INRIA

INSTITUT NATIONAL DE RECHERCHE EN INFORMATIQUE ET EN AUTOMATIQUE

A Mean Field Theory of Nonlinear Filtering

Pierre Del Moral — Frédéric Patras — Sylvain Rubenthaler

N° 6437

February 2008

Thème NUM

Rapport de recherche

ISSN 0249-6399 ISRN INRIA/RR--6437--FR+ENG

INSTITUT NATIONAL DE RECHERCHE EN INFORMATIQUE ET EN AUTOMATIQUE

INRIA

centre de recherche BORDEAUX - SUD OUEST

A Mean Field Theory of Nonlinear Filtering

Pierre Del Moral*, Frédéric Patras† , Sylvain Rubenthaler ‡

Thème NUM — Systèmes numériques
Équipes-Projets CQFD

Rapport de recherche n° 6437 — February 2008 — 36 pages

Abstract: We present a mean field particle theory for the numerical approximation of Feynman-Kac path integrals in the context of nonlinear filtering. We show that the conditional distribution of the signal paths given a series of noisy and partial observation data is approximated by the occupation measure of a genealogical tree model associated with mean field interacting particle model. The complete historical model converges to the McKean distribution of the paths of a nonlinear Markov chain dictated by the mean field interpretation model. We review the stability properties and the asymptotic analysis of these interacting processes, including fluctuation theorems and large deviation principles. We also present an original Laurent type and algebraic tree-based integral representations of particle block distributions. These sharp and non asymptotic propagations of chaos properties seem to be the first result of this type for mean field and interacting particle systems.

Key-words: Feynman-Kac measures, nonlinear filtering, interacting particle systems, historical and genealogical tree models, central limit theorems, Gaussian fields, propagations of chaos, trees and forests, combinatorial enumeration.

* Centre INRIA Bordeaux et Sud-Ouest & Institut de Mathématiques de Bordeaux , Université de Bordeaux I, 351 cours de la Libération 33405 Talence cedex, France, Pierre.Del-Moral@inria.fr.

† CNRS UMR 6621, Université de Nice, Laboratoire de Mathématiques J. Dieudonné, Parc Valrose, 06108 Nice Cedex 2, France.

‡ CNRS UMR 6621, Université de Nice, Laboratoire de Mathématiques J. Dieudonné, Parc Valrose, 06108 Nice Cedex 2, France.

Centre de recherche INRIA Bordeaux – Sud Ouest
Domaine Universitaire - 351, cours de la Libération 33405 Talence Cedex
Téléphone : +33 5 40 00 69 00

Une théorie champ moyen du filtrage non linéaire

Résumé : Nous exposons ici une théorie particulaire de type champ moyen pour la résolution numérique des intégrales de chemins de Feynman utilisées en filtrage non linéaire. Nous démontrons que les lois conditionnelles des trajectoires d'un signal bruité et partiellement observé peuvent être calculées à partir des mesures d'occupation d'arbres généalogiques associés à des systèmes de particules en interaction. Le processus historique caractérisant l'évolution ancestrale complète converge vers la mesure de McKean des trajectoires d'une chaîne de Markov non linéaire dictée par l'interprétation champ moyen du modèle de filtrage. Nous passons en revue les propriétés de stabilité et les résultats d'analyse asymptotique de ces processus en interaction, avec notamment des théorèmes de fluctuations et des principes de grandes déviations. Nous exposons aussi des développements faibles et non asymptotiques des distributions de blocs de particules en termes combinatoire de forêts et d'arbres de coalescences. Ces propriétés fines de propagations du chaos semblent être les premiers résultats de ce type pour des systèmes de particules en interaction de type champ moyen.

Mots-clés : Mesures de Feynman-Kac, filtrage non linéaire, systèmes de particules en interaction, processus historique et modèles d'arbres généalogiques, théorèmes de la limite centrale, champs gaussiens, propriétés de propagations du chaos, combinatoire d'arbres et de forêt.

1 Introduction

1.1 A mean field theory of nonlinear filtering

The filtering problem consists in computing the conditional distributions of a state signal given a series of observations. The signal/observation pair sequence $(X_n, Y_n)_{n\geq 0}$ is defined as a Markov chain which takes values in some product of measurable spaces $(E_n \times F_n)_{n\geq 0}$. We further assume that the initial distribution ν_0 and the Markov transitions P_n of the pair process (X_n, Y_n) have the form

$$\begin{aligned} \nu_0(d(x_0, y_0)) &= g_0(x_0, y_0)\ \eta_0(dx_0)\ q_0(dy_0) \\ P_n((x_{n-1}, y_{n-1}), d(x_n, y_n)) &= M_n(x_{n-1}, dx_n)\ g_n(x_n, y_n)\ q_n(dy_n) \end{aligned} \tag{1.1}$$

where g_n are strictly positive functions on $(E_n \times F_n)$ and q_n is a sequence of measures on F_n. The initial distribution η_0 of the signal X_n, the Markov transitions M_n and the likelihood functions g_n are assumed to be known.

