On the Concentration Properties of Interacting Particle Processes

On the Concentration Properties of Interacting Particle Processes

Pierre Del Moral

INRIA, Université Bordeaux 1
France

Peng Hu

INRIA, Université Bordeaux 1
France

Liming Wu

Institute of Applied Mathematics
AMSS, Chinese Academy of Sciences
China

Université Blaise Pascal
France

now
the essence of knowledge

Boston – Delft

Foundations and Trends® in Machine Learning

Published, sold and distributed by:
now Publishers Inc.
PO Box 1024
Hanover, MA 02339
USA
Tel. +1-781-985-4510
www.nowpublishers.com
sales@nowpublishers.com

Outside North America:
now Publishers Inc.
PO Box 179
2600 AD Delft
The Netherlands
Tel. +31-6-51115274

The preferred citation for this publication is P. Del Moral, P. Hu, and L. Wu, On the Concentration Properties of Interacting Particle Processes, Foundation and Trends® in Machine Learning, vol 3, nos 3–4, pp 225–389, 2010.

ISBN: 978-1-60198-512-5

Foundations and Trends® in Machine Learning
Volume 3 Issues 3–4, 2010

Editorial Board

Editorial Scope

Foundations and Trends® in Machine Learning will publish survey and tutorial articles in the following topics:

- Adaptive control and signal processing
- Applications and case studies
- Behavioral, cognitive and neural learning
- Bayesian learning
- Classification and prediction
- Clustering
- Data mining
- Dimensionality reduction
- Evaluation
- Game theoretic learning
- Graphical models
- Independent component analysis
- Inductive logic programming
- Kernel methods
- Markov chain Monte Carlo
- Model choice
- Nonparametric methods
- Online learning
- Optimization
- Reinforcement learning
- Relational learning
- Robustness
- Spectral methods
- Statistical learning theory
- Variational inference
- Visualization

Information for Librarians
Foundations and Trends® in Machine Learning, 2010, Volume 3, 4 issues. ISSN paper version 1935-8237. ISSN online version 1935-8245. Also available as a combined paper and online subscription.

Foundations and Trends® in
Machine Learning
Vol. 3, Nos. 3–4 (2010) 225–389

DOI: 10.1561/2200000026

now
the essence of knowledge

On the Concentration Properties of Interacting Particle Processes

Pierre Del Moral[1], Peng Hu[2] and Liming Wu[3]

[1] *INRIA, Bordeaux-Sud Ouest Center, INRIA, Bordeaux-Sud Ouest Center & Bordeaux Mathematical Institute, Université Bordeaux 1, 351, Cours de la Libération, Talence, 33405, France, pierre.del-moral@inria.fr*
[2] *INRIA, Bordeaux-Sud Ouest Center, INRIA, Bordeaux-Sud Ouest Center & Bordeaux Mathematical Institute, Université Bordeaux 1, 351, Cours de la Libération, Talence, 33405, France, peng.hu@inria.fr*
[3] *Academy of Mathematics and Systems Science & Laboratoire de Mathématiques, Université Blaise Pascal, 63177 AUBIERE, France, Institute of Applied Mathematics, AMSS, CAS, Siyuan Building, Beijing, China, & Université Blaise Pascal, Aubière 63177, France, wuliming@amt.ac.cn, Li-Ming.Wu@math.univ-bpclermont.fr*

Abstract

This monograph presents some new concentration inequalities for Feynman-Kac particle processes. We analyze different types of stochastic particle models, including particle profile occupation measures, genealogical tree based evolution models, particle free energies, as well as backward Markov chain particle models. We illustrate these results with a series of topics related to computational physics and biology, stochastic optimization, signal processing and Bayesian statistics, and many other probabilistic machine learning algorithms. Special emphasis is given to the stochastic modeling, and to the quantitative performance analysis of a series of advanced Monte Carlo methods, including particle filters, genetic type island models, Markov bridge models, and interacting particle Markov chain Monte Carlo methodologies.

Contents

1

Stochastic Particle Methods

1.1 Introduction

Stochastic particle methods have come to play a significant role in applied probability, numerical physics, Bayesian statistics, probabilistic machine learning, and engineering sciences.

They are increasingly used to solve a variety of problems, including nonlinear filtering equations, data assimilation problems, rare event sampling, hidden Markov chain parameter estimation, stochastic control problems and financial mathematics. To name a few, They are also used in computational physics for free energy computations, and Schrödinger operator's ground states estimation problems, as well as in computational chemistry for sampling the conformation of polymers in a given solvent.

To illustrate these methods, we start with a classical filtering example. We consider a Markov chain X_k taking values in $\mathbb{R}^d$, with prior transitions given by

$$\mathbb{P}(X_k \in dx_k \mid X_{k-1} = x_{k-1}) = p_k(x_k|x_{k-1})\, dx_k, \tag{1.1}$$

Using some slight abuse of Bayesian notation, the observations Y_k are $\mathbb{R}^{d'}$-valued random variables defined in terms of the likelihood functions

$$\mathbb{P}(Y_k \in dy_k \mid X_k = x_k) = p_k(y_k|x_k)\ dy_k, \tag{1.2}$$

In the above display, dx_k and dy_k stand for the Lebesgue measures in $\mathbb{R}^d$ and $\mathbb{R}^{d'}$. To compute the conditional distribution of the signal path sequence $(X_0, \ldots, X_n)$, given the observations $(Y_0, \ldots, Y_n)$, we can use the genealogical tree model associated with a genetic type interacting particle model. This genetic algorithm is defined with mutation transitions according to 1.1, and proportional selections with regard to (w.r.t.) the fitness functions 1.2. The occupation measures of the corresponding genealogical tree provides an approximation of the desired conditional distributions of the signal. More generally, for any function f on the path space we have

$$\lim_{N\uparrow\infty} \frac{1}{N} \sum_{1}^{N} f(line_n(i)) = \mathbb{E}(f(X_0, \ldots, X_n)|Y_0 = y_0, \ldots, Y_n = y_n) \tag{1.3}$$

where $\text{line}_n(i)$ stands for the $i-$th ancestral line of the genealogical tree, at time n.

More refined particle filters can be designed, including fixed parameter estimates in hidden Markov chain models, unbiased particle estimates of the density of the observation sequence, and backward smoothing models based on complete ancestral trees. Section 2 presents a more rigorous and detailed discussion on these topics.

Rigorous understanding of these new particle Monte Carlo methodologies leads to fascinating mathematics related to Feynman-Kac path integral theory and their interacting particle interpretations [17, 20, 38]. In the last two decades, this line of research has been developed by using methods from stochastic analysis of interacting particle systems and nonlinear semigroup models in distribution spaces, but it has also generated difficult questions that cannot be addressed without developing new mathematical tools.

Let us survey some of the important challenges that arise.

For numerical applications, it is essential to obtain nonasymptotic quantitative information on the convergence of the algorithms. For instance, in the filtering problem presented at beginning of this section,

it is important to quantify the performance of the empirical particle estimate in 1.3. Asymptotic theory, including central limit theorems, moderate deviations, and large deviations principles have clearly limited practical values. An overview of these asymptotic results in the context of mean field and Feynman-Kac particle models can be found in the series of articles [13, 28, 29, 33, 41, 43].

Furthermore, when solving a given concrete problem, it is important to obtain explicit nonasymptotic error bounds estimates to ensure that the stochastic algorithm is provably correct. While non asymptotic propagation of chaos results provides some insights on the bias properties of these models, it rarely provides useful effective convergence rates.

Last but not least, it is essential to analyze the robustness properties, and more particularly the uniform performance of particle algorithms w.r.t. the time horizon. By construction, these important questions are intimately related to the stability properties of complex nonlinear Markov chain semigroups associated with the limiting measure valued process. In the filtering example illustrated in this section, the limiting measure valued process is given by the so-called nonlinear filtering equation. In this context, the stability property of these equations ensures that the optimal filter will correct any erroneous initial conditions. This line of thought has been further developed in the articles [13, 31, 38, 40], and in the books [17, 20].

Without any doubt, one of the most powerful mathematical tools to analyze the deviations of Monte Carlo based approximations is the theory of empirical processes and measure concentration theory. In the last two decades, these new tools have become one of the most important steps forward in infinite dimensional stochastic analysis, advanced machine learning techniques, as well as in the development of a statistical non asymptotic theory.

In recent years, much effort has been devoted to describing the behavior of the supremum norm of empirical functionals around the mean value of the norm. For an overview of these subjects, we refer the reader to the seminal books of Pollard [81], Van der Vaart and Wellner [93], Ledoux and Talagrand [72], the remarkable articles by Giné [56], Ledoux [70, 71], and Talagrand [90, 91, 92], and the more

recent article by Adamczak [1]. The best constants in Talagrand's concentration inequalities were obtained by Klein and Rio [67]. In this article, the authors proved the functional version of Bennett's and Bernstein's inequalities for sums of independent random variables.

Two main difficulties we encountered in applying these concentration inequalities to interacting particle models are of different order:

First, all of the concentration inequalities developed in the literature on empirical processes still involve the mean value of the supremum norm empirical functionals. In practical situations, these tail style inequalities can only be used if we have some precise information on the magnitude of the mean value of the supremum norm of the functionals.

On the other hand, the range of application of the theory of empirical processes and measure concentration theory is restricted to independent random samples, or equivalently product measures, and more recently to mixing Markov chain models. In the reverse angle, stochastic particle techniques are not based on fully independent sequences, nor on Markov chain Monte Carlo principles, but on interacting particle samples combined with complex nonlinear Markov chain semigroups. More precisely, in addition to the fact that particle models are built sequentially using conditionally independent random samples, their respective conditional distributions are still random. Also, in a nonlinear way, they strongly depend on the occupation measure of the current population.

