an general framework which allows to treat path-valued processes. This also provides a natural way to analyze the asymptotic behavior of historical processes and the convergence of genealogical trees as we did in the discrete time case [45].
We also analyze in this work strong propagations of chaos. We obtain an estimate in total variation norm of the same order as the one obtained by Méléard [67] and Graham-Méléard [59] for generalized and spatially homogeneous Boltzmann equations.
Our strategy differs from the interacting Boltzmann trees couplings used by Graham and Méléard. We propose in [40] a more flexible strategy based on an extension of the semigroup and martingale technics presented in [42]. For any time subdivision $0 \leq t_1 < \ldots < t_n \leq t$ we obtain the next key estimates
$$\|\mathbb{E}(\eta^N_{t_1} \otimes \ldots \otimes \eta^N_{t_n}) - \eta_{t_1} \otimes \ldots \otimes \eta_{t_n}\|_{tv} \leq \frac{n^2}{N}\; c(t).$$

- The numerical performance of these particle schemes has led us to extend in [41] these models to more general interacting branching rules. In this situation the size of the population is no more fixed but it forms an integer valued martingale. Thus, these branching schemes now take values in the state space $\mathcal{E} = \cup_{p\geq 0} E^p$, with a usual cemetery point $E^0 = \{\Delta\}$. The corresponding generator is defined, with some usual conventions, for any $x \in E^p$ and $p \geq 1$ as follows
$$\mathcal{G}_t(f)(x) = \sum_{i=1}^{p} L_t^{(i)}(f)(x) + \lambda_t(x) \int_{\mathcal{E}} (f(y) - f(x))\; \mathcal{Q}_t(x, dy).$$
The jump intensity is defined by $\lambda_t(x) = \sum_{i=1}^p V_t(x_i)$ and the transition kernel
$$\begin{aligned} \mathcal{Q}_t(f)(x) &= \frac{1}{p}\sum_{i=1}^{p} \int_{E\times\mathbb{N}} f(x(i,q,u))\; Q_t(x, x_i, d(u,q)) \\ x(i,q,u) &= (x_1, \ldots, x_{i-1}, \underbrace{u, \ldots, u}_{\text{q-times}}, x_{i+1}, \ldots, x_p) \end{aligned}$$
satisfies the compatibility condition
$$\frac{1}{p}\sum_{i=1}^{p} \int_{E\times\mathbb{N}} (\varphi(u)\; q)\; Q_t(x, x_i, d(u,q)) = \sum_{i=1}^{p} \frac{V_t(x_i)}{\sum_{j=1}^p V_t(x_j)}\; \varphi(x_i).$$

- In an adjacent study [54] we propose large deviation principles for continuous time interacting systems. Although the approach applies to a class of diffusion or jump type models it does not include Feynman-Kac particle approximations. This article presents a novel and relatively simple method based on an original extension/complement of the integral lemma of Laplace-Varadhan. To describe this new way to transfer large deviations principles suppose $\mathbb{P}^N_{m,\alpha}$ and $\mathbb{Q}^N_m$ are collections of probability measures on some measurable space Ω^N (which may depend on N) and indexed respectively by a real number $\alpha \in \mathbb{R}$ and a parameter m in some metric space (M, d). We also suppose that $\mathbb{P}^N_{m,1}$ does not depend on the parameter m and we write $\mathbb{P}^N_{m,1} = \mathbb{P}^N_1$. We also suppose that for any index pair $(\alpha, m) \in (\mathbb{R} \times M)$ we have $\mathbb{P}^N_{m,\alpha} \sim \mathbb{Q}^N_m$ and for $\mathbb{Q}^N_m$-almost every $x \in \Omega^N$
$$\frac{d\mathbb{P}^N_{m,\alpha}}{d\mathbb{Q}^N_m}(x) = \exp\left(N[\alpha S_N(x,m) + C_\alpha(\pi_N(x), m)]\right) \tag{1}$$
for some measurable functions $S_N : \Omega^N \times M \to \mathbb{R}$ and $C_\alpha : M \times M \to \mathbb{R}$.

  For some measurable mapping $\pi_N : \Omega^N \to M$ and for any $(\alpha, m) \in (\mathbb{R} \times M)$ we define the image measures
$$\overline{\mathbb{P}}^N_{m,\alpha} = \mathbb{P}^N_\alpha \circ \pi_N^{-1} \quad \text{and} \quad \overline{\mathbb{Q}}^N_m = \mathbb{Q}^N_m \circ \pi_N^{-1}$$
  Suppose the sequence of probability measures $\overline{\mathbb{Q}}^N_m$, $N \geq 1$ satisfies an LDP with good rate function $H_m : M \to [0, \infty]$, for each $m \in M$. Also assume that the mappings $\{C_\alpha(., m)\ ;\ \alpha \in \mathbb{R}\}$, are continuous at each m, $C_\alpha(m, m) = 0$ and the following exponential moment condition holds
$$\limsup_{N\to\infty} \frac{1}{N} \log \int_{\Omega^N} \exp \gamma N[S_N(x,m) + C_1(\pi_N(x), m)] \ \ d\mathbb{Q}^N_m(x) < \infty \tag{2}$$
  for some $(m, \gamma) \in M \times (1, \infty)$. Then $\overline{\mathbb{P}}^N_1 = \mathbb{P}^N_1 \circ \pi_N^{-1}$ satisfies an LDP with good rate function
$$I : m \in M \to I(m) = H_m(m) \in [0, \infty]$$
  In [54] we apply this transfer lemma to derive large deviations principles in path space for a class of regular interacting jump or McKean diffusions processes. We prove that the law of the empirical measures of the path particles satisfy large deviations principles with a good rate function I given for any distribution μ on càdlàg trajectories by
$$I(\mu) = \mathrm{Ent}(\mu \mid \mathbb{Q}_{\eta_0, \phi(\mu)})$$
  where $\phi(\mu) = (\mu_t)_{t\geq 0}$ denotes the distribution flow of the time marginals of μ and $\mathbb{Q}_{\eta_0,\phi(\mu)}$ the solution of the non homogenous and non linear martingale problem associated to the corresponding sequence of generators L_{μ_t}.