The main advantage of this general and abstract set-up comes from the fact that it applies directly without further work to traditional real valued or multidimensional problems, as well as to smoothing and path estimation filtering models. For instance the signal

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

may represents the path from the origin up to time n of an auxiliary Markov chain X'_n taking values in some measurable state space $(E'_n, \mathcal{E}'_n)$. A version of the conditional distributions η_n of the signal states X_n given their noisy observations $Y_p = y_p$ up to time $p < n$ is expressed in terms of a flow of Feynman-Kac measures associated with the distribution of the paths of the signal and weighted by the collection of likelihood potential functions.

To better connect the mean field particle approach to existing alternatives methods, it is convenient at this point to make a couple of remarks.

Firstly, using this functional representation it is tempting to approximate this measures using a crude Monte Carlo method based on independent simulations of the paths of the signal weighted by products of the likelihood functions from the origin up to time n. This strategy gives satisfactory results when the signal paths are sufficiently stable using a simple weight regularization to avoid their degeneracy with respect to the time parameter. We refer to [12, 14, 33] for further details. Precise comparisons of the fluctuation variances associated with this Monte Carlo method and the mean feld particle models presented in the present article have been studied in [23].

Another commonly used strategy in Bayesian statistics and in stochastic engineering is the well-known Monte Carlo Markov chain method (abbreviate MCMC methods). The idea is to interpret the conditional distributions η_n as the invariant measure of a suitably chosen Markov chain. This technique as to main drawbacks. The first one is that we need to run the underlying chain for very long times. This burning period is difficult to estimate, and it is often too long to tackle filtering problems with high frequency observation sequences such as those arising in radar processing. The second difficulty is that the conditional distributions η_n vary with the time parameter and we need to chose at each time an appropriate MCMC algorithm.

In contrast to the two ideas discussed above, the mean field particle strategy presented in this article can be interpreted as a stochastic and adaptive grid approximation model. Loosely speaking, this particle technique consists in interpreting the weights likelihood as branching selection rates. No calibration of the convergence to equilibrium is needed, the conditional measures variations are automatically updated by the stochastic particle model. The first rigorous study in this field seems to be the article [15] published in 1996 on the applications of interacting particle methods to nonlinear estimation problems. This article provides the first convergence results for a series of heuristic genetic type schemes presented in the beginning of the 1990's in three independent chains of articles on nonlinear filtering [36, 35], [39], and [2, 40, 30, 32]. In the end of the 1990's, three other independent works [6, 7, 8] proposed another class of particle branching variants for solving continuous-time filtering problems.

For a more thorough discussion on the origins and the analysis these models, we refer the reader to the research monograph of the first author [11]. The latter is dedicated to Feynman-Kac and Boltzmann-Gibbs measures with their genealogical and interacting particle system interpretations, as well as their applications in physics, in biology and in engineering sciences. Besides their important application in filtering, Feynman-Kac type particle models also arise in the spectral analysis of Schrdinger type operators, rare events estimation, as well as in macromolecular and directed polymers simulations. In this connection, we also mention that the mean field theory presented here applies without further work to study the analysis of a variety of heuristic like models introduced in stochastic engineering since the beginning of the 1950's, including genetic type algorithms, quantum and sequential Monte Carlo methods, pruned enrichment molecular simulations, bootstrap and particle filters, and many others. For a rather thorough discussion on these rather well known application areas, the interested reader is also recommended to consult the book [34], and the references therein. These study can be completed with the more recent articles of the first author with A. Doucet and A. Jasra [17, 18] on sequential Monte Carlo techniques and their applications in Bayesian computation. We also mention that the continuous time version of the material presented in this review article can also be found in the series of articles of the first author with L. Miclo [26, 27, 28, 29]. For instance, the first reference [27] provides an original mean field particle interpretation of the robust nonlinear filtering equation.

This first key idea towards a mean field particle approach to nonlinear filtering is to recall that the flow of conditional measures η_n satisfies a dynamical system in distribution space, often called the nonlinear filtering equations. As in physics and more particularly in fluid mechanics, the second step is to interpret these equations as the evolution of he laws of a nonlinear Markov process. More formally, this preliminary stage simply consists in expressing the nonlinear filtering equations in terms of a transport model associated with a collection of Markov kernels K_{n+1,η_n} indexed by the time parameter n and the set of measures η_n on the space E_n; that is, we have that

$$\eta_{n+1} = \eta_n K_{n+1,\eta_n} \tag{1.2}$$

The mean field particle interpretation of this nonlinear measure valued model is an E_n^N-valued Markov chain $\xi_n^{(N)} = \left(\xi_n^{(N,i)}\right)_{1\leq i\leq N}$, with elementary transitions defined as