In summary, the concentration analysis of interacting particle processes requires the development of new stochastic perturbation style techniques to control the interaction propagation and the degree of independence between the samples.

Del Moral and Ledoux [36] extend empirical processes theory to particle models. In this work, the authors proved Glivenko-Cantelli and Donsker theorems under entropy conditions, as well as nonasymptotic exponential bounds for Vapnik-Cervonenkis classes of sets or functions. Nevertheless, in practical situations these non asymptotic results tend to be a little disappointing, with very poor constants that degenerate w.r.t. the time horizon.

The second most important result on the concentration properties of the mean field particle model is found in [40]. This article is only

concerned with the finite marginal model. The authors generalize the classical Hoeffding, Bernstein and Bennett inequalities for independent random sequences to interacting particle systems.

In this monograph, we survey some of these results, and we provide new concentration inequalities for interacting empirical processes. We emphasize that this review does not give a comprehensive treatment of the theory of interacting empirical processes. To name a few missing topics, we do not discuss large deviation principles w.r.t. the strong τ-topology, Donsker type fluctuation theorems, moderate deviation principles, and continuous time models. The first two topics are developed [17], the third one is developed in [32], the last one is still an open research subject.

Here, we emphasize a single stochastic perturbation method, with second-order expansion entering the stability properties of the limiting Feynman-Kac semigroups. The concentration results attained are probably not the best possible of their kind. We have chosen to strive for just enough generality to derive useful and *uniform concentration inequalities w.r.t. the time horizon*, without having to impose complex and often unnatural regularity conditions to squeeze them into the general theory of empirical processes.

Some of the results are borrowed from [40], and many others are new. This monograph should be complemented with the books and articles [17, 20, 31, 44]. A very basic knowledge in statistics and machine learning theory will be useful, but not necessary. Good backgrounds in Markov chain theory and in stochastic semigroup analysis are necessary.

We have done our best to give a self-contained presentation, with detailed proofs. However, we assume some familiarity with Feynman-Kac models, and basic facts on the theory of Markov chains on abstract state spaces. Only in subsection 4.6.1, have we skipped the proof of some tools from convex analysis. We hope that the essential ideas are still accessible to the readers.

It is clearly not the scope of this monograph to give an exhaustive list of references to articles in computational physics, engineering sciences, and machine learning, presenting heuristic-like particle algorithms to solve a specific estimation problem. With a few exceptions, we have only provided references to articles with rigorous and well founded

mathematical treatments on particle models. We apologize in advance for possible errors, or for references that have been omitted due to the lack of accurate information.

This monograph grew from series of lectures the first author gave in the Computer Science and Communications Research Unit, of the University of Luxembourg in February and March 2011. They were reworked, with the addition of new material on the concentration of empirical processes for a course given at the Sino-French Summer Institute in Stochastic Modeling and Applications (CNRS-NSFC Joint Institute of Mathematics), held at the Academy of Mathematics and System Science, Beijing, in June 2011. The Summer Institute was ably organized by Fuzhou Gong, Ying Jiao, Gilles Pagès, and Mingyu Xu, and the members of the scientific committee, including Nicole El Karoui, Zhiming Ma, Shige Peng, Liming Wu, Jia-An Yan, and Nizar Touzi. The first author is grateful to them for giving to him the opportunity to experiment on a receptive audience with material not entirely polished.

In reworking the lectures, we have tried to resist the urge to push the analysis to general classes of mean field particle models, in the spirit of the recent joint article with E. Rio [40]. Our principal objective has been to develop just enough analysis to handle four types of Feynman-Kac interacting particle processes, namely, genetic dynamic population models, genealogical tree based algorithms, particle free energies, as well as backward Markov chain particle models. These application models do not exhaust the possible uses of the theory developed in these lectures.

1.2 A Brief Review on Particle Algorithms

Stochastic particle methods belong to the class of Monte Carlo methods. They can be thought of as a universal particle methodology for sampling complex distributions in highly dimensional state spaces.

We can distinguish two different classes of models, namely, diffusion type interacting processes, and interacting jump particle models. Feynman-Kac particle methods belongs to the second class of models, with rejection-recycling jump type interaction mechanisms. In contrast

to conventional acceptance-rejection type techniques, Feynman-Kac particle methods are equipped with an adaptive and interacting recycling strategy.

The common central feature of all the Monte Carlo particle methodologies developed so far is to solve discrete generation, or continuous time integro-differential equations in distribution spaces. The first heuristic-like description of these probabilistic techniques in mathematical physics goes back to the Los Alamos report [49], and the article by Everett and Ulam in 1948 [48], and the short article by Metropolis and Ulam [79], published in 1949.

In some instances, the flow of measures is dictated by the problem at hand. In advanced signal processing, the conditional distributions of the signal, given partial and noisy observations, are given by the so-called nonlinear filtering equation in distribution space (see for instance [15, 16, 17, 20, 38], and references therein).

Free energies and Schrödinger operator's ground states are given by the quasi-invariant distribution of a Feynman-Kac conditional distribution flow of non absorbed particles in absorbing media. We refer the reader to the articles by Cancès, Jourdain and Lelièvre [5], El Makrini, Jourdain and Lelièvre [46], Rousset [85], the pair of articles of Del Moral with Miclo [38, 39], with Doucet [19], and the book [17], and the references therein.

In mathematical biology, branching processes and infinite population models are also expressed by nonlinear parabolic type integro-differential equations. Further details on this subject can be found in the articles by Dawson and his co-authors [11, 12, 14], the works of Dynkin [45], and Le Gall [69], and more particularly the seminal book of Ethier and Kurtz [47], and the pioneering article by Feller [50].

In other instances, we formulate a given estimation problem in terms of a sequence of distributions with increasing complexity on state space models with increasing dimension. These stochastic evolutions can be related to decreasing temperature schedules in Boltzmann-Gibbs measures, multilevel decompositions for rare event excursion models on critical level sets, decreasing subsets strategies for sampling tail style distributions, and many other sequential

importance sampling plans. For a more thorough discussion on these models we refer the reader to [21].

From a purely probabilistic point of view, any flow of probability measures can be interpreted as the evolution of the laws of the random states of a Markov process. In contrast to conventional Markov chain models, the Markov transitions of these chains may depend on the distribution of the current random state. The mathematical foundations of these discrete generation models began in 1996 in [15] within the context of nonlinear filtering problems. Further analysis was developed in [38]. For a more thorough discussion on the origin and the performance analysis of these discrete generation models, we also refer the reader to the book [17], and the joint articles Del Moral with Guionnet [28, 29, 30, 31], and with Kouritzin [35].

The continuous time version of these nonlinear type Markov chain models take their origins from the 1960s, with the development of fluid mechanisms and statistical physics. We refer the reader to the pioneering works of McKean [61, 63], as well as the more recent treatments by Bellomo and Pulvirenti [3, 4], the series of articles by Graham and Méléard on interacting jump models [58, 59, 82], the articles by Méléard on Boltzmann equations [75, 76, 77, 78], and the lecture notes of Sznitman [89], and references therein.

In contrast to conventional Markov chain Monte Carlo techniques, these McKean type nonlinear Markov chain models can be thought of as perfect importance sampling strategies, in the sense that the desired *target measures coincide* at any time step with the law of the random states of a Markov chain. Unfortunately, as we mentioned above, the transitions of these chains depend on the distributions of their random states. Thus, they cannot be sampled without an additional level of approximation. One natural solution is to use a mean field particle interpretation model. These stochastic techniques belong to the class of stochastic population models, with free evolutions mechanisms, coupled with branching and/or adaptive interacting jumps. At any time step, the occupation measure of the population of individuals approximates the solution of the nonlinear equation, when the size of the system tends to ∞.

In genetic algorithms and sequential Monte Carlo literature, the reference free evolution model is interpreted as a reference sequence of twisted Markov chain samplers. These chains are used to perform the mutation/proposal transitions. As in conventional Markov chain Monte Carlo methods, the interacting jumps are interpreted as an acceptance-rejection transition, equipped with sophisticated interacting and adaptive recycling mechanism. In Bayesian statistics and engineering sciences, the resulting adaptive particle sampling model is often coined as a sequential Monte Carlo algorithm, genetic procedure, or simply a Sampling Importance Resampling method, mainly because it is based on importance sampling plans and online approximations of a flow of probability measures.

Since the 1960s, the adaptive particle recycling strategy has also been associated, in biology and engineering science, with several heuristic-like paradigms, with a proliferation of botanical names, depending on the application area in which they are considered: bootstrapping, switching, replenishing, pruning, enrichment, cloning, reconfigurations, resampling, rejuvenation, acceptance/rejection, spawning.

Of course, the idea of duplicating online better-fitted individuals and moving them one step forward to explore state-space regions is the basis of various stochastic search algorithms, such as:

Particle and bootstrap filters, Rao-Blackwell particle filters, sequential Monte Carlo methods, sequentially interacting Markov chain Monte Carlo methods, genetic type search algorithms, Gibbs cloning search techniques, interacting simulated annealing algorithms, sampling-importance resampling methods, quantum Monte Carlo walkers, adaptive population Monte Carlo sampling models, and many others evolutionary type Monte Carlo methods.

For a more detailed discussion on these models, with precise references, we refer the reader to the three books [17, 20, 44].