2 Genetic algorithms and global optimization

Although genetic algorithms arise naturally in filtering and in the analysis of Feynman-Kac formulae they have been introduced originally in 1975 by Holland in [60] to find the extrema of a numerical function. In his PhD dissertation [7] Cerf presents in 1994 the first rigorous and well founded convergence as the time parameter tends to infinity of such schemes on finite state space and discrete time by introducing an auxiliary cooling schedule. His main result can be stated as follows:
For sufficiently large systems with irreducible mutations the genetic populations with logarithmic temperature concentrate as time tends to infinity to the set of global extrema of the potential function.
His proof of this result is partly based on the work of Trouvé on the generalized simulated annealing [73] and partly on a precise analysis of virtual communication cost functions.

In [39] we simplify and extend this study and we analyze generalized simulated annealing models with time non homogeneous potentials. With the help of this extension we construct a class of stochastic algorithms which converges to the global extrema of integral potentials having the integral form

$$U(x) = \int \mathcal{U}(x,z)\ \mu(dz)$$

where μ is a given probability measure. The idea is to approximate U along the evolution of the algorithm with the flow of non homogeneous and stochastic potentials given by

$$U_t(x) = \frac{1}{t}\int_0^t \mathcal{U}(x,Z_s)\ ds$$

and associated to a Markov process Z with invariant measure μ. We also simplify the large deviation technics presented in Trouvé [73]. Here we present an alternative semigroup approach based on the relative entropy and log-Sobolev constants. We also propose a natural test set strategy to analyze the extrema of a virtual energy function.

In the second part of the article we apply the above analysis to study the convergence of genetic algorithms. We discuss two different types of Markov transitions

$$G_\beta^{(1)} = M_\beta S_\beta \qquad \text{and} \qquad G_\beta^{(2)} = \alpha M_\beta + (1-\alpha)\ S_\beta, \qquad \alpha \in (0,1),$$

with the mutation/selection kernels:

$$M_\beta(x,y) = \prod_{i=1}^N\ K_\beta(x^i,y^i)\,, \qquad S_\beta(x,y) = \prod_{i=1}^N\ \sum_{j=1}^N \frac{e^{-\beta U(x^j)}}{\sum_{k=1}^N e^{-\beta U(x^k)}}\ 1_{x^j}(y^i).$$

The mutations K_β are built in terms of an irreducible kernel $K_0(x,y)$ and an U-compatible equivalence relation $a : E \times E \to \mathbb{R}_+$ (that is $a(x,y) = 0 \Rightarrow U(x) = U(y)$). As soon as $a(x,y) > 0$ it takes the form

$$K_\beta(x,y) = K_0(x,y)\ e^{-\beta\, a(x,y)}.$$

For each $\alpha = 1,2$ we denote by $\xi_t^{(\alpha)}$ the continuous time genetic algorithm with infinitesimal generator $L_{\beta_t}^{(\alpha)} = G_{\beta_t}^{(\alpha)} - Id$ and inverse cooling schedule $\beta(t) = \frac{1}{T}\log(e+t)$, $\beta(0) = 1/T < \infty$. In [39] we prove that there exist an initial temperature T_0 (which can be explicitly described) and two critical population sizes $N^{(\alpha)}(a,U)$, $\alpha = 1,2$, such that

$$(\ N \geq N^{(\alpha)}(a,U) \quad \text{and} \quad T > T_0\ ) \implies \lim_{t\to\infty} \mathbb{P}(\xi_t^{(\alpha)} \in U_a^\star) = 1$$

where

$$U_a^\star = \{x \in\ E^N\ ;\ a(x^i,x^j) = 0 \text{ and } U(x^i) = \inf_E U\}.$$

In filtering settings the one by one mutation/selection order is dictated by the corresponding prediction/updating stages. In numerical function analysis there is no reasons to follow this order and it may even be useful to bypass this strict rule. In this sense the second genetic algorithm $\xi_t^{(2)}$ is more flexible since it permits random sequences of mutations and selections. Its numerical analysis at low temperature is even more clear and simple. Although the exact values of the critical population sizes $N^{(\alpha)}(a,U)$ are not known we proved that

$$N^{(1)}(a,U) = N^{(2)}(a,U) \times (1 + \delta(U)/\delta(a))\ >\ N^{(2)}(a,U)$$

with $\delta(U) = \sup\{|U(x) - U(y)|\ ;\ x,y \in E\}$ et $\delta(a) = \sup\{a(x,y)\ ;\ x,y \in E\}$.

These observations lead us to think that there exists situations with less individuals in which the second genetic algorithm converges to the desired extrema but the first one fails.

In the last few years genetic algorithm theory has been developed in two distinct directions:

1. The first one is related to numerical function analysis. It is essentially founded on the theory of stochastic algorithms with rare transitions.

2. The second one is a little more recent and it corresponds to applications to non linear filtering problems. This subject is founded on Monte-carlo methods and branching and interacting particle systems theory.

In a conference and invited survey paper [41] we explain in some detail these two approaches. We also shed some light on different continuous or discrete branching particle models arising in filtering which may be used in global optimization problems. The convergence analysis has been started in [5].

In a complementary paper [34] we compare various existing approaches proposed in genetic algorithm literature and numerical function analysis. We study the asymptotic behavior of a general class of discrete genetic models including cross-over mechanisms as the population size tends to infinity.

3 Contraction properties of Markov operators and non linear semigroups

This thematic again originates from non linear filtering and more precisely from the study of the optimal filter asymptotic stability. We have developed our research in this field in two main directions:

1. First we have extended this rather restrictive filtering framework to study of the stability properties of general class non linear Feynman-Kac-Schrödinger type semigroups. Then we have also proposed a original strategy partly based on Dobrushin ideas related to the stability of Markov chains.
   The discrete time situation has been first considered in the papers [28, 29, 42]. In [38] we have extended this approach to treat Feynman-Kac semigroups associated to a abstract multiplicative function for discrete or continuous time models.