$$\mathbb{P}\left(\xi_{n+1}^{(N)} \in d(x^1,\ldots,x^N)|\xi_n^{(N)}\right) = \prod_{i=1}^{N} K_{n+1,\eta_n^N}(\xi_n^{(N,i)}, dx^i) \tag{1.3}$$

with

$$\eta_n^N := \frac{1}{N}\sum_{j=1}^{N}\delta_{\xi_n^{(N,j)}}$$

The initial system $\xi_0^{(N)}$ consists of N independent and identically distributed random variables with common law η_0. The state components of this Markov chain are called particles or sometimes walkers. The rationale behind this is that η_{n+1}^N is the empirical measure associated with N independent variables with distributions $K_{n+1,\eta_n^N}(\xi_n^{(N,i)}, .)$, so as soon as η_n^N is a good approximation of η_n then, in view of (1.3), η_{n+1}^N should be a good approximation of η_{n+1}. This strategy is not restricted to nonlinear filtering models and it applies under appropriate regularity conditions to any kind of measure valued model of the form (1.2). In our context, the evolution of the particles is dictated by the well known pair prediction-updating transitions associated with the optimal filter equations. The prediction transition is associated with a mutation transition of the whole population of particles. During this stage, the particles explore the state space independently of one another, according to the same probability transitions as the signal. The filter updating stage is associated with a branching type selection transition. During this stage, each particle evaluates the relative likelihood value of its location. The ones with poor value are killed, while the ones with high potential duplicate.

Using this branching particle interpretation, we notice that the ancestral lines of each particle form a genealogical tree evolving as above by tracking back in time the whole ancestor line of current individuals. In other words, the genealogical tree model associated with the branching process described above is the mean field particle interpretation of the nonlinear filtering equation but in path space. Using this simple observation, we readily prove that the genealogical tree occupation measure converges, as the population size tends to infinity, to the conditional distribution of the paths the signal $(X_p)_{0\leq p\leq n}$ given the observations delivered by the sensors up to time n. Another important mathematical object is the complete genealogical tree structure defined by the whole population model from the origin up to the current time horizon. The occupation measure of this ancestral tree keep tracks of the whole history of the particle and its convergence is technically more involved. We can prove that it converges, as the population size tends to infinity, to the McKean distribution of the paths of a nonhomogeneous Markov chain with transition probabilities K_{n+1,η_n}, and starting with the initial distribution η_0.

In the present article, we provide a synthetic review of the stability properties and the convergence analysis of these mean field interacting particle models going from the traditional law of large numbers to more sophisticated empirical process theorems, uniform estimates with respect to the time parameter, central limit theorems, Donsker type theorems, as well as large-deviation principles. We also analyze the increasing and strong propagation of chaos properties, including an original algebraic tree-based functional representations of particle block distributions, stripped of all analytical superstructure, and probabilistic irrelevancies. These Laurent type integral representations seems to be the first sharp and precise propagations of chaos estimates for this type of mean field particle models. We emphasized that most of the material presented in this review is taken from the book of the first author [11], and results from various collaboration with Donald Dawson, Jean Jacod, Michel Ledoux, Laurent Miclo, and Alice Guionnet. We refer to [11] for a detailed historical account with precise reference pointers. The sharp propagation of chaos expansions presented in the second part of this article are taken from the article [31]. The latter also discusses Hilbert series techniques for counting forests with prescribed numbers of vertices at each level, or with prescribed coalescence degrees, as well as new wreath products interpretations of vertex permutation groups. In this connection, we mention that forests and their combinatorics have also appeared recently in various fields such as in theoretical physics and gaussian matrix integral models [37], renormalization theory in high-energy physics or Runge-Kutta methods, two fields where the structure and complexity of perturbative expansions has required the development of new tools [1, 3].

The article is divided into four main parts, devoted respectively to the precise description of Feynman-Kac path integral models and their mean field particle interpretations, to the stability analysis of non linear semigroups, to the asymptotic analysis of mean field interacting processes, and to propagation of chaos properties.