1.3 Feynman-Kac Path Integrals

Feynman-Kac measures represent the distribution of the paths of a Markov process, weighted by a collection of potential functions. These functional models encapsulate traditional changes of probability

measures, commonly used in importance sampling, posterior distributions in Bayesian statistics, and the optimal filter in nonlinear filtering problems.

These stochastic models are defined in terms of only two ingredients:

A Markov chain X_n, with Markov transition M_n on some measurable state spaces $(E_n, \mathcal{E}_n)$ with initial distribution η_0, and a sequence of $(0,1]$-valued potential functions G_n on the set E_n.

The Feynman-Kac path measure associated with the pairs (M_n, G_n) is the probability measure $\mathbb{Q}_n$ on the product state space

$$\mathbf{E}_n := (E_0 \times \ldots \times E_n)$$

defined by the following formula

$$d\mathbb{Q}_n := \frac{1}{\mathcal{Z}_n} \left\{ \prod_{0 \leq p < n} G_p(X_p) \right\} d\mathbb{P}_n \tag{1.4}$$

where $\mathcal{Z}_n$ is a normalizing constant and $\mathbb{P}_n$ is the distribution of the random paths

$$\mathbf{X_n} = (X_0, \ldots, X_n) \in \mathbf{E}_n$$

of the Markov process X_p from the origin $p = 0$ with initial distribution η_0, up to the current time $p = n$. We also denote by

$$\Gamma_n = \mathcal{Z}_n \mathbb{Q}_n \tag{1.5}$$

its unnormalized version.

The prototype model we have in mind is the traditional particle absorbed Markov chain model

$$X_n^c \in E_n^c := E_n \cup \{c\} \xrightarrow{\textit{absorption } \sim(1-G_n)} \widehat{X}_n^c \xrightarrow{\textit{exploration } \sim M_{n+1}} X_{n+1}^c. \tag{1.6}$$

The chain X_n^c starts at some initial state X_0^c randomly chosen with distribution η_0. During the absorption stage, we set $\widehat{X}_n^c = X_n^c$ with probability $G_n(X_n)$, otherwise we put the particle in an auxiliary cemetery state $\widehat{X}_n^c = c$. When the particle $\widehat{X}_n^c$ is still alive (that is, if we have $\widehat{X}_n^c \in E_n$), it performs an elementary move $\widehat{X}_n^c \rightsquigarrow X_{n+1}^c$ according to the

Markov transition M_{n+1}. Otherwise, the particle is absorbed and we set $X_p^c = \widehat{X}_p^c = c$, for any time $p > n$.

If we let T be the first time $\widehat{X}_n^c = c$, then we have the Feynman-Kac representation formulae

$$\mathbb{Q}_n = \mathrm{Law}((X_0^c, \ldots, X_n^c) \mid T \geq n) \quad \text{and} \quad \mathcal{Z}_n = \mathrm{Proba}(T \geq n).$$

For a more thorough discussion on the variety of application domains of Feynman-Kac models, we refer the reader to Section 2.

We also denote by η_n and γ_n, the n-th time marginal of $\mathbb{Q}_n$ and Γ_n. It is a simple exercise to check that

$$\gamma_n = \gamma_{n-1} Q_n \quad \text{and} \quad \eta_{n+1} = \Phi_{n+1}(\eta_n) := \Psi_{G_n}(\eta_n) M_{n+1} \tag{1.7}$$

with the positive integral operator

$$Q_n(x, dy) = G_{n-1}(x)\ M_n(x, dy)$$

and the Boltzmann-Gibbs transformation

$$\Psi_{G_n}(\eta_n)(dx) = \frac{1}{\eta_n(G_n)} G_n(x) \eta_n(dx). \tag{1.8}$$

In addition, the normalizing constants $\mathcal{Z}_n$ can be expressed in terms of the flow of marginal measures η_p, from the origin $p = 0$ up to the current time n, with the following multiplicative formulae:

$$\mathcal{Z}_n := \gamma_n(\mathbb{1}) = \mathbb{E}\left(\prod_{0 \leq p < n} G_p(X_p) \right) = \prod_{0 \leq p < n} \eta_p(G_p). \tag{1.9}$$

This multiplicative formula is easily checked using the induction

$$\gamma_{n+1}(1) = \gamma_n(G_n) = \eta_n(G_n) \gamma_n(1).$$

The abstract formulae discussed above are more general than they may appear. For instance, they can be used to analyze, without further work, path spaces models, including historical processes or transition space models, as well as finite excursion models. These functional models also encapsulate quenched Feynman-Kac models, Brownian type bridges and linear Gaussian Markov chains conditioned on starting and end points.

For a more thorough discussion on these path space models, we refer the reader to subsections 2.4 and 2.6, Chapters 11–12 in the book [17], as well as to the Section 2, in this monograph.

When the Markov transitions M_n are absolutely continuous with respect to some measures λ_n on E_n, and for any $(x,y) \in (E_{n-1} \times E_n)$ we have

$$H_n(x,y) := \frac{dM_n(x,.)}{d\lambda_n}(y) > 0. \tag{1.10}$$

We also have the following backward formula

$$\mathbb{Q}_n(d(x_0,\ldots,x_n)) = \eta_n(dx_n) \prod_{q=1}^{n} \mathbb{M}_{q,\eta_{q-1}}(x_q, dx_{q-1}) \tag{1.11}$$

with the collection of Markov transitions defined by

$$\mathbb{M}_{n+1,\eta_n}(x,dy) \propto G_n(y) H_{n+1}(y,x) \eta_n(dy). \tag{1.12}$$

The proof of this formula is postponed to subsection 3.2.

Before launching into the description of the particle approximation of these models, we end this subsection with some connexions between discrete generation Feynman-Kac models and more conventional continuous time models arising in physics and scientific computing.

The Feynman-Kac models presented above play a central role in the numerical analysis of certain partial differential equations, offering a natural way to solve these functional integral models by simulating random paths of stochastic processes. These Feynman-Kac models were originally presented by Mark Kac in 1949 [66] for continuous time processes.

These continuous time models are used in molecular chemistry and computational physics to calculate the ground state energy of some Hamiltonian operators associated with some potential function V describing the energy of a molecular configuration (see, for instance, [5, 17, 39, 46, 85], and references therein). To better connect these partial differential equation models with (1.4), let us assume that $M_n(x_{n-1}, dx_n)$ is the Markov probability transition $X_n = x_n \rightsquigarrow X_{n+1} = x_{n+1}$ coming from a discretization in time $X_n = X'_{t_n}$ of a continuous time E-valued Markov process X'_t on a given time mesh $(t_n)_{n\geq 0}$ with

a given time step $(t_n - t_{n-1}) = \Delta t$. For potential functions of the form $G_n = e^{-V\Delta t}$, the measures $\mathbb{Q}_n \simeq_{\Delta t \to 0} \mathbb{Q}_{t_n}$ represent the time discretization of the following distribution:

$$d\mathbb{Q}_t = \frac{1}{\mathcal{Z}_t} \exp\left(-\int_0^t V(X'_s)ds\right) d\mathbb{P}_t^{X'}$$

where $\mathbb{P}_t^{X'}$ stands for the distribution of the random paths $(X'_s)_{0\leq s\leq t}$ with a given infinitesimal generator L. The marginal distributions γ_t at time t of the unnormalized measures $\mathcal{Z}_t\ d\mathbb{Q}_t$ are the solution of the so-called imaginary time Schrödinger equation, given in weak formulation on sufficiently regular function f by the following integro–differential equation

$$\frac{d}{dt}\ \gamma_t(f) := \gamma_t(L^V(f)) \quad \text{with} \quad L^V = L - V.$$

The errors introduced by the discretization of the time are well understood for regular models; we refer the interested reader to [34, 42, 68, 80] in the context of nonlinear filtering.

1.4 Interacting Particle Systems

Our aim here is to design an interacting particle approximation of the Feynman-Kac measures introduced in the previous subsection. These particle methods can be interpreted in different ways, depending on the application domain in which they are considered.

In the filtering example presented at the beginning of this monograph, these particle algorithms can be seen as a stochastic adaptive fixed approximation of the filtering equations. From a purely statistical point of view, these algorithms can also be seen as a sophisticated acceptance-rejection technique with an interacting recycling transition.

The particle model is defined as follows:

We start with a population of N candidate possible solutions $(\xi_0^1, \ldots, \xi_0^N)$ randomly chosen w.r.t. some distribution η_0.

The coordinates ξ_0^i are also called individuals or phenotypes, with $1 \leq N$. The random evolution of the particles is decomposed into two main steps : the free exploration and the adaptive selection transition.

During the updating-selection stage, multiple individuals in the current population $(\xi_n^1, \ldots, \xi_n^N)$ at time $n \in \mathbb{N}$ are stochastically selected based on the fitness function G_n. In practice, we choose a random proportion B_n^i of an existing solution ξ_n^i in the current population with a mean value $\propto G_n(\xi_n^i)$ to breed a brand new generation of "improved" solutions $(\widehat{\xi}_n^1, \ldots, \widehat{\xi}_n^N)$. For instance, for every index i, with a probability $\epsilon_n G_n(\xi_n^i)$, we set $\widehat{\xi}_n^i = \xi_n^i$, otherwise we replace ξ_n^i with a new individual $\widehat{\xi}_n^i = \xi_n^j$ randomly chosen from the whole population with a probability proportional to $G_n(\xi_n^j)$. The parameter $\epsilon_n \geq 0$ is a tuning parameter that must satisfy the constraint $\epsilon_n G_n(\xi_n^i) \leq 1$, for every $1 \leq i \leq N$. During the prediction-mutation stage, every selected individual $\widehat{\xi}_n^i$ moves to a new solution $\xi_{n+1}^i = x$ randomly chosen in E_{n+1}, with a distribution $M_{n+1}(\widehat{\xi}_n^i, dx)$.