2. In a more recent article [37] we examine the contraction properties of Dobrushin's coefficient and we analyze the regularity of a given Markov operator with respect to a class of relative entropy. We develop several strategies to estimate these contraction coefficients in terms of spectral gaps and classical or modified log-Sobolev constants.

3.1 Stability properties of non linear Feynman-Kac semigroups

To shed some light on our contributions in this field it is useful to connect our research with related literature on filtering. First we recall that the optimal filter is solution of a non linear measure valued equation. The initial condition of these evolution equations is the initial distribution of the signal process. In general this distribution is unknown and the optimal filter and any of its approximating model started as some erroneous initial point. To show that the optimal filter forgets and corrects this error is equivalent to study the asymptotic stability of the underlying non linear filtering equation in distribution space.

Originally the long time behavior of the non linear filtering equation has first been studied by Kunita [63, 64] in continuous time settings and then by Stettner [71, 72] for discrete time models. These two authors essentially show that if the Feller-semigroup of the signal has an invariant measure μ then the optimal filter is a measure valued Markov process with an unique invariant measure with barycenter μ. In [69] Ocone and Pardoux precise that the optimal filter will correct an erroneous initial data as soon as the (usually unknown) initial measure of the signal is absolutely continuous with respect to this wrong starting point. This asymptotic stability property gives a first partial answer to the filter initialization problem but it clearly relies on a condition which may be difficult to check in practice. Following these pioneering studies arise a chain of papers proving correction decays, in alphabetical order Atar-Zeitouni [1, 2], Budhiraja-Ocone [4], Delyon-Zeitouni [55], Da Prato-Furhman-Malliavin [14]. Without getting into the details of each work we only underline that all these studies without exceptions are based on projective properties, Hilbert metric and Birkhoff's ergodic coefficient.

At the same time and independently of the previously referenced articles we proposed in [28, 29] an alternative approach based on the use of the total variation distance and Dobrushin's contraction coefficient $\beta(K)$ of a Markov operator K acting on the set of probability measures. We recall that it can be alternatively defined by one of the formulae

$$\beta(K) = \sup_{\mu_1,\mu_2} \frac{\|\mu_1 K - \mu_2 K\|_{tv}}{\|\mu_1 - \mu_2\|_{tv}} = \sup_{x_1,x_2,A} |K(x_1, A) - K(x_2, A)|.$$

Thus, under some rather strong mixing conditions it is possible to estimate the Lipschitz-contraction. For instance we have that

$$(\mathcal{K})_\epsilon : \qquad \frac{dK(x,.)}{dK(x',.)}(y) \geq \epsilon > 0 \qquad \Longrightarrow \qquad \alpha(K) \geq \epsilon.$$

In [28, 29] we discuss the asymptotic stability of discrete time models. We develop a semigroup approach which does not need to interpret the filtering equation in projective spaces. In a complementary article [38] we extend this study to general discrete or continuous models of the form

$$\eta_t(f) = \gamma_t(f)/\gamma_t(1) \qquad \text{with} \quad \gamma_t(f) = \mathbb{E}(f(X_t)\ Z_{0,t})$$

and related to an abstract multiplicative function $Z_{r,t} = Z_{r,s}\ Z_{s,t}$, $r \leq s \leq t$, which is predictable with respect to the natural filtration of X. The main idea is to study for each finite time horizon t the stability properties of the process X under the new probability measure $\mathbb{Q}_t$ defined by the Radon-Nykodim derivative

$$\frac{d\mathbb{Q}_t}{d\mathbb{P}_t} = \frac{Z_{0,t}}{\mathbb{E}_{\mathbb{P}}(Z_{0,t})}.$$

where $\mathbb{P}_t$ stands for the distribution of X on this time horizon. Let $K^{(t)}_{s,r}(x, dy) = \mathbb{Q}_t(X_r \in dy | X_s = x)$ be the transition semigroup of X under $\mathbb{Q}_t$. The stability properties of the non linear semigroup $\Phi_{s,t}(\eta_s) = \eta_t$ and the Dobrushin contraction coefficient $\beta(K^{(t)}_{s,r})$ are connected by the formula

$$\beta(K^{(t)}_{s,t}) = \sup_{\mu_1,\mu_2} \|\Phi_{s,t}(\mu_1) - \Phi_{s,t}(\mu_2)\|_{tv}.$$

The regularity properties of the semigroup $P_{s,r}$ of X under $\mathbb{P}$ are, in a second stage, transfered to the non homogeneous semigroup $K^{(t)}_{s,r}$ of X under $\mathbb{Q}_t$ according to the representation formula

$$K^{(t)}_{s,r}(f)(x) = \mathbb{E}_{\mathbb{P}}(f(X_r)\ Z_{s,t}|X_s = x)/\mathbb{E}_{\mathbb{P}}(Z_{s,t}|X_s = x).$$

For homogeneous potentials and semigroups $P_{s,t} = P_{t-s}$ we prove for instance that:

$$\frac{dP_1(x,.)}{dP_1(x',.)}(y) \geq \epsilon > 0 \Longrightarrow \frac{1}{t}\log\beta(K^{(t)}_{0,t}) \leq -c\ \epsilon^2.$$

For discrete time models we can choose $c = 1$. In the continuous time situation the above estimate is satisfied for any $t \geq 2$ with a constant $-\log(2c) = \mathrm{osc}(V)$ which depends on the potential function V.

In filtering settings, this approach with abstract multiplicative functions is not restricted to discrete or continuous Kallianpur-Striebel representation but it also allows to treat robust equations.

We have also used this general formulation to study the asymptotic stability of a Markov particle evolving in an environment with obstacles and absorption rates related to a potential function.

3.2 Contraction properties of Markov operators

In this other research direction we have extended the study of the contraction properties of Markov transitions for the total variation distance to other kind of criteria. The corresponding article presenting these extensions is [37].
Our strategy differs from traditional approaches on the stability of Markov chaîns by the notable fact that it doesn't use (whenever it exists) the invariant measure of the underlying transitions. The proximity criteria we have analyzed are defined in terms of an abstract ϕ-relative entropy. This class of entropy includes for instance the L_p-norms, the total variation and Kakutani-Hellinger distance as well as the Boltzmann-Shannon-Kullback or Havrda-Charvat entropy.