Contents

1.2 Notation and conventions

For the convenience of the reader we have collected some of the main notation and conventions used in the article. We denote respectively by $\mathcal{M}(E)$, $\mathcal{P}(E)$, and $\mathcal{B}(E)$, the set of all finite signed measures on some measurable space $(E, \mathcal{E})$, the convex subset of all probability measures, and the Banach space of all bounded and measurable functions f on E, equipped with the uniform norm $\|f\| = \sup_{x\in E} |f(x)|$. We also denote $\mathrm{Osc}_1(E)$, the convex set of $\mathcal{E}$-measurable functions f with oscillations less than one; that is,

$$\mathrm{osc}(f) = \sup\{|f(x) - f(y)| \ ; \ x, y \in E\} \leq 1$$

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 $\mu \in \mathcal{M}(E)$, and we equip the Banach space $\mathcal{M}(E)$ with the total variation norm $\|\mu\|_{\mathrm{tv}} = \sup_{f\in \mathrm{Osc}_1(E)} |\mu(f)|$. We recall that a integral operator Q from a measurable space E_1 into an auxiliary measurable space E_2 into itself, is an operator $f \mapsto Q(f)$ from $\mathcal{B}(E_2)$ into $\mathcal{B}(E_1)$ such that the functions

$$Q(f)(x) = \int_{E_2} Q(x_1, dx_2)\ f(x_2)\ \in \mathbb{R}$$

are $\mathcal{E}_1$-measurable and bounded, for any $f \in \mathcal{B}(E_2)$. It also generates a dual operator $\mu \mapsto \mu Q$ from $\mathcal{M}(E_1)$ into $\mathcal{M}(E_2)$ defined by $(\mu Q)(f) := \mu(Q(f))$. We denote by $\beta(M) := \sup_{x,y\in E_1} \|M(x, .) - M(y, .)\|_{\mathrm{tv}} \in [0, 1]$ the Dobrushin coefficient associated with a Markov transition M from E_1 into E_2. For further use in various places of this article we recall that $\beta(M)$ is the norm of both the operator $\mu \mapsto \mu M$ on the set of measures with null mass or the operator $f \mapsto M(f)$ on the set of functions $f \in \mathrm{Osc}_1(E_1)$. Thus, we have that

$$\forall f \in \mathrm{Osc}_1(E_1) \qquad \mathrm{osc}(M(f)) \ \leq \ \beta(M)\ \mathrm{osc}(M(f))$$

For a proof of this rather well known inequality, we refer the reader to proposition 4.2.1 [11]. We also simplify the notation and we write for any $f, g \in \mathcal{B}(E_2)$ and $x \in E_1$

$$Q[(f - Qf)\ (g - Qg)](x)$$

instead of

$$Q[(f - Q(f)(x))\ (g - Q(g)(x))](x) = Q(fg)(x) - Q(f)(x)\ Q(g)(x)$$

The q-tensor power $Q^{\otimes q}$, with $q \geq 1$, represents the bounded integral operator on E_1^q into E_2^q, defined for any $F \in \mathcal{B}(E_2^q)$ by

$$Q^{\otimes q}(F)(x^1, \ldots, x^q) = \int_{E_2^q} \left[Q(x^1, dy^1)\ \ldots\ Q(x^q, dy^q)\right]\ F(y^1, \ldots, y^q)$$

For a pair of integral operators Q_1 from E_1 into E_2, and Q_2 from E_2 into E_3, we denote by Q_1Q_2 the composition integral operator from E_1 into E_3 defined for any $f \in \mathcal{B}(E_3)$ by $(Q_1Q_2)(f) := Q_1(Q_2(f))$.

With respect to the N-particle model introduced in (1.2), as soon as there is no possible confusion we simplify notation and suppress the index $(.)^{(N)}$ and write (ξ_n, ξ_n^i) instead of $(\xi_n^{(N)}, \xi_n^{(N,i)})$. To clarify the presentation, we often suppose that all the random processes including the signal and the particle processes are defined on some common probability space, and we denote by $\mathbb{E}(.)$ and $\mathbb{P}(.)$ the integral expectation and the probability measure on this common probability space. Last but not least, we fix a series of observations $Y_n = y_n$, with $n \geq 0$, and unless otherwise is stated all the results presented in this article are quenched in the sense that they depend of the observation sequence. The following classical conventions $(\sum_\emptyset, \prod_\emptyset) = (0, 1)$ are also used.

2 Feynman-Kac and mean field particle models

This section is concerned with Feynman-Kac formulations of the filtering equations with their mean field particle interpretations. The functional representations of both the one-step predictor and the optimal filter are discussed in the first subsection. In the second subsection, we discuss some advantages of these abstract models in the analysis of smoothing and path space filtering problems including continuous time signal-observation models. The mean field interacting particle systems and the corresponding genealogical tree based models are described in the final subsection.