If we interpret the updating-selection transition as a birth and death process, then the important notion of the ancestral line of a current individual arises. More precisely, when a particle $\widehat{\xi}_{n-1}^i \longrightarrow \xi_n^i$ evolves to a new location ξ_n^i, we can interpret $\widehat{\xi}_{n-1}^i$ as the parent of ξ_n^i. Looking backwards in time and recalling that the particle $\widehat{\xi}_{n-1}^i$ has selected a site ξ_{n-1}^j in the configuration at time $(n-1)$, we can interpret this site ξ_{n-1}^j as the parent of $\widehat{\xi}_{n-1}^i$ and therefore as the ancestor denoted $\xi_{n-1,n}^i$ at level $(n-1)$ of ξ_n^i. Running backwards in time we may trace the whole ancestral line as

$$\xi_{0,n}^i \longleftarrow \xi_{1,n}^i \longleftarrow \cdots \longleftarrow \xi_{n-1,n}^i \longleftarrow \xi_{n,n}^i = \xi_n^i. \tag{1.13}$$

Most of the terminology we have used is drawn from filtering and genetic evolution theories.

In the filtering example presented in the subsection 1.1, the former particle model is dictated by the two steps prediction-updating learning equations of the conditional distributions of the signal process X_k, given some noisy and partial observations Y_k. In this setting, the potential functions represent the likelihood function of the current observation, while the free exploration transitions are related to the Markov transitions of the signal process. More formally, using the notation we used in example (1.2), we have:

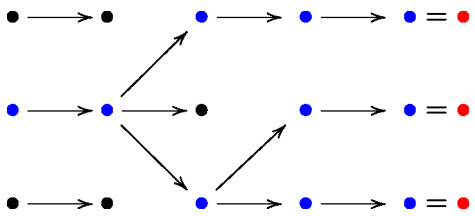
$$dp_k(x_k|x_{k-1}) = M_k(x_{k-1}|x_k)$$

and

$$p_k(y_k|x_k) = G_k(x_k).$$

In biology, the mutation-selection particle model presented above is used to mimic genetic evolutions of biological organisms and, more generally, natural evolution processes. For instance, in gene analysis, each population of individuals represents a chromosome and each individual particle is called a gene. In this setting the fitness potential function is usually time-homogeneous and it represents the quality and the adaptation potential value of the set of genes in a chromosome [62]. These particle algorithms are also used in population analysis to model changes in the structure of population in time and in space.

The different types of particle approximation measures associated with the genetic type particle model described above are summarized in the following synthetic picture corresponding to the case $N = 3$.



In the next four subsections we give an overview of the four particle approximation measures that can be extracted from the interacting population evolution model described above. We also provide some basic formulation of the concentration inequalities that will be treated in greater detail later. As a service to the reader we also provide precise pointers to their location within each section of the monograph.

We have already mentioned that the proofs of these results are quite subtle. In the further development of the next subsections, c_1 *stands for a finite constant related to* the bias *of the particle model, while* c_2 *is related to* the variance *of the scheme. The value of these constants may vary from one line to another, but in all the situations they do not depend on the time parameter.*

The precise form of the constants in these exponential inequalities depends on the contraction properties of Feynman-Kac flows. Our

stochastic analysis requires us to combine the stability properties of the nonlinear semigroup of the Feynman-Kac distribution flow η_n, with the deep convergence results of empirical processes theory associated with interacting random samples.

1.4.1 Last Population Models

The occupation measures of the current population, represented by the red dots in the above figure

$$\eta_n^N := \frac{1}{N}\sum_{i=1}^{N} \delta_{\xi_n^i}$$

converge to the n-th time marginals η_n of the Feynman-Kac measures $\mathbb{Q}_n$. We shall measure the performance of these particle estimates through several concentration inequalities, with a special emphasis on uniform inequalities w.r.t. the time parameter. Our results will basically be stated as follows.

1) For any time horizon $n \geq 0$, any bounded function f, any $N \geq 1$, and for any $x \geq 0$, the probability of the event

$$[\eta_n^N - \eta_n](f) \leq \frac{c_1}{N}(1 + x + \sqrt{x}) + \frac{c_2}{\sqrt{N}}\sqrt{x}$$

is greater than $1 - e^{-x}$.

We have already mentioned one important consequence of these uniform concentration inequalities for time homogeneous Feynman-Kac models. Under some regularity conditions, the flow of measures η_n tends to some fixed point distribution η_∞, in the sense that

$$\|\eta_n - \eta_\infty\|_{\mathrm{tv}} \leq c_3\ e^{-\delta n} \tag{1.14}$$

for some finite positive constants c_3 and δ. In the above display $\|\nu - \mu\|_{tv}$ stands for the total variation distance. The connexions between these limiting measures and the top of the spectrum of Schrödinger operators is discussed in subsection 2.7.1. We also refer the reader to subsection 2.7.2 for a discussion on these quasi-invariant measures and Yaglom limits. Quantitative contraction theorems for Feynman-Kac semigroups are developed in subsection 3.4.2. As a direct consequence of the above inequalities, we find that for any $x \geq 0$, the

probability of the following events is greater than $1 - e^{-x}$:

$$[\eta_n^N - \eta_\infty](f) \leq \frac{c_1}{N}(1 + x + \sqrt{x}) + \frac{c_2}{\sqrt{N}}\sqrt{x} + c_3 e^{-\delta n}.$$

2) For any $x = (x_i)_{1 \leq i \leq d} \in E_n = \mathbb{R}^d$, we set $(-\infty, x] = \prod_{i=1}^d (-\infty, x_i]$ and we consider the repartition functions

$$F_n(x) = \eta_n(1_{(-\infty,x]}) \quad \text{and} \quad F_n^N(x) = \eta_n^N(1_{(-\infty,x]}).$$

The probability of the following event

$$\sqrt{N}\|F_n^N - F_n\| \leq c\sqrt{d(x+1)}$$

is greater than $1 - e^{-x}$, for any $x \geq 0$, for some universal constant $c < \infty$ that does not depend on the dimension, nor on the time parameter. In the above display $\|F\| = \sup_x |F(x)|$ stands for the uniform norm. Furthermore, under the stability properties (1.14), if we set

$$F_\infty(x) = \eta_\infty(1_{(-\infty,x]})$$

then, the probability of the following event

$$\|F_n^N - F_\infty\| \leq \frac{c}{\sqrt{N}}\sqrt{d(x+1)} + c_3 e^{-\delta n}$$

is greater than $1 - e^{-x}$, for any $x \geq 0$, for some universal constant $c < \infty$ that does not depend on the dimension.

For more precise statements, we refer the reader to Corollary 6.4, and Corollary 6.9, respectively.

The concentration properties of the particle measures η_n^N around their limiting values are developed in Section 6. In subsection 6.3, we design a stochastic perturbation analysis that allows us to enter the stability properties of the limiting Feynman-Kac semigroup. Finite marginal models are discussed in subsection 6.4.1. Subsection 6.4.2 is concerned with the concentration inequalities of interacting particle processes w.r.t. some collection of functions.

1.4.2 Particle Free Energy Models

Mimicking the multiplicative formula (1.9), we set

$$\mathcal{Z}_n^N = \prod_{0 \leq p < n} \eta_p^N(G_p) \quad \text{and} \quad \gamma_n^N(dx) = \mathcal{Z}_n^N \times \eta_n^N(dx). \tag{1.15}$$

We have already mentioned that these rather complex particle models provide an unbiased estimate of the unnormalized measures. That is, we have that

$$\mathbb{E}\left(\eta_n^N(f_n)\prod_{0\leq p<n}\eta_p^N(G_p)\right)=\mathbb{E}\left(f_n(X_n)\prod_{0\leq p<n}G_p(X_p)\right). \tag{1.16}$$

The concentration properties of the *unbiased* particle free energies $\mathcal{Z}_n^N$ around their limiting values $\mathcal{Z}_n$ are developed in subsection 6.5. Our results will basically be stated as follows.

For any $N\geq 1$, and any $\epsilon\in\{+1,-1\}$, the probability of each of the following events

$$\frac{\epsilon}{n}\log\frac{\mathcal{Z}_n^N}{\mathcal{Z}_n}\leq\frac{c_1}{N}(1+x+\sqrt{x})+\frac{c_2}{\sqrt{N}}\sqrt{x}$$

is greater than $1-e^{-x}$. A more precise statement is provided in Corollary 6.14.

1.4.3 Genealogical Tree Model

The occupation measure of the N-genealogical tree model represented by the lines linking the blue dots converges as $N\to\infty$ to the distribution $\mathbb{Q}_n$

$$\lim_{N\to\infty}\frac{1}{N}\sum_{i=1}^{N}\delta_{(\xi_{0,n}^i,\xi_{1,n}^i,\ldots,\xi_{n,n}^i)}=\mathbb{Q}_n. \tag{1.17}$$

Our concentration inequalities will basically be stated as follows. A more precise statement is provided in Corollary 6.5.

For any $n\geq 0$, any bounded function $\mathbf{f_n}$ on the path space $\mathbf{E_n}$, such that (s.t.) $\|\mathbf{f_n}\|\leq 1$, and any $N\geq 1$, the probability of each of the following events

$$\left[\frac{1}{N}\sum_{i=1}^{N}\mathbf{f_n}(\xi_{0,n}^i,\xi_{1,n}^i,\ldots,\xi_{n,n}^i)-\mathbb{Q}_n(\mathbf{f_n})\right]$$
$$\leq c_1\frac{n+1}{N}(1+x+\sqrt{x})+c_2\sqrt{\frac{(n+1)}{N}}\sqrt{x}$$

is greater than $1-e^{-x}$.