These ϕ-relative entropies are defined in terms of an homogeneous convex function $\phi : \mathbb{R}^2_+ \to \mathbb{R} \cup \{\infty\}$ with $\phi(ax, ay) = a\phi(x,y)$ and $\phi(1,1) = 0$. With some usual conventions they are given by the formula

$$H_\phi(\mu,\nu) = \int\ \phi(d\mu, d\nu).$$

Let $\beta_\phi(K,\nu) = 1 - \alpha_\phi(K,\nu)$ be the contraction coefficient of a Markov kernel K around the measure ν and with respect to this criterium H_ϕ. In other words this coefficient is the best possible constant such that for any probability measure μ the following inequality holds

$$H_\phi(\mu K, \nu K) \leq \ (1 - \alpha_\phi(K,\nu)) \ H_\phi(\mu,\nu).$$

If we choose $\phi_0(x,1) = |x-1|$ we obtain the traditional total variation distance and in this situation the Dobrushin ergodic coefficient $\alpha(K)$ is the global contraction coefficient $\alpha(K) = \inf_\nu \alpha_{\phi_0}(K,\nu)$.

Our study of the contraction regularity of Markov operators [37] has been developped in two directions:

1. Our first result is an extension of a collective paper of Cohen, Iwasa, Rautu, Ruskai, Seneta, Zbaganu [61] on finite state spaces to general and abstract measurable spaces. We prove that the contraction coefficient of a given Markov transition K on any measurable space with respect to any ϕ-relative entropy can always be estimated with Dobrushin coefficient and
$$\alpha(K) \leq \alpha_\phi(K,\nu).$$
Therefore if the transition K satisfies the above mixing condition $(\mathcal{K})_\epsilon$ then we obtain easily the lower estimate $\alpha_\phi(K,\nu) \geq \epsilon$. Notice also that $\mathbb{L}_p$-spaces correspond to the choices $\phi(x,1) = |x-1|^p$. Under this mixing condition we also discuss contraction properties on more general Orlicz spaces $(\mathbb{L}^\psi(\nu), \|.\|_{\psi,\nu})$ associated to a given Young function ψ. In this situation we prove that for any $\mu << \nu$, $d\mu/d\nu \in \mathbb{L}^\psi(\nu)$ we still have the inequality
$$(\mathcal{K})_\epsilon \quad \Longrightarrow \quad \|1 - d\mu K/d\nu K\|_{\psi,\nu K} \leq (1-\epsilon) \ \|1 - d\mu/d\nu\|_{\psi,\nu}.$$

2. In the second part of this paper we analyze in more details the contraction properties with respect to the $\mathbb{L}_2$-norm and the traditional relative entropy. In the above framework the above two criteria correspond to the ϕ-relative entropies associated to the functions $\phi_1(x,1) = |x-1|^2$ and $\phi_2(x,1) = x\log x$. To present a part of our result it is convenient to introduce some additional notations. For each K-invariant probability measure ν we denote respectively by $\lambda(\nu,K)$, $ln(\nu,K)$ and $lm(\nu,K)$ the spectral gap, the classical and the modified log-Sobolev constant defined by
$$\begin{aligned}
\lambda(\nu,K) &= \inf\{\mathcal{E}_{\nu,K}(f,f)/\mathrm{var}_\nu(f) \ ; \ f \in \mathbb{L}^2(\nu) - \mathrm{Vect}(1)\} \\
ln(\nu,K) &= \inf\{ \ \ \mathcal{E}_{\nu,K}(f,f)/\mathrm{ent}_\nu(f^2) \ ; \ f \in \mathbb{L}^2(\nu) - \mathrm{Vect}(1)\} \\
lm(\nu,K) &= \inf\{\mathcal{E}_{\nu,K}(f^2,\log f^2)/\mathrm{ent}_\nu(f^2) \ ; \ f \in \mathbb{L}^{\psi_m}(\nu) - \mathrm{Vect}(1)\}
\end{aligned}$$
with $\psi_m(x) = x^2\log(x^2)$ and $\mathrm{ent}_\nu(f) = \int f\log(f/\nu(f)) \ d\nu$, for any $\sqrt{f} \in \mathbb{L}^\psi(\nu)$. We prove that the contraction coefficient for the $\mathbb{L}_2$-norm and

the traditional relative entropy can be estimated according to the following inequalities

$$\begin{array}{rcl}\alpha_{\phi_1}(K,\nu) & = & \lambda(\nu, KK^\star_\nu) \ \ (\geq lm(\nu,K)/2 \geq ln(\nu,K)) \\ \alpha_{\phi_2}(K,\nu) & \geq & ln(\nu, KK^\star_\nu)\end{array}$$

Using modified log-Sobolev constants we also obtain the existence of some number $\rho \in (0,1)$ such that

$$\rho\ lm(\nu, KK^\star_\nu) \leq \alpha_{\phi_2}(K,\nu) \leq lm(\nu, KK^\star_\nu).$$

These last estimates are often more useful than the first ones. For instance Miclo proved in [68] that $ln(\nu, KK^\star_\nu) = 0$ whenever ν is not a finite linear combination of Dirac masses. On the other hand from the definition of the spectral gap and the modified log-Sobolev constant we readily obtain that

$$(\mathcal{K})_\epsilon \quad \Longrightarrow \quad \lambda(\nu, KK^\star_\nu) \geq \epsilon \quad \text{and} \quad lm(\nu, KK^\star_\nu) \geq \epsilon.$$

4 Idempotent probability calculus

This last research theme is directly connected to my phD dissertation. It essentially concerns the development of an idempotent probability calculus which allows to extend probability intuition, ideas and technics to deterministic optimal control.
The key idea is to replace in the structural axioms of probability theory the semiring of real numbers with traditional operations $(+, \times)$ by the semiring of extended real numbers with the two operations. $(\oplus = max, \otimes = +)$. This approach is partly described in [47, 20] and in the appendix I have written for the book of Kolokoltsov and Maslov [19]. In [53] we use this framework to discuss the connections between optimal control regulation with filtering. Using this approach we have adapted in a natural way particle filtering technics to optimal regulation control around a reference path.