2.1 Description of the models

We let G_n be the non homogeneous function on E_n defined for any $x_n \in E_n$ by

$$G_n(x_n) = g_n(x_n, y_n). \tag{2.1}$$

Note that G_n depends on the observation value y_n at time n. To simplify the presentation, and avoid unnecessary technical discussion, we shall suppose that the likelihood functions are chosen so that there exists a sequence of strictly positive constants $\epsilon_n(G) \in (0, 1]$ that may depend on the observation values y_n and such that for any $x_n, x'_n \in E_n$ we have that

$$G_n(x_n) \;\geq\; \epsilon_n(G)\; G_n(x'_n) \;> 0$$

The archetype of such nonlinear filtering model is the situation where Y_n is a real valued observation sequence described by a dynamical equation

$$Y_n = h_n(X_n) + V_n$$

where V_n represents a sequence of centered gaussian random variables with $\mathbb{E}(V_n^2) := \sigma_n > 0$ and $h_n \in \mathcal{B}(E_n)$. In this situation, (1.1) and the above condition are clearly met with the Lebesgue measure q_n on $\mathbb{R}$ and the likelihood potential function given by

$$G_n(x_n) = \exp\left(-\frac{1}{2\sigma_n^2}(y_n - h_n(x_n))^2\right)$$

Notice that in this case we have

$$\epsilon_n(G) \geq \exp\left(\sigma_n^{-2} \;\mathrm{osc}(h_n)\; (y_n + \|h_n\|)\right)$$

In this notation, versions of the one step predictor $\eta_n \in \mathcal{P}(E_n)$ and the optimal filter $\widehat{\eta}_n \in \mathcal{P}(E_n)$ given for any $f_n \in \mathcal{B}(E_n)$ by

$$\eta_n(f_n) := \mathbb{E}(f_n(X_n) \mid Y_p = y_p,\; 0 \leq p < n)$$

and

$$\widehat{\eta}_n(f_n) := \mathbb{E}(f_n(X_n) \mid Y_p = y_p,\ 0 \leq p \leq n)$$

have the functional representations

$$\eta_n(f_n) = \gamma_n(f_n)/\gamma_n(1) \quad \text{and} \quad \widehat{\eta}_n(f_n) = \widehat{\gamma}_n(f_n)/\widehat{\gamma}_n(1) \tag{2.2}$$

with the Feynman-Kac measures γ_n and $\widehat{\gamma}_n$ defined by the formulae

$$\gamma_n(f_n) = \mathbb{E}[f_n(X_n) \textstyle\prod_{0 \leq k < n} G_k(X_k)] \quad \text{and} \quad \widehat{\gamma}_n(f_n) = \gamma_n(f_n G_n) \tag{2.3}$$

A simple change of measure shows that the optimal filter can also be rewritten in the following prediction type form :

$$\widehat{\eta}_n(f_n) = \mathbb{E}[f_n(\widehat{X}_n) \prod_{0 \leq k < n} \widehat{G}_k(\widehat{X}_k)]/\mathbb{E}[\prod_{0 \leq k < n} \widehat{G}_k(\widehat{X}_k)] \tag{2.4}$$

In the above display, the signal $\widehat{X}_n$ is a Markov chain with initial distribution $\widehat{\eta}_0$ and the elementary transitions

$$\widehat{M}_n(x_{n-1}, dx_n) = \frac{M_n(x_{n-1}, dx_n)\ G_n(x_n)}{M_n(G_n)(x_{n-1})} \quad \text{and} \quad \widehat{G}_n(x_n) = M_{n+1}(G_{n+1})(x_n) \tag{2.5}$$

This simple observation shows that all the analysis on the one step predictor flow remains valid without further work to study the optimal filter. Another simple calculation shows that the unnormalized flow can be computed in terms of the normalized distributions. More precisely, we easily deduce the following multiplicative formulae :

$$\gamma_n(f) = \eta_n(f) \prod_{p=0}^{n-1} \eta_p(G_p) \tag{2.6}$$

2.2 Path space and related filtering models

Firstly, it is important to notice that the abstract Feynman-Kac formulation presented in section 2.1 is particularly useful for describing Markov motions on path spaces. For instance, X_n may represent the historical process

$$X_n = (X'_0, \ldots, X'_n) \in E_n = (E'_0 \times \ldots \times E'_n) \tag{2.7}$$

associated with an auxiliary Markov chain X'_n which takes values in some measurable state spaces E'_n. In this situation, we have that

$$\widehat{\eta}_n = \text{Law}((X'_0, \ldots, X'_n) \mid Y_p = y_p,\ 0 \leq p \leq n)$$

As we shall see, this apparently innocent observation is essential for modelling and analyzing genealogical evolution processes.