The concentration properties of genealogical tree occupation measures can be derived more or less directly from those of the current population models. This rather surprising assertion comes from the fact that the n-th time marginal η_n of a Feynman-Kac measure associated with a reference historical Markov process has the same form as in the measure (1.4). This equivalence principle between $\mathbb{Q}_n$ and the marginal measures are developed in subsection 3.2, dedicated to historical Feynman-Kac models.

Using these properties, we prove concentration properties for interacting empirical processes associated with genealogical tree models. Our concentration inequalities will basically be stated as follows. A more precise statement is provided in subsection 6.4.2. We let $\mathcal{F}_n$ be the set of product functions of cell indicators in the path space $\mathbf{E_n} = (\mathbb{R}^{d_0} \times \ldots, \times \mathbb{R}^{d_n})$, for some $d_p \geq 1$, $p \geq 0$. We also denote by η_n^N the occupation measure of the genealogical tree model. In this notation, the probability of the following event

$$\sup_{\mathbf{f_n} \in \mathcal{F}_n} |\eta_n^N(\mathbf{f_n}) - \mathbb{Q}_n(\mathbf{f_n})| \leq c(n+1)\sqrt{\frac{\sum_{0 \leq p \leq n} d_p}{N}(x+1)}$$

is greater than $1 - e^{-x}$, for any $x \geq 0$, for some universal constant $c < \infty$ that does not depend on the dimension.

1.4.4 Complete Genealogical Tree Models

Mimicking the backward model (1.11) and the above formulae, we set

$$\Gamma_n^N = \mathcal{Z}_n^N \times \mathbb{Q}_n^N \tag{1.18}$$

with

$$\mathbb{Q}_n^N(d(x_0, \ldots, x_n)) = \eta_n^N(dx_n) \prod_{q=1}^{n} \mathbb{M}_{q, \eta_{q-1}^N}(x_q, dx_{q-1}).$$

Notice that the computation of sums w.r.t. these particle measures are reduced to summations over the particle locations ξ_n^i. It is therefore natural to identify a population of individuals $(\xi_n^1, \ldots, \xi_n^N)$ at time n to the ordered set of indexes $\{1, \ldots, N\}$. In this case, the occupation

measures and the functions are identified with the following line and column vectors

$$\eta_n^N := \left[\frac{1}{N}, \ldots, \frac{1}{N}\right] \quad \text{and} \quad \mathrm{f}_n := \begin{pmatrix} f_n(\xi_n^1) \\ \vdots \\ f_n(\xi_n^N) \end{pmatrix}$$

and the matrices $\mathbb{M}_{n,\eta_{n-1}^N}$ by the $(N \times N)$ matrices

$$\mathbb{M}_{n,\eta_{n-1}^N} := \begin{pmatrix} \mathbb{M}_{n,\eta_{n-1}^N}(\xi_n^1, \xi_{n-1}^1) & \cdots & \mathbb{M}_{n,\eta_{n-1}^N}(\xi_n^1, \xi_{n-1}^N) \\ \vdots & \vdots & \vdots \\ \mathbb{M}_{n,\eta_{n-1}^N}(\xi_n^N, \xi_{n-1}^1) & \cdots & \mathbb{M}_{n,\eta_{n-1}^N}(\xi_n^N, \xi_{n-1}^N) \end{pmatrix} \tag{1.19}$$

with the (i,j)-entries

$$\mathbb{M}_{n,\eta_{n-1}^N}(\xi_n^i, \xi_{n-1}^j) = \frac{G_{n-1}(\xi_{n-1}^j) H_n(\xi_{n-1}^j, \xi_n^i)}{\sum_{k=1}^N G_{n-1}(\xi_{n-1}^k) H_n(\xi_{n-1}^k, \xi_n^i)}.$$

For instance, the $\mathbb{Q}_n$-integration of normalized additive linear functionals of the form

$$\mathbf{f}_n(x_0, \ldots, x_n) = \frac{1}{n+1} \sum_{0 \le p \le n} f_p(x_p) \tag{1.20}$$

is given the particle matrix approximation model

$$\mathbb{Q}_n^N(\mathbf{f}_n) = \frac{1}{n+1} \sum_{0 \le p \le n} \eta_n^N \mathbb{M}_{n,\eta_{n-1}^N} \mathbb{M}_{n-1,\eta_{n-2}^N} \ldots \mathbb{M}_{p+1,\eta_p^N}(f_p).$$

These type of additive functionals arise in the calculation of the sensitivity measures discussed in subsection 2.4.1.

The concentration properties of the particle measures $\mathbb{Q}_n^N$ around the Feynman-Kac measures $\mathbb{Q}_n$ are developed in subsection 6.6. Special emphasis is given to the additive functional models (1.20). In subsection 6.6.3, we extend the stochastic perturbation methodology developed in subsection 6.3 for time marginal models to the particle backward Markov chain associated with the random stochastic matrices (1.19). This technique allows us to enter not only the stability properties of the limiting Feynman-Kac semigroup, but also those of the particle backward Markov chain model.

Our concentration inequalities will basically be stated as follows. A more precise statement is provided in Corollary 6.20 and in Corollary 6.24.

For any $n \geq 0$, any normalized additive functional of the form (1.20), with $\max_{0 \leq p \leq n} \|f_p\| \leq 1$, and any $N \geq 1$, the probability of each of the following events

$$[\mathbb{Q}_n^N - \mathbb{Q}_n](\overline{\mathbf{f}}_n) \leq c_1 \frac{1}{N} \left(1 + x + \sqrt{x}\right) + c_2 \sqrt{\frac{x}{N(n+1)}}$$

is greater than $1 - e^{-x}$.

For any $a = (a_i)_{1 \leq i \leq d} \in E_n = \mathbb{R}^d$, we denote by C_a the cell

$$C_a := (-\infty, a] = \prod_{i=1}^{d} (-\infty, a_i]$$

and $\mathbf{f}_{a,n}$ the additive functional

$$\mathbf{f}_{a,n}(x_0, \ldots, x_n) = \frac{1}{n+1} \sum_{0 \leq p \leq n} 1_{(-\infty,a]}(x_p).$$

The probability of the following event

$$\sup_{a \in \mathbb{R}^d} |\mathbb{Q}_n^N(\mathbf{f}_{a,n}) - \mathbb{Q}_n(\mathbf{f}_{a,n})| \leq c \sqrt{\frac{d}{N}(x+1)}$$

is greater than $1 - e^{-x}$, for any $x \geq 0$, for some constant $c < \infty$ that does not depend on the dimension, nor on the time horizon.

Remark 1.1. One way to turn all of these inequalities into of Bernstein style concentration inequalities is as follows. For any exponential inequality of the form

$$\forall x \geq 0 \quad \mathbb{P}(X \leq ax + \sqrt{2bx} + c) \leq 1 - e^{-x}$$

for some non negative constants (a, b, c), we also have

$$\forall y \geq 0 \quad \mathbb{P}(X \leq y + c) \leq 1 - \exp\left(-\frac{y^2}{2\,(b + ay)}\right).$$

A proof of this result is provided in Lemma 4.13.

1.5 Basic Notation

Here we provide some background from stochastic analysis and integral operator theory, which we require for our proofs. Most of the results with detailed proofs can be located in the book [17] on Feynman-Kac formulae and interacting particle methods. Our proofs also contain cross-references to this rather well known material, so the reader may wish to skip this subsection and proceed directly to Section 2, which is dedicated to some application domains of Feynman-Kac models.

1.5.1 Integral Operators

We denote respectively by $\mathcal{M}(E)$, $\mathcal{M}_0(E)$, $\mathcal{P}(E)$, and $\mathcal{B}(E)$, the set of all finite signed measures on some measurable space $(E,\mathcal{E})$, the convex subset of measures with null mass, the set of all probability measures, and the Banach space of all bounded and measurable functions f equipped with the uniform norm $\|f\|$. We also denote by $\mathrm{Osc}_1(E)$, and by $\mathcal{B}_1(E)$ the set of $\mathcal{E}$-measurable functions f with oscillations $\mathrm{osc}(f) \leq 1$, and respectively with $\|f\| \leq 1$. We let

$$\mu(f) = \int \mu(dx) f(x)$$

be the Lebesgue integral of a function $f \in \mathcal{B}(E)$, with respect to a measure $\mu \in \mathcal{M}(E)$.

We recall that the total variation distance on $\mathcal{M}(E)$ is defined for any $\mu \in \mathcal{M}(E)$ by

$$\|\mu\|_{\mathrm{tv}} = \frac{1}{2} \sup_{(A,B)\in\mathcal{E}^2} (\mu(A) - \mu(B)).$$

We recall that a bounded integral operator M turned from a measurable space $(E,\mathcal{E})$ into an auxiliary measurable space $(F,\mathcal{F})$ is an operator $f \mapsto M(f)$ from $\mathcal{B}(F)$ into $\mathcal{B}(E)$ such that the functions

$$M(f)(x) := \int_F M(x,dy) f(y)$$

are $\mathcal{E}$-measurable and bounded, for any $f \in \mathcal{B}(F)$. A Markov kernel is a positive and bounded integral operator M with $M(1) = 1$. Given a pair

of bounded integral operators (M_1, M_2), we let $(M_1 M_2)$ represent the composition operator defined by $(M_1 M_2)(f) = M_1(M_2(f))$. For time homogeneous state spaces, we denote by $M^m = M^{m-1}M = MM^{m-1}$ the m-th composition of a given bounded integral operator M, with $m \geq 1$. A bounded integral operator M turned from a measurable space $(E, \mathcal{E})$ into an auxiliary measurable space $(F, \mathcal{F})$ also generates a dual operator

$$\mu(dx) \mapsto (\mu M)(dx) = \int \mu(dy) M(y, dx)$$

from $\mathcal{M}(E)$ into $\mathcal{M}(F)$ defined by $(\mu M)(f) := \mu(M(f))$. We also used the notation

$$K([f - K(f)]^2)(x) := K([f - K(f)(x)]^2)(x)$$

for some bounded integral operator K and some bounded function f.