In [23] we complete these studies. We present a rigorous construction of idempotent probability calculus based on the theory of idempotent measures of Maslov introduced in [66] and on large deviation calculus.
In this framework Bellman evolutions correspond to Chapman-Kolmogorov transport equations and Hamiltonian operators can be regarded as infinitesimal generators of decision processes. These two probability interpretations allow to define in a natural way weak solutions of Hamilton-Jacobi equations. These solutions are the analog of the weak solution of the evolution equation of the probability distributions of a given Markov process.

To throw some light on this approach we have chosen to present a model which allows to analyze these equations. To this end, we denote by $\Omega = \mathbb{L}^\infty((0,1), \mathcal{P}(U))$

the space of relaxed controls taking values in some compact set U with the weak-$\star$ topology. For each $x \in \mathbb{R}^d$ we denote by X^x the decision process taking values in $\mathcal{C} = \mathcal{C}((0,1), \mathbb{R}^d)$ and defined for any $t \in (0,1)$ and any admissible control $\omega \in \Omega$ by the formula

$$X_t^x(\omega) = x + \int_0^t \int_U f(X_s^x(\omega), u)\ \omega_s(du)\ ds$$

where f is some regular drift. On this set of controls Ω we defined an idempotent performance measure $\mathbf{P}_x$ by setting

$$\mathbf{P}_x(A) = \sup_{\omega \in A} \mathbf{p}_x(\omega) \quad \text{with} \quad \mathbf{p}_x(\omega) = \int_0^1 \int_U l(X_t^x(\omega), u)\ \omega_t(du)\ dt$$

where l denotes a regular Lagrangian. The idempotent distribution of the path X^x on $\mathcal{C}((0,1), \mathbb{R}^d)$ and the conditional distribution of X_t^x given $X_s = y$ (with $s \leq t$) are respectively defined in this framework as follows:

$$\begin{aligned} \mathbf{p}_x^X(z) &= \mathbf{P}_x((X^x)^{-1}(\{z\})) = \sup\{\mathbf{p}_x(\omega)\ ,\ X^x(\omega) = z\} \\ \mathbf{p}_{s,t}^X(x,y) &= \sup\{\int_s^t \int_U l(X_r^{s,x}(\omega), u)\ \omega_r(du)\ dr\ ;\ \omega \in \Omega\ ,\ X_t^{s,x}(\omega) = y\}. \end{aligned}$$

In the same way the mathematical expectation with respect to these idempotent measures are defined according the idempotent integral formula:

$$\mathbf{E}_x(f(X)) = \int^{\oplus} f(z) \odot \mathbf{P}_x^X(dz) = \sup_{z \in \mathcal{C}} (f(z) + \mathbf{p}_x^X(z)).$$

With these notations, the Hamiltonian H associated to the previously defined decision process can be regarded as the local mean drift of the process and

$$\lim_{\epsilon \to 0} \epsilon^{-1}\ \mathbf{E}_x(\langle p, X_\epsilon^x - x \rangle) = H(x,p) = \sup_{u \in U} (\langle p, f(x,u) \rangle + l(x,u)).$$

For any bounded Lipschitz test function $f \in \mathrm{Lip}(\mathbb{R}^d)$ we can prove that the mapping ϕ defined by

$$\phi(t,x) = \mathbf{P}_t(f)(x) = \mathbf{E}_x(f(X_t^x))$$

is again Lipschitz and by homogeneity we conclude that

$$\phi(t,x) = \mathbf{E}_x(\mathbf{E}_x(f(X_t^x)|X_\epsilon^x)) = \mathbf{E}_x(\phi(t-\epsilon, X_\epsilon^x)).$$

Rademacher's theorem now indicates that ϕ is differentiable for almost every (t,x) and from the above estimates we have that

$$\epsilon^{-1}\ \mathbf{E}_x(\langle D_x\phi(t,x), X_\epsilon^x - x \rangle) = \frac{\partial \phi}{\partial t}(t,x) + \mathrm{O}(\epsilon) \to H(x, D\phi(t,x)) = \frac{\partial \phi}{\partial t}(t,x).$$

More generally ϕ is the unique solution of the Hamilton-Jacobi equation (in the sense of super and sub-gradients) and $\{\mathbf{P}_t\ ;\ t \in (0,1)\}$ is a monotone semigroup of

$(\oplus, \odot)$-linear operators.
This semigroup also acts in a dual way on upper semi continuous densities (u.s.c.) μ according to the integral formula

$$\mu \mathbf{P}_t(y) = \int^{\oplus} \mu(dx) \odot \mathbf{p}_{0,t}(x,y) = \sup_{x \in \mathbb{R}^d} (\mu(x) + \mathbf{p}_{0,t}(x,y)).$$

In this way we construct the weak solution $\psi(t,y) = \mu \mathbf{P}_t(y)$ of the Hamilton-Jacobi equation corresponding to the dual operator $H^\star(x,p) = H(x,-p)$ starting at $\psi(0,.) = \mu$ and defined by

$$\frac{\partial \psi}{\partial t}(t,x) = H^\star(x, D\psi(t,x)).$$

In [24] we present some situations in which these weak solutions coincide with the viscosity solutions introduced by Crandall and Lions in [8].

We also apply this framework in image processing and mathematical morphology. This gives an alternative approach to the use of morphological Beuchler gradients or the super and sub-derivations proposed by Brockett and Maragos in [3] and by Van der Boomgaard in [74].