Another important state space enlargement allowed by our abstract formulation is the following. We let p_n be an increasing sequence of integers such that $p_0 = 0$, and we consider the pair signal observation model (X_n, Y_n) is given by

$$X_n = (X'_q)_{p_n \leq q < p_{n+1}} \in E_n := \prod_{p_n \leq q < p_{n+1}} E'_q$$

and

$$Y_n = (Y'_q)_{p_n \le q < p_{n+1}} \in F_n := \prod_{p_n \le q < p_{n+1}} F'_q$$

We further assume that the auxiliary pair signal observation model (X'_n, Y'_n) is defined as in (1.1) with a pair transition-likelihood function (M'_n, g'_n). In this situation, if we choose in (2.2) the multiplicative likelihood potential functions

$$G_n(X_n) = \prod_{p_n \le q < p_{n+1}} g'_q(x'_q, y'_q)$$

then we have that

$$\widehat{\eta}_n = \mathrm{Law}((X'_q)_{0 \le q \le p_{n+1}} \mid Y'_q = y'_q,\ 0 \le q < p_{n+1}) \tag{2.8}$$

As we shall see in the further development of section 2.3, one advantage of the mean field particle interpretation of this model comes from the fact that it only updates the sampled path predictions at the chosen times p_n. The "optimal" choice of the updating times depend on the filtering problem. For instance, in radar processing we current observation measures a noisy distance to the target. Thus, the speed and acceleration components are observable only after a series of three observations.

The traditional nonlinear filtering problem in continuous time is again defined in terms of a pair signal/observation Markov process (S_t, Y_t) taking values $\mathbb{R}^{d+d'}$. The signal S_t is given by a time homogeneous Markov process with right continuous and left limited paths taking values in some Polish space E, and the observation process is an $\mathbb{R}^{d'}$-valued process defined by

$$dY_t = h_t(S_t)\, dt + \sigma\, dV_t$$

where V_t is a d'-vector standard Wiener process independent of the signal, and h_t is a bounded measurable function from E into $\mathbb{R}^{d'}$. The Kallianpur-Striebel formula (see for instance [38]) states that there exists a reference probability measure $\mathbb{P}_0$ under which the signal and the observations are independent. In addition, for any measurable function f_t on the space $D([0,t], \mathbb{R}^d)$ of $\mathbb{R}^d$-valued càdlàg paths from 0 to t, we have that

$$\mathbb{E}(f_t((S_s)_{s \le t}) \mid \mathcal{Y}_t) = \frac{\mathbb{E}_0(f_t((S_s)_{s \le t})\ Z_t(S,Y) \mid \mathcal{Y}_t)}{\mathbb{E}_0(Z_t(S,Y) \mid \mathcal{Y}_t)} \tag{2.9}$$

where $\mathcal{Y}_t = \sigma(Y_s,\ s \le t)$ represents the sigma-field generated by the observation process and

$$\log Z_t(S,Y) = \int_0^t H_s^\star(S_s)\ dY_s - \int_0^t H_s^\star(S_s) H_s(S_s)\ ds$$

In the above definition $(.)^\star$ stands for the transposition operator. We let t_n, $n \ge 0$, be a given time mesh with $t_0 = 0$ and $t_n \le t_{n+1}$. Also let X'_n be the sequence of random variables defined by

$$X'_n = S_{[t_n, t_{n+1}]}$$

By construction, X'_n is a nonhomogeneous Markov chain taking values at each time n in the space $E'_n = D([t_n, t_{n+1}], \mathbb{R}^d)$. From previous considerations, the observation process Y_t can be regarded as a random environment. Given the observation path, we define the "random" potential functions G_n on

$$E_n = (E'_0 \times \ldots \times E'_n)$$

by setting for any $x_n = (x'_0, \ldots, x'_n)$, with $x'_p = (x'_p(s))_{t_p \leq s \leq t_{p+1}} \in E'_p$, and $0 \leq p \leq n$

$$\begin{array}{rcl} G_n(x_n) & = & G'_n(x'_n) \\ & := & \exp\left(\int_{t_n}^{t_{n+1}} H^\star_s(x'_n(s))\ dY_s - \int_{t_n}^{t_n} H^\star_s(x'_n(s)) H_s(x'_n(s))\ ds\right) \end{array}$$

By construction, we can check that the quenched Feynman-Kac path measures (2.2) associated with the pair (X_n, G_n) coincide with the Kallianpur-Striebel representation and by (2.9) we prove that

$$\widehat{\eta}_n = \text{Law}\left(S_{[t_0,t_1]}, \ldots, S_{[t_n,t_{n+1}]} \mid \mathcal{Y}_{t_n}\right) \tag{2.10}$$

Next, we suppose that the observations are only delivered by the sensors at some fixed times t_n, $n \geq 0$, with $t_n \leq t_{n+1}$. To analyze this situation, we first notice that

$$(Y_{t_{n+1}} - Y_{t_n}) = \int_{t_n}^{t_{n+1}} H_s(S_s)\, ds + \sigma\ (V_{t_{n+1}} - V_{t_n})$$

and $\sigma\ (V_{t_{n+1}} - V_{t_n})$ are independent and random variables with Gaussian density q_n. Arguing as before and using the same notation as there, we introduce the "random" potential functions G_n defined by

$$G_n(x_n) = q_n\left(Y_{t_{n+1}} - Y_{t_n} - \int_{t_n}^{t_{n+1}} H_s(x'_n(s))\, ds\right)$$

By construction, the quenched Feynman-Kac path measures (2.2) associated with the pair (X_n, G_n) now have the following interpretation

$$\widehat{\eta}_n = \text{Law}\left(S_{[t_0,t_1]}, \ldots, S_{[t_n,t_{n+1}]} \mid Y_{t_1}, \ldots, Y_{t_n}\right) \tag{2.11}$$

We end this section with a couple of remarks.