We prefer to avoid unnecessary abstraction and technical assumptions, so we frame the standing assumption that all the test functions are in the unit sphere, and the integral operators, and all the random variables are sufficiently regular so that we are justified in computing integral transport equations, regular versions of conditional expectations, and so forth.

1.5.2 Contraction Coefficients

When the bounded integral operator M has a constant mass, that is, when $M(1)(x) = M(1)(y)$ for any $(x, y) \in E^2$, the operator $\mu \mapsto \mu M$ maps $\mathcal{M}_0(E)$ into $\mathcal{M}_0(F)$. In this situation, we let $\beta(M)$ be the Dobrushin coefficient of a bounded integral operator M defined by the formula

$$\beta(M) := \sup \{\mathrm{osc}(M(f));\ f \in \mathrm{Osc}(F)\}.$$

Notice that $\beta(M)$ is the operator norm of M on $\mathcal{M}_0(E)$, and we have the equivalent formulations

$$\begin{aligned} \beta(M) &= \sup \{\|M(x, .) - M(y, .)\|_{\mathrm{tv}} \,;\, (x, y) \in E^2\} \\ &= \sup_{\mu \in \mathcal{M}_0(E)} \|\mu M\|_{\mathrm{tv}} / \|\mu\|_{\mathrm{tv}}. \end{aligned}$$

A detailed proof of these well known formulae can be found in [17].

Given a positive and bounded potential function G on E, we also denote by Ψ_G the Boltzmann-Gibbs mapping from $\mathcal{P}(E)$ into itself, defined for any $\mu \in \mathcal{P}(E)$ by

$$\Psi_G(\mu)(dx) = \frac{1}{\mu(G)} G(x)\mu(dx).$$

For $[0,1]$-valued potential functions, we also mention that $\Psi_G(\mu)$ can be expressed as a non linear Markov transport equation

$$\Psi_G(\mu) = \mu S_{\mu,G} \tag{1.21}$$

with the Markov transitions

$$S_{\mu,G}(x,dy) = G(x)\delta_x(dy) + (1 - G(x))\Psi_G(\mu)(dy).$$

We notice that

$$\Psi_G(\mu) - \Psi_G(\nu) = (\mu - \nu)S_\mu + \nu(S_\mu - S_\nu)$$

and

$$\nu(S_\mu - S_\nu) = (1 - \nu(G))[\Psi_G(\mu) - \Psi_G(\nu)]$$

from which we find the formula

$$\Psi_G(\mu) - \Psi_G(\nu) = \frac{1}{\nu(G)}\ (\mu - \nu)S_\mu.$$

In addition, using the fact that

$$\forall (x,A) \in (E,\mathcal{E}) \quad S_\mu(x,A) \geq (1 - \|G\|)\Psi_G(\mu)(A)$$

we prove that $\beta(S_\mu) \leq \|G\|$ and

$$\|\Psi_G(\mu) - \Psi_G(\nu)\|_{\mathrm{tv}} \leq \frac{\|G\|}{\mu(G) \vee \nu(G)}\ \|\mu - \nu\|_{\mathrm{tv}}.$$

If we set $\Phi(\mu) = \Psi_G(\mu)M$, for some Markov transition M, then we have the decomposition

$$\Phi(\mu) - \Phi(\nu) = \frac{1}{\nu(G)}\ (\mu - \nu)S_\mu M \tag{1.22}$$

for any couple of measures ν, μ on E. From the previous discussion, we also find the following Lipschitz estimates

$$\|\Phi(\mu) - \Phi(\nu)\|_{\mathrm{tv}} \leq \frac{\|G\|}{\mu(G) \vee \nu(G)} \beta(M) \|\mu - \nu\|_{\mathrm{tv}}. \tag{1.23}$$

We end this subsection with an interesting contraction property of a Markov transition

$$M_G(x, dy) = \frac{M(x, dy)G(y)}{M(G)(x)} = \Psi_G(\delta_x M)(dy) \tag{1.24}$$

associated with a $]0,1]$-valued potential function G, with

$$g = \sup_{x,y} G(x)/G(y) < \infty. \tag{1.25}$$

It is easily checked that

$$\begin{aligned} |M_G(f)(x) - M_G(f)(y)| &= |\Psi_G(\delta_x M)(f) - \Psi_G(\delta_y M)(f)| \\ &\leq g \|\delta_x M - \delta_y M\|_{\mathrm{tv}} \end{aligned}$$

from which we conclude that

$$\beta(M_G) \leq g \beta(M). \tag{1.26}$$

1.5.3 Orlicz Norms and Gaussian Moments

We let $\pi_\psi[Y]$ be the Orlicz norm of an $\mathbb{R}$-valued random variable Y associated with the convex function $\psi(u) = e^{u^2} - 1$, and defined by

$$\pi_\psi(Y) = \inf \{a \in (0, \infty) : \mathbb{E}(\psi(|Y|/a)) \leq 1\}$$

with the convention $\inf_\emptyset = \infty$. Notice that

$$\pi_\psi(Y) \leq c \Longleftrightarrow \mathbb{E}(\psi(Y/c)) \leq 1.$$

For instance, the Orlicz norm of a Gaussian and centered random variable U, s.t. $E(U^2) = 1$, is given by $\pi_\psi(U) = \sqrt{8/3}$. We also recall that

$$\begin{aligned} \mathbb{E}(U^{2m}) &= b(2m)^{2m} := (2m)_m \, 2^{-m} \\ \mathbb{E}(|U|^{2m+1}) &\leq b(2m+1)^{2m+1} := \frac{(2m+1)_{(m+1)}}{\sqrt{m + 1/2}} \, 2^{-(m+1/2)} \end{aligned} \tag{1.27}$$

with $(q+p)_p := (q+p)!/q!$. The second assertion comes from the fact that

$$\mathbb{E}(U^{2m+1})^2 \leq \mathbb{E}(U^{2m})\mathbb{E}(U^{2(m+1)})$$

and therefore

$$\begin{aligned} b(2m+1)^{2(2m+1)} &= \mathbb{E}(U^{2m})\mathbb{E}(U^{2(m+1)}) \\ &= 2^{-(2m+1)}\ (2m)_m\ (2(m+1))_{(m+1)}. \end{aligned}$$

This formula is a direct consequence of the following decompositions

$$(2(m+1))_{(m+1)} = \frac{(2(m+1))!}{(m+1)!} = 2\frac{(2m+1)!}{m!} = 2(2m+1)_{(m+1)}$$

and

$$(2m)_m = \frac{1}{2m+1}\ \frac{(2m+1)!}{m!} = \frac{1}{2m+1}\ (2m+1)_{(m+1)}.$$

We also mention that

$$b(m) \leq b(2m). \tag{1.28}$$

Indeed, for even numbers $m = 2p$ we have

$$b(m)^{2m} = b(2p)^{4p} = \mathbb{E}(U^{2p})^2 \leq \mathbb{E}(U^{4p}) = b(4p)^{4p} = b(2m)^{2m}$$

and for odd numbers $m = (2p+1)$, we have

$$\begin{aligned} b(m)^{2m} &= b(2p+1)^{2(2p+1)} = \mathbb{E}(U^{2p})\mathbb{E}(U^{2(p+1)}) \\ &\leq \mathbb{E}\left((U^{2p})^{\frac{(2p+1)}{p}}\right)^{\frac{p}{2p+1}} \mathbb{E}\left((U^{2(p+1)})^{\frac{(2p+1)}{p+1}}\right)^{\frac{p+1}{2p+1}} \\ &= \mathbb{E}(U^{2(2p+1)}) = b(2(2p+1))^{2(2p+1)} = b(2m)^{2m}. \end{aligned}$$

References

[1] R. Adamczak, "A tail inequality for suprema of unbounded empirical processes with applications to Markov chains," *Electronic Journal of Probability*, vol. 13, no. 34, pp. 1000–1034, 2008.

[2] C. Andrieu, A. Doucet, and R. Holenstein, "Particle Markov chain Monte Carlo methods," *Journal Royal Statistical Society B*, vol. 72, no. 3, pp. 269–342, 2010.

[3] N. Bellomo and M. Pulvirenti, "Generalized kinetic models in applied sciences," in *Modeling in Applied Sciences,* Modeling and Simulation in Science, Engineering, and Technology, pp. 1–19, Birkhäuser, Boston, 2000.

[4] N. Bellomo and M. Pulvirenti, eds., *Modeling in Applied Sciences,* Modeling and Simulation in Science, Engineering, and Technology. Birkhäuser, Boston, 2000.

[5] E. Cancès, B. Jourdain, and T. Lelièvre, "Quantum Monte Carlo simulations of fermions. A mathematical analysis of the fixed-node approximation," *ESAIM: M2AN*, vol. 16, no. 9, pp. 1403–1449, 2006.