In [25] we present the version of Dynkin's formula for Markov processes. For a similar decision process and for a regular function $f \in \mathcal{C}_b^1(\mathbb{R}^d)$ we have that

$$\mathbf{E}_x(f(X_t^x) - \int_0^t H(X_s^x, Df(X_s^x))\, ds\ ) = f(x).$$

As in the construction of Lyapunov stochastic functions we use in [15] and [25] this formula to estimate the performance of some classes of admissible controls. For instance, suppose we can find some regular mapping V such that

$$H(x, DV(x)) \leq -a$$

on some Borel subset A and for some positive number $a > 0$. In this situation it is possible to calibrate the performance of the subset of admissible controls for which the decision process X^x, starting at $x \in A$, remains in the same set. More precisely Dynkin's formula leads to the formula

$$\mathbf{P}_x(\forall 0 \leq s \leq t\ ;\ X_s^x \in A) \leq \varphi(x) - \inf_A \varphi - a\ t.$$

In [25] we also discuss the stability of decision processes. We propose an explicit description of the fixed point of a class of non expansive mappings. By analogy with the classical Markov chaîn theory, the idea is to interpret these fixed points as the invariant measures of a markovian decision process. To present this result we proceed as in probability and we associate to a canonical process X with initial distribution μ and elementary transitions $\mathbf{p}$ an idempotent measure on path space by setting

$$\mathbf{P}_\mu((X_0, \ldots, X_n) \in A) = \sup\{\mu(x_0) + \sum_{p=1}^{n} \mathbf{p}(x_{n-1}, x_n) \; ; \; (x_0, \ldots, x_n) \in A\}$$

When the state space is compact and for Lipschitz communication costs functions $\mathbf{p}$ we define two integral operations. One is acting on continuous functions f, the other one on u.s.c. densities μ and they are defined by

$$\mathbf{P}(f)(x) = \sup_{y \in E}(\mathbf{p}(x, y) + f(y)) \quad \text{and} \quad (\mu\mathbf{P})(y) = \sup_{x \in E}(\mu(x) + \mathbf{p}(x, y)).$$

To describe in some details the corresponding invariant measures we introduce by analogy with Markov chaîns the taboo communication measures

$$x\mathbf{P}^n(x, y) = \mathbf{P}_x(X_n = y \, , \tau_x \geq n) \quad \text{with} \quad \tau_x = \inf\{n \geq 1 \; ; \; X_n = x\}$$

with

$$\pi_x(y) = \oplus_{n \geq 1} \; x\mathbf{P}^n(x, y) = \mathbf{E}_x(\oplus_{n=1}^{\tau_x} 1_{X_n = x})$$

where 1_A stands for the $(\max, +)$-indicatrix $1_A(x) = 0$ if $x \in A$ and $1_A(x) = -\infty$ otherwise. In this idempotent probability calculus we prove that for ant $x \in E$

$$\pi_x \quad \mathbf{P} - \text{invariant} \iff \mathbf{P}_x(\tau_x < \infty) = 0 \iff \oplus_{n \geq 1}\mathbf{P}^n(x, x) = 0.$$

References

[1] R. Atar and O. Zeitouni. Exponential stability for nonlinear filtering. *Annales de l'Institut Henri Poincaré*, 33(6):697–725, 1997.

[2] R. Atar and O. Zeitouni. Lyapunov exponents for finite state space nonlinear filtering. *Society for Industrial and Applied Mathematics. Journal on Control and Optimization*, 35(1):36–55, January 1997.

[3] R.W. Brockett and P. Maragos. Evolution equations for continuous scale morphological filtering. *IEEE Trans. on signal Processing*, 42(12), 1994.

[4] A. Budhiraja and D. Ocone. Exponential stability of discrete time filters for bounded observation noise. *Systems and Control Letters*, 30:185–193, 1997.

[5] R. Carmona and P. Del Moral. Interacting particle methods and non linear estimation. Preprint, 2002.

[6] H. Carvalho, P. Del Moral, A. Monin, and G. Salut. Optimal nonlinear filtering in gps/ins integration. *Institute of Electrical and Electronics Engineers. Transactions on Aerospace and Electronic Systems*, 33(3):835–850, July 1997.

[7] R. Cerf. *Une théorie asymptotique des algorithmes génétiques.* Thèse de doctorat, Université Montpellier II, March 1994.

[8] M.G. Crandall and P.L. Lions. Viscosity solutions of hamilton-jacobi equations. *Trans. Amer. Math. Soc.*, 277:1–42, 1983.

[9] D. Crisan, J. Gaines, and T.J. Lyons. A particle approximation of the solution of the kushner-stratonovitch equation. *Society for Industrial and Applied Mathematics. Journal on Applied Mathematics*, 58(5):1568–1590, 1998.

[10] D. Crisan and T.J. Lyons. Nonlinear filtering and measure valued processes. *Probability Theory and Related Fields*, 109:217–244, 1997.

[11] D. Crisan and T.J. Lyons. A particle approximation of the solution of the Kushner-Stratonovitch equation. *Probability Theory and Related Fields*, 115(4):549–578., 1999.

[12] D. Crisan, P. Del Moral, and T.J. Lyons. Interacting particle systems approximations of the Kushner-Stratonovitch equation. *Advances in Applied Probability*, 31(3):819–838, September 1999.

[13] D. Crisan, P. Del Moral, and T.J. Lyons. Non linear filtering using branching and interacting particle systems. *Markov Processes and Related Fields*, 5(3):293–319, 1999.

[14] G. Da Prato, M. Furhman, and P. Malliavin. Asymptotic ergodicity for the Zakai filtering equation. *Comptes Rendus de l'Académie des Sciences de Paris. Série I. Mathématique*, 321(5):613–616, 1995.

[15] P. Del Moral. On the stability of Maslov optimization processes. In *Proceedings XXXIV IEEE Conference on Decision and Control*, New Orleans, CDC, December 1995.

[16] P. Del Moral. Non-linear filtering: interacting particle resolution. *Markov Processes and Related Fields*, 2(4):555–581, 1996.

[17] P. Del Moral. Nonlinear filtering using random particles. *Theory of Probability and its applications*, 40(4):690–701, 1996.

[18] P. Del Moral. Filtrage non linéaire par systèmes de particules en interaction. *Comptes Rendus de l'Académie des Sciences de Paris. Série I. Mathématique*, 325:653–658, 1997.