To compute the Feynman-Kac distributions (2.10) and (2.11), it is tempting to use directly the mean field particle approximation model introduced in (1.3). Unfortunately, in practice the signal semigroup and the above integrals are generally not known exactly and another level of approximation is therefore needed. As shown in [22], the use of Euler approximation schemes introduces a deterministic bias in the fluctuation of the particle measures, but the resulting approximation models can be analyzed using the same perturbation analysis as the one we shall describe in section 4.1 (see also [5, 11], for a more thorough analysis of the discrete time schemes and genetic type particle approximations).

To solve the continuous time problem, another strategy is to use a fully continuous time particle approximation model of the robust equations. Loosely speaking, the geometric acceptance rates of the discrete generation particle models are replaced by exponential clocks with an appropriate stochastic intensity dictated by the log-likelihood potential functions. For an introduction to interacting particle interpretation of continuous time Feynman-Kac models we recommend the review article on genetic type models [25] as well as the series of articles [27, 28, 29].

2.3 McKean interpretations

By the Markov property and the multiplicative structure of (2.2), it is easily checked that the flow of measures $(\eta_n)_{n\geq 0}$ satisfies the following equation

$$\eta_{n+1} = \Phi_n(\eta_{n-1}) := \Psi_n(\eta_n)M_{n+1} \tag{2.12}$$

with the updating Bayes or the Boltzmann-Gibbs transformation $\Psi_n : \mathcal{P}(E_n) \to \mathcal{P}(E_n)$ defined by

$$\Psi_n(\eta_n)(dx_n) := \frac{1}{\eta_n(G_n)}\ G_n(x_n)\ \eta_n(dx_n)\,(=\widehat{\eta}_n(dx_n))\,.$$

The mean field particle approximation of the flow (2.12) depends on the choice of the McKean interpretation model. As we mentioned in the introduction, these stochastic models amount of choosing a suitably defined Markov chain $\overline{X}_n$ with the prescribed evolution (2.12) of the laws of its states. More formally, these probabilistic interpretations consist of a chosen collection of Markov transitions K_{n+1,η_n}, indexed by the time parameter n and the set of probability measures $\eta \in \mathcal{P}(E_n)$, and satisfying the compatibility condition

$$\Phi_{n+1}(\eta) = \eta K_{n+1,\eta} \tag{2.13}$$

The choice of these collections is not unique. We can choose, for instance the composition transition operator

$$K_{n+1,\eta_n} = S_{n,\eta_n} M_{n+1}$$

with the updating Markov transition S_{n,η_n} from E_n into itself defined by the following formula

$$S_{n,\eta_n}(x_n, dy_n) = \epsilon_n(\eta_n)\ G_n(x_n)\ \delta_{x_n}(dy_n) + (1 - \epsilon_n(\eta_n)\ G_n(x_n))\ \Psi_n(\eta_n)(dy_n) \tag{2.14}$$

In the above display, $\epsilon_n(\eta_n)$ represents *any* possibly null constant that may depends on the current distribution η_n, and such that $\epsilon_n(\eta_n)\|G_n\| \leq 1$. For instance, we can choose $1/\epsilon_n(\eta_n) = \eta_n - \text{ess} - \sup G_n$. The corresponding nonlinear transport equation

$$\eta_{n+1} = \eta_n K_{n+1,\eta_n}$$

can be interpreted as the evolution of the laws η_n of the states of a Markov chain $\overline{X}_n$ whose elementary transitions K_{n+1,η_n} depend on the law of the current state; that is, we have

$$\mathbb{P}(\overline{X}_{n+1} \in dx_{n+1} \mid \overline{X}_n = x_n) = K_{n+1,\eta_n}(x_n, dx_{n+1}) \quad \text{with} \quad \mathbb{P} \circ \overline{X}_n^{-1} = \eta_n \tag{2.15}$$

The law $\mathbb{K}_n$ of the random path $(\overline{X}_p)_{0\leq p\leq n}$ is called the McKean measure associated with the Markov transitions $(K_{n,\eta})_{n\geq 0,\eta\in\mathcal{P}(E_n)}$ and the initial distribution η_0. This measure on path space is explicitly defined by the following formula

$$\mathbb{K}_n(d(x_0,\ldots,x_n)) = \eta_0(dx_0)\ K_{1,\eta_0}(x_0,dx_1)\ \ldots\ K_{n,\eta_{n-1}}(x_{n-1},dx_n)$$