[6] R. Carmona, J.-P. Fouque, and D. Vestal, "Interacting particle systems for the computation of rare credit portfolio Losses," *Finance and Stochastics*, vol. 13, no. 4, pp. 613–633, 2009.

[7] F. Caron, P. Del Moral, A. Tantar, and E. Tantar, "Simulation particulaire, Dimensionnement en conditions extrêmes," Research contract no. 2010-IFREMER-01 with IFREMER, (92p) September, 2010.

[8] F. Cerou, P. Del Moral, Fr. Le Gland, and P. Lezaud, "Genealogical models in entrance times rare event analysis," *Alea, Latin American Journal of Probability And Mathematical Statistics*, vol. 1, pp. 181–203, 2006.

[9] F. Cerou, P. Del Moral, T. Furon, and A. Guyader, "Sequential Monte Carlo for rare event estimation," *Statistics and Computing*, to appear, 2011.

[10] F. Cerou, P. Del Moral, A. Guyader, F. LeGland, P. Lezaud, and H. Topart, "Some recent improvements to importance splitting," in (preliminary version), *Proceedings of 6th International Workshop on Rare Event Simulation*, Bamberg, Germany, 2006.

[11] D. A. Dawson, "Stochastic evolution equations and related measure processes," *J. Multivariate Anal.*, vol. 3, pp. 1–52, 1975.

[12] D. A. Dawson, "Measure-valued Markov processes," in *École d'Été de Probabilités de Saint-Flour 1991,* Lecture Notes Math. 1541, pp. 1–260, Springer, Berlin, 1993.

[13] D. A. Dawson and P. Del Moral, "Large deviations for interacting processes in the strong topology," in *Statistical Modeling and Analysis for Complex Data Problem*, (P. Duchesne and B. Rmillard, eds.), pp. 179–209, Springer, 2005.

[14] D. A. Dawson, K. J. Hochberg, and V. Vinogradov, "On weak convergence of branching particle systems undergoing spatial motion," in *Stochastic Analysis: Random Fields and Measure-valued Processes*, pp. 65–79, *Israel Math. Conf. Proc. 10. Bar-Ilan Univ., Ramat Gan.*, 1996.

[15] P. Del Moral, "Non linear filtering: Interacting particle solution," *Markov Processes and Related Fields*, vol. 2, no. 4, pp. 555–580, 1996.

[16] P. Del Moral, "Measure valued processes and interacting particle systems. Application to nonlinear filtering problems," *Annals of Applied Probability*, vol. 8, no. 2, pp. 438–495, 1998.

[17] P. Del Moral, *Feynman-Kac Formulae: Genealogical and Interacting Particle Systems with Applications*. New York: Springer-Verlag, 2004.

[18] P. Del Moral and A. Doucet, "On a class of genealogical and interacting Metropolis Models," in *Séminaire de Probabilités* XXXVII, (J. Azéma, M. Emery, M. Ledoux, and M. Yor, eds.), pp. 415–446, Lecture Notes in Mathematics 1832, Springer-Verlag, Berlin, 2003.

[19] P. Del Moral and A. Doucet, "Particle motions in absorbing medium with hard and soft obstacles," *Stochastic Analysis and Applications*, vol. 22, no. 5, pp. 1175–1207, 2004.

[20] P. Del Moral and A. Doucet, *Sequential Monte Carlo & Genetic Particle Models. Theory and Practice*. Chapman & Hall, Green series, Statistics and Applied Probability, to appear, 2012.

[21] P. Del Moral, A. Doucet, and A. Jasra, "Sequential Monte Carlo samplers," *J. Royal Statist. Soc. B*, vol. 68, pp. 411–436, 2006.

[22] P. Del Moral, A. Doucet, and A. Jasra, "An adaptive sequential Monte Carlo method for approximate Bayesian computation," *Research report, Imperial College*, 2008.

[23] P. Del Moral, A. Doucet, and S. S. Singh, "A backward particle interpretation of Feynman-Kac formulae," *ESAIM M2AN*, vol. 44, no. 5, pp. 947–976, 2010.

[24] P. Del Moral, A. Doucet, and S. S. Singh, "Forward smoothing using sequential Monte Carlo," Technical Report 638. Cambridge University Engineering Department., 2010.

[25] P. Del Moral, A. Doucet, and S. S. Singh, "Uniform stability of a particle approximation of the optimal filter derivative," Technical Report CUED/F-INFENG/TR 668. Cambridge University Engineering Department., 2011.

[26] P. Del Moral and J. Garnier, "Genealogical particle analysis of rare events," *Annals of Applied Probability*, vol. 15, no. 4, pp. 2496–2534, 2005.
[27] P. Del Moral and J. Garnier, "Simulations of rare events in fiber optics by interacting particle systems," (preliminary version), *Optics Communications*, Elsevier, 2006.
[28] P. Del Moral and A. Guionnet, "Large deviations for interacting particle systems. Applications to nonlinear filtering problems," *Stochastic Processes Appl.*, vol. 78, pp. 69–95, 1998.
[29] P. Del Moral and A. Guionnet, "A central limit theorem for nonlinear filtering using interacting particle systems," *Ann. Appl. Probab.*, vol. 9, no. 2, pp. 275–297, 1999.
[30] P. Del Moral and A. Guionnet, "On the stability of measure valued processes with applications to filtering," *C.R. Acad. Sci. Paris*, vol. t. 329, no. Serie I, pp. 429–434, 1999.
[31] P. Del Moral and A. Guionnet, "On the stability of interacting processes with applications to filtering and genetic algorithms," *Ann. Inst. Henri Poincaré*, vol. 37, no. 2, pp. 155–194, 2001.
[32] P. Del Moral, S. Hu, and L. M. Wu, "Moderate deviations for mean field particle models," in preparation, 2011.
[33] P. Del Moral and J. Jacod, "The Monte-Carlo method for filtering with discrete-time observations: Central limit theorems," in *Numerical Methods and Stochastics (Toronto, ON, 1999)*, Vol. 34 of Fields Inst. Commun., pp. 29–53, American Mathematical Society, Providence, RI, 2002.
[34] P. Del Moral, J. Jacod, and P. Protter, "The Monte-Carlo method for filtering with discrete-time observations," *Probability Theory and Related Fields*, vol. 120, pp. 346–368, 2001.
[35] P. Del Moral, M. A. Kouritzin, and L. Miclo, "On a class of discrete generation interacting particle systems," *Electronic Journal of Probability*, vol. 16, pp. 1–26, 2001.
[36] P. Del Moral and M. Ledoux, "On the convergence and the applications of empirical processes for interacting particle systems and nonlinear filtering," *Journal of Theoretical Probability*, vol. 13, no. 1, pp. 225–257, 2000.
[37] P. Del Moral and P. Lezaud, "Branching and interacting particle interpretation of rare event probabilities," in *Stochastic Hybrid Systems: Theory and Safety Critical Applications*, (H. Blom and J. Lygeros, eds.), Springer-Verlag, Heidelberg, 2006.
[38] P. Del Moral and L. Miclo, "Branching and interacting particle systems approximations of Feynman-Kac formulae with applications to nonlinear filtering," in *Seminaire de Probabilités* XXXIV, Vol. 1729, (J. Azma, M. Emery, M. Ledoux, and M. Yor, eds.), pp. 1–145, Lecture Notes in Mathematics, Springer-Verlag Berlin, 2000.
[39] P. Del Moral and L. Miclo, "Particle approximations of Lyapunov exponents connected to Schrödinger operators and Feynman-Kac semigroups," *ESAIM Probability and Statistics*, vol. 7, pp. 169–207, 2003.
[40] P. Del Moral and E. Rio, "Concentration inequalities for mean field particle models," *Annals of Applied Probability*, vol. 21, no. 3, pp. 1017–1052, 2011.