[19] P. Del Moral. Maslov optimization theory: optimality versus randomness. In V.N. Kolokoltsov and V.P. Maslov, editors, *Idempotency Analysis and its Applications*, Mathematics and its Applications 401, pages 243–302. Kluwer Academic Publishers, Dordrecht/Boston/London, 1997.

[20] P. Del Moral. Maslov optimization theory: topological aspects. In J. Gunawardena, editor, *Idempotency*, Publications of the Newton Institute, pages 354–383. Cambridge University Press, 1998.

[21] P. Del Moral. Measure valued processes and interacting particle systems. Application to non linear filtering problems. *The Annals of Applied Probability*, 8(2):438–495, 1998.

[22] P. Del Moral. A uniform convergence theorem for the numerical solving of non linear filtering problems. *Journal of Applied Probability*, 35:873–884, 1998.

[23] P. Del Moral and M. Doisy. Maslov idempotent probability calculus. part I. *Theory of Probability and Its Applications*, 43(4):735–751, 1998.

[24] P. Del Moral and M. Doisy. Maslov idempotent probability calculus. part II. *Theory of Probability and Its Applications*, 44(2):384–400, 1999.

[25] P. Del Moral and M. Doisy. On the applications of Maslov optimization theory. *Mathematical Notes*, 69(2):232–244, February 2001.

[26] P. Del Moral and A. Guionnet. Large deviations for interacting particle systems. Applications to non linear filtering problems. *Stochastic Processes and their Applications*, 78:69–95, 1998.

[27] P. Del Moral and A. Guionnet. A central limit theorem for non linear filtering using interacting particle systems. *The Annals of Applied Probability*, 9(2):275–297, 1999.

[28] P. Del Moral and A. Guionnet. On the stability of measure valued processes with applications to filtering. *Comptes Rendus de l'Académie des Sciences de Paris. Série I. Mathématique*, 329:429–434, 1999.

[29] P. Del Moral and A. Guionnet. On the stability of interacting processes with applications to filtering and genetic algorithms. *Annales de l'Institut Henri Poincaré*, 37(2):155–194, 2001.

[30] P. Del Moral and J. Jacod. Interacting particle filtering with discrete observations. In N. J. Gordon A. Doucet, J. F. G. de Freitas, editor, *Sequential Monte Carlo Methods in Practice*. Springer-Verlag, 2001.

[31] P. Del Moral and J. Jacod. Interacting particle filtering with discrete observations: asymptotic behaviour in the gaussian case. In H. Kunita B.S. Rajput S. Watanabe J. Xiong T. Hida, R. Karandikar, editor, *Stochastics in finite and infinite dimensions. in the honour of G. Kallianpur*, pages 101–122. Birkhaüser, Trends in Mathematics, 2001.

[32] P. Del Moral and J. Jacod. The monte-carlo method for filtering with discrete time observations. central limit theorems. In *Proceedings of the Workshop on Numerical Methods and Stochastics*. Fields Institute, Toronto, 2001 (to appear).

[33] P. Del Moral, J. Jacod, and P. Protter. The Monte Carlo method for filtering with discrete time observations. *Probability Theory and Related Fields*, 120:346–368, 2001.

[34] P. Del Moral, L. Kallel, and J. Rowe. Modeling genetic algorithms with interacting particle systems. *Revista de mathemática: Teoria y Aplicationes*, 8(2), July 2001.

[35] P. Del Moral, M.A. Kouritzin, and L. Miclo. On a class of discrete generation interacting particle systems. *Electronic Journal of Probability*, 6(16):1–26, 2001.

[36] P. Del Moral and M. Ledoux. Convergence of empirical processes for interacting particle systems with applications to nonlinear filtering. *Journal of Theoretical Probability*, 13(1):225–257, 2000.

[37] P. Del Moral, M. Ledoux, and L. Miclo. About super-contraction properties of markov kernels. Publications du Laboratoire de Statistique et Probabilités, no 01-01, Université Paul Sabatier, 2001.

[38] P. Del Moral and L. Miclo. Asymptotic stability of non linear semigroup of Feynman-Kac type. Préprint, publications du Laboratoire de Statistique et Probabilités, no 04-99, 1999.

[39] P. Del Moral and L. Miclo. On the convergence and the applications of the generalized simulated annealing. *SIAM Journal on Control and Optimization*, 37(4):1222–1250, 1999.

[40] P. Del Moral and L. Miclo. About the strong propagation of chaos for interacting particle approximations of Feynman-Kac formulae. Publications du Laboratoire de Statistique et Probabilités, no 08-00, Université Paul Sabatier, 2000.

[41] P. Del Moral and L. Miclo. Asymptotic results for genetic algorithms with applications to non linear estimation. In L. Kallel B. Naudts, editor, *Proceedings Second EvoNet Summer School on Theoretical Aspects of Evolutionary Computing*, Natural Computing Series. Springer-Verlag, 2000.

[42] P. Del Moral and L. Miclo. Branching and interacting particle systems approximations of feynman-kac formulae with applications to non-linear filtering. In J. Azéma, M. Emery, M. Ledoux, and M. Yor, editors, *Séminaire de Probabilités XXXIV*, Lecture Notes in Mathematics 1729, pages 1–145. Springer-Verlag, Berlin, 2000.

[43] P. Del Moral and L. Miclo. A Moran particle system approximation of Feynman-Kac formulae. *Stochastic Processes and their Applications*, 86:193–216, 2000.

[44] P. Del Moral and L. Miclo. Particle approximations of Lyapunov exponents connected to Feynman-Kac-Scrhödinger semigroups. Préprint, publications du Laboratoire de Statistique et Probabilités no 01-08, 2001.

[45] P. Del Moral and L. Miclo. Genealogies and increasing propagations of chaos for Feynman-Kac and genetic models. *The Annals of Applied Probability*, 2001 (to appear).

[46] P. Del Moral, J.C. Noyer, G. Rigal, and G. Salut. Traitement particulaire du signal radar, détection, estimation et reconnaissance de cibles aériennes. Technical report, LAAS/CNRS, Toulouse, 1992.