[41] P. Del Moral and T. Zajic, "A note on Laplace-Varadhan's integral lemma," *Bernoulli*, vol. 9, no. 1, pp. 49–65, 2003.
[42] G. B. Di Masi, M. Pratelli, and W. G. Runggaldier, "An approximation for the nonlinear filtering problem with error bounds," *Stochastics*, vol. 14, no. 4, pp. 247–271, 1985.
[43] R. Douc, A. Guillin, and J. Najim, "Moderate deviations for particle filtering," *The Annals of Applied Probability*, vol. 15, no. 1B, pp. 587–614, 2005.
[44] A. Doucet, N. De Freitas, and N. Gordon, eds., *Sequential Monte Carlo Methods in Pratice,* Statistics for Engineering and Information Science. New York: Springer, 2001.
[45] E. B. Dynkin, "An introduction to branching measure-valued processes," in *CRM Monograph Series*, Vol. 6. Amer. Math. Soc., Providence, 1994.
[46] M. El Makrini, B. Jourdain, and T. Lelièvre, "Diffusion Monte Carlo method: Numerical analysis in a simple case," *ESAIM: M2AN*, vol. 41, no. 2, pp. 189–213, 2007.
[47] S. N. Ethier and T. G. Kurtz, *Markov Processes: Characterization and Convergence*. Wiley, New York, 1995.
[48] C. J. Everett and S. Ulam, *Proc. Nat. Acad. Sciences*, vol. 403, no. 33, 1948.
[49] C. J. Everett and S. Ulam, *USAEC, Los Alamos report, LADC-533, LADC-534*. declassified, 1948.
[50] W. Feller, "Diffusion processes in genetics," in *Proc. Second Berkeley Symp. Math. Statist. Prob.*, pp. 227–246, Berkeley: University of California Press, 1951.
[51] P. A. Ferrari, H. Kesten, S. Martínez, and P. Picco, "Existence of quasi-stationary distributions. A renewal dynamical approach," *Ann. Probab.*, vol. 23, pp. 501–521, 1995.
[52] Fr. Caron, P. Del Moral, A. Doucet, and M. Pace. On the Conditional Distributions of Spatial Point Processes *HAL-INRIA RR-7232*. To appear in *Advances in Applied Probability*, 2011.
[53] Fr. Caron, P. Del Moral, A. Doucet, and M. Pace, "Particle approximations of a class of branching distribution flows arising in multi-target tracking," *HAL-INRIA RR-7233 (2010) [29p], SIAM Journal on Control and Optimization*, vol. 49, no. 4, pp. 1766–1792, 2011.
[54] Fr. Caron, P. Del Moral, M. Pace, and B. N. Vo, "On the stability and the approximation of branching distribution flows, with applications to nonlinear multiple target filtering," *Stoch. Analysis and Applications*, vol. 29, no. 6, pp. 951–997, 2011.
[55] Fr. Cerou, A. Guyader, R. Rubinstein, and R. Vaismana, "On the use of smoothing to improve the performance of the splitting method," *Stochastic Models*, vol. 27, no. 4, pp. 629–650, 2011.
[56] E. Giné, "Lectures on some aspects of the bootstrap," in *Lecture Notes in Maths*, Vol. 1665, Springer, 1997.
[57] F. Gosselin, "Asymptotic behavior of absorbing Markov chains conditional on nonabsorption for applications in conservation biology," *Ann. Appl. Probab.*, vol. 11, pp. 261–284, 2001.
[58] C. Graham and S. Méléard, "Stochastic particle approximations for generalized Boltzmann models and convergence estimates," *Ann. Probab.*, vol. 25, no. 1, pp. 115–132, 1997.

[59] C. Graham and S. Méléard, "Probabilistic tools and Monte-Carlo approximations for some Boltzmann equations," in *CEMRACS 1999 (Orsay)*, Vol. 10 of ESAIM Proceedings, pp. 77–126, (electronic). Société de Mathématiques Appliquées et Industrielles, Paris, 1999.

[60] A. Guyader, N. Hengartner, and E. Matzner-Løber, "Simulation and estimation of extreme quantiles and extreme probabilities," *Applied Mathematics & Optimization*, vol. 64, no. 2, pp. 171–196, 2011.

[61] H. P. McKean, Jr, "A class of Markov processes associated with nonlinear parabolic equations," *Proc. Natl. Acad. Sci. U.S.A.*, vol. 56, pp. 1907–1911, 1966.

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

[63] H.P. McKean, Jr, "Propagation of chaos for a class of non-linear parabolic equations," in *Stochastic Differential Equations (Lecture Series in Differential Equations, Session 7, Catholic University, 1967)*, pp. 41–57, Air Force Office of Scientific Research, Arlington, VA, 1967.

[64] A. M. Johansen, P. Del Moral, and A. Doucet, "Sequential Monte Carlo samplers for rare events," in (preliminary version), *Proceedings of 6th International Workshop on Rare Event Simulation*, Bamberg, Germany, 2006.

[65] W. B. Johnson, G. Schechtman, and J. Zinn, "Best constants in moment inequalities for linear combinations of independent and exchangeable random variables," *Ann. Probab.*, vol. 13, no. 1, pp. 234–253, 1985.

[66] M. Kac, "On distributions of certain Wiener functionals," *Trans. American Math. Soc.*, vol. 61, no. 1, pp. 1–13, 1949.

[67] T. Klein and E. Rio, "Concentration around the mean for maxima of empirical processes," *Ann. Probab.*, vol. 33, no. 3, pp. 1060–1077, 2005.

[68] H. Korezlioglu and W. J. Runggaldier, "Filtering for nonlinear systems driven by nonwhite noises: An approximating scheme," *Stochastics and Stochastics Rep.*, vol. 44, no. 1–2, pp. 65–102, 1983.

[69] J. F. Le Gall, "Spatial branching processes, random snakes, and partial different equations," in *Lectures Given in the Nachdiplomsvorlesung, ETH Zürich*, Published in the series, Lectures in Mathematics, ETH Zürich, Birkhauser, 1999.

[70] M. Ledoux, "On Talagrand's deviation inequalities for product measures," *ESAIM Probability and Statistics*, vol. 1, pp. 63–87, 1996.

[71] M. Ledoux, "The concentration of measure phenomenon," AMS Monographs, Providence, 2001.

[72] M. Ledoux and M. Talagrand, *Probability in Banach Spaces*. Springer-Verlag, 1991.

[73] R. Mahler, "A theoretical foundation for the Stein-Winter Probability Hypothesis Density (PHD) multi-target tracking approach," in *Proc. MSS Nat'l Symp. on Sensor and Data Fusion*, Vol. 1, 2000.

[74] R. Mahler, "Multi-target Bayes filtering via first-order multi-target moments," *IEEE Trans. Aerospace & Electronic Systems*, vol. 39, no. 4, pp. 1152–1178, 2003.

[75] S. Méléard, "Asymptotic behaviour of some interacting particle systems; McKean-Vlasov and Boltzmann models," in *Probabilistic Models for Nonlinear*

Partial Differential Equations, Montecatini Terme, 1995, (D. Talay and L. Tubaro, eds.), Lecture Notes in Mathematics 1627. Springer-Verlag, Berlin, 1996.

[76] S. Méléard, "Convergence of the fluctuations for interacting diffusions with jumps associated with Boltzmann equations," *Stochastics Stochastics Rep.*, vol. 63, no. 3–4, pp. 195–225, 1998.

[77] S. Méléard, "Probabilistic interpretation and approximations of some Boltzmann equations," in *Stochastic models (Spanish) (Guanajuato, 1998)*, Vol. 14 of *Aportaciones Mat. Investig.*, pp. 1–64, Soc. Mat. Mexicana, México, 1998.

[78] S. Méléard, "Stochastic approximations of the solution of a full Boltzmann equation with small initial data," *ESAIM Probab. Stat.*, vol. 2, pp. 23–40, 1998.

[79] N. Metropolis and S. Ulam, "The Monte Carlo method," *Journal of the American Statistical Association*, vol. 44, no. 247, pp. 335–341, 1949.

[80] J. Picard, "Approximation of the non linear filtering problems and order of convergence," in *Lecture Notes in Control and Inf. Sc*, Vol. 61, 1984.

[81] D. Pollard, *Empirical Processes: Theory and Applications,* Vol. 2. NSF-CBMS Regional Conference Series in Probability and Statistics. Publication of the Institute of Mathematical Statistics. Hayward, CA, 1990.

[82] M. Pulvirenti, "Kinetic limits for stochastic particle systems," in *Probabilistic Models for Nonlinear Partial Differential Equations (Montecatini Terme, 1995)*, Vol. 1627 of Lecture Notes in Mathematics, pp. 96–126, Springer, Berlin, 1996.

[83] S. T. Rachev, *Probability Metrics and the Stability of Stochastic Models*. Wiley, New York, 1991.

[84] E. Rio, "Local invariance principles and their applications to density estimation," *Probability and related Fields*, vol. 98, pp. 21–45, 1994.

[85] M. Rousset, "On the control of an interacting particle approximation of Schrödinger ground states," *SIAM J. Math. Anal.*, vol. 38, no. 3, pp. 824–844, 2006.

[86] W. Runggaldier and L. Stettner, *Approximations of Discrete Time Partially Observed Control Problems*. Applied Mathematics Monographs CNR, Giardini Editori, Pisa, 1994.

[87] Z. Skolicki and K. De Jong, "The influence of migration sizes and intervals on island models," in *Proceedings of the Genetic and Evolutionary Computation Conference (GECCO-2005)*, ACM Press, 2005.

[88] D. Steinsaltz and S. N. Evans, "Markov mortality models: Implications of quasistationarity and varying initial distributions," *Theor. Pop. Biol.*, vol. 65, pp. 319–337, 2004.

[89] A. S. Sznitman, "Topics in propagation of chaos," in *Ecole d'Eté de Probabilités de Saint-Flour XIX-1989*, (P. L. Hennequin, ed.), Lecture Notes in Mathematics 1464. Springer-Verlag, Berlin, 1991.

[90] M. Talagrand, "Concentration of measure and isoperimetric inequalities in product spaces," *Inst. Hautes Études Sci. Publ. Math.*, vol. 81, pp. 73–205, 1995.

[91] M. Talagrand, "New concentration inequalities in product spaces," *Invent. Math.*, vol. 125, pp. 505–563, 1996.

[92] M. Talagrand, "A new look at independence," *Ann. Probab.*, vol. 24, pp. 1–34, 1996.

[93] A. N. Van der Vaart and J. A. Wellner, *Weak Convergence and Empirical Processes with Applications to Statistics*. New York: Springer Series in Statistics. Springer, 1996.

[94] J. Vanneste, "Estimating generalized Lyapunov exponents for products of random matrices," *Phys. Rev. E*, vol. 81, no. 3, p. 036701, 2010.

[95] J.-P. Vila and V. Rossi, "Nonlinear filtering in discrete time: A particle convolution approach, Biostatistic group of Montpellier," Technical Report 04-03, (available at http://vrossi.free.fr/recherche.html), 2004.

[96] D. Whitley, S. Rana, and R. B. Heckendorn, "The island model genetic algorithm: On separability, population size and convergence," *Journal of Computing and Information Technology*, vol. 7, no. 1, pp. 33–47, 1999.