[47] P. Del Moral, J.C. Noyer, and G. Salut. Maslov optimization theory: Stochastic interpretation, particle resolution. In *Proceedings de XI Conf. Internationale sur l'Analyse et l'Optimisation des Systèmes, Ecole des Mines de Sophia Antipolis*, volume 199, pages 312–318. Lecture Note in Control and Information Sciences, Springer-Verlag, 1994.

[48] P. Del Moral, J.C. Noyer, and G. Salut. Résolution particulaire et traitement non-linéaire du signal : application radar/sonar. In *Traitement du signal*, September 1995.

[49] P. Del Moral, G. Rigal, and G. Salut. Estimation et commande optimale non linéaire. Technical Report 2, LAAS/CNRS, Toulouse, March 1992. Contract D.R.E.T.-DIGILOG.

[50] P. Del Moral, G. Rigal, and G. Salut. Estimation et commande optimale non-linéaire : un cadre unifié pour la résolution particulaire. Technical report, LAAS/CNRS, Toulouse, 1992. Contrat DRET-DIGILOG-LAAS/CNRS.

[51] P. Del Moral, G. Rigal, and G. Salut. Filtrage non-linéaire non-gaussien appliqué au recalage de plates-formes inertielles. Technical report, LAAS/CNRS, Toulouse, 1992. STCAN/DIGILOG-LAAS/CNRS contract no. A.91.77.013.

[52] P. Del Moral and G. Salut. Nonlinear filtering using monte carlo methods. *C.R. Académie des Sciences, Paris, Série I*, 320:1147–1152, 1995.

[53] P. Del Moral and G. Salut. Particle interpretation of nonlinear filtering and optimization. *Russian Journal of Mathematical Physics*, 5(3):355–372, 1998.

[54] P. Del Moral and T. Zajic. A note on Laplace-Varadhan's integral lemma. Préprint, publications du Laboratoire de Statistique et Probabilités no 01-09 2001, http://www.lsp.ups-tlse.fr/Delmoral/preprints.html.

[55] B. Delyon and O. Zeitouni. Liapunov exponents for filtering problems. In M.H.A. Davis and R.J. Elliot, editors, *Applied Stochastic Analysis*, pages 511–521. Springer-Verlag, 1991.

[56] E.B. Dynkin and A. Mandelbaum. Symmetric statistics, Poisson processes and multiple Wiener integrals. *The Annals of Statistics*, 11:739–745, 1983.

[57] A. Etheridge and P. March. A note on superprocesses. *Probability Theory and Related Fields*, 89:141–147, 1991.

[58] N.J. Gordon, D.J. Salmon, and A.F.M. Smith. Novel approach to nonlinear/non-Gaussian Bayesian state estimation. *IEE Proceedings F*, 140:107–113, 1993.

[59] C. Graham and S. Méléard. Stochastic particle approximations for generalized Boltzmann models and convergence estimates. *The Annals of Probability*, 25(1):115–132, 1997.

[60] J.H. Holland. *Adaptation in Natural and Artificial Systems.* University of Michigan Press, Ann Arbor, 1975.

[61] G. Rautu M.B. Ruskai E. Seneta G. Zbaganu J.E. Cohen, Y. Iwasa. Relative entropy under mappings by stochastic matrices. *Linear Algebra and Applications*, 179:211–235, 1993.

[62] G. Kitagawa. Monte-Carlo filter and smoother for non-Gaussian nonlinear state space models. *Journal on Computational and Graphical Statistics*, 5(1):1–25, 1996.

[63] H. Kunita. Asymptotic behavior of nonlinear filtering errors of Markov processes. *Journal of Multivariate Analysis*, 1(4):365–393, 1971.

[64] H. Kunita. Ergodic properties nonlinear filtering processes. In K.C. Alexander and J.C. Watkins, editors, *Spatial Stochastic Processes*. Birkhaüser Boston, Boston, MA, 1991.

[65] T. Kurtz and J; Xiong. Particle representations for a class of nonlinear spdes. *Stochastic Processes and their Applications*, 83:103–126, 1999.

[66] V.P. Maslov. *Méthodes opératorielles*. Edition Mir, Moscow, 1987.

[67] S. Méléard. Asymptotic behaviour of some interacting particle systems; McKean-Vlasov and Boltzmann models. In D. Talay and L. Tubaro, editors, *Probabilistic Models for Nonlinear Partial Differential Equations, Montecatini Terme, 1995*, Lecture Notes in Mathematics 1627. Springer-Verlag, 1996.

[68] L. Miclo. Remarques sur l'hypercontractivité et l'évolution de l'entropie pour des chaînes de Markov finies. In J. Azéma, M. Emery, and M. Yor, editors, *Séminaire de Probabilités XXXI*, Lecture Notes in Mathematics 1655, pages 136–167. Springer-Verlag, Berlin, 1997.

[69] D. Ocone and E Pardoux. Asymptotic stability of the optimal filter with respect to its initial condition. *Society for Industrial and Applied Mathematics. Journal on Control and Optimization*, 34:226–243, 1996.

[70] T. Shiga and H. Tanaka. Central limit theorem for a system of markovian particles with mean field interaction. *Zeitschrift für Wahrscheinlichkeitstheorie verwandte Gebiete*, 69, 1985. 439–459.

[71] L. Stettner. On invariant measures of filtering processes. In K. Helmes and N. Kohlmann, editors, *Stochastic Differential Systems, Proc. 4th Bad Honnef Conf.*, Lecture Notes in Control and Inform. Sci., pages 279–292, 1989.

[72] L. Stettner. Invariant measures of pair state/approximate filtering process. In *Colloq. Math. LXII*, pages 347–352, 1991.

[73] A. Trouvé. *Parallélisation massive du recuit simulé*. PhD thesis, Université Paris 11, Janvier 1993. Thèse de doctorat.

[74] R. Van der Boomgaard. *Mathematical morphology*. Phd thesis, University of Amsterdam, 1992.