2.4 Mean field particle and genealogical tree based models

The N-particle model associated with a given collection of Markov transitions satisfying the compatibility condition (2.13) is the Markov chain introduced in (1.2). By the definition of the updating transitions (2.14), it appears that the mean field interacting particle model (1.2) is the combination of simple selection/mutation genetic transitions. The selection stage consists of N randomly evolving path-particles $\xi^i_{n-1} \rightsquigarrow \widehat{\xi}^i_{n-1}$ according to the update transition $S_{n,\eta^N_{n-1}}(\xi^i_{n-1}, .)$. In other words, with probability $\epsilon_{n-1}(\eta^N_{n-1})G_{n-1}(\xi^i_{n-1})$, we set $\widehat{\xi}^i_{n-1} = \xi^i_{n-1}$; otherwise, the particle jumps to a new location, randomly drawn from the discrete distribution $\Psi_{n-1}(\eta^N_{n-1})$. During the mutation stage, each of the selected particles $\widehat{\xi}^i_{n-1} \rightsquigarrow \xi^i_n$ evolves according to the Markov transition M_n.

For any sufficiently regular transitions $K_{n,\eta_{n-1}}$ satisfying the compatibility condition (2.13), we can prove that for any time horizon n

$$\mathbb{K}^N_n := \frac{1}{N}\sum_{i=1}^N \delta_{(\xi^i_0,\ldots,\xi^i_n)} \longrightarrow_{N\to\infty} \mathbb{K}_n \quad \text{and} \quad \eta^N_n := \frac{1}{N}\sum_{i=1}^N \delta_{\xi^i_n} \longrightarrow_{N\to\infty} \eta_n$$

Mimicking formula (2.6), we also construct *an unbiased* estimate for the unnormalized model and we have that

$$\gamma^N_n(.) := \eta^N_n(.) \prod_{p=0}^{n-1} \eta^N_p(G_p) \longrightarrow_{N\to\infty} \gamma_n(.) = \eta_n(.) \prod_{p=0}^{n-1} \eta_p(G_p)$$

The convergence above can be understood in various ways. A variety of estimates going from the traditional $\mathbb{L}_p$-mean error bounds and exponential inequalities to fluctuation and large deviation theorems are provided in section 4 dedicated to the asymptotic behavior of these particle measures.

If we interpret the selection transition as a birth and death process, then arises the important notion of the ancestral line of a current individual. More precisely, when a particle $\widehat{\xi}^i_{n-1} \longrightarrow \xi^i_n$ evolves to a new location ξ^i_n, we can interpret $\widehat{\xi}^i_{n-1}$ as the parent of ξ^i_n. Looking backwards in time and recalling that the particle $\widehat{\xi}^i_{n-1}$ has selected a site ξ^j_{n-1} in the configuration at time $(n-1)$, we can interpret this site ξ^j_{n-1} as the parent of $\widehat{\xi}^i_{n-1}$ and therefore as the ancestor $\xi^i_{n-1,n}$ at level $(n-1)$ of ξ^i_n. Running back in time we trace mentally the whole ancestral line of each current individual :

$$\xi^i_{0,n} \longleftarrow \xi^i_{1,n} \longleftarrow \ldots \longleftarrow \xi^i_{n-1,n} \longleftarrow \xi^i_{n,n} = \xi^i_n$$

If we consider the historical process formulation (2.7) of a given signal model X'_n, then the mean field particle model associated with the Feynman-Kac measures on path spaces consists of N path particles evolving according to the same selection/mutation transitions. It is rather clear that the resulting path particle model can also be interpreted as the evolution of a genealogical tree model. This shows that the occupation measures of the corresponding N-genealogical tree model

$$\eta^N_n = \frac{1}{N}\sum_{i=1}^N \delta_{(\xi^i_{0,n},\xi^i_{1,n},\ldots,\xi^i_{n,n})}$$

converge as $N \to \infty$ to the Feynman-Kac path measures η_n defined as in (2.2) with the pair of mathematical objects (X_n, G_n) given by

$$X_n = (X'_0, \ldots, X'_n) \quad \text{and} \quad G_n(X_n) = g_n(X'_n, y_n)$$

From previous considerations, we already mention that the mathematical techniques developed to study the convergence of the particle measures η_N^N apply directly without further work to analyze the asymptotic behavior of this class of genealogical tree particle models. The occupation measures $\mathbb{K}_n^N$ of the complete genealogical tree $(\xi_0^i, \ldots, \xi_n^i)_{1\leq i\leq N}$ keep track of the whole descendant history of the initial population individuals ξ_0. The asymptotic analysis of these particle measures strongly depends on the choice of the McKean interpretation model and it requires more attention.

Last but not least, the mean filed particle methodology we have developed applies directly to solve the path-space filtering models we have presented in section 2.2. We leave the interested reader to write down the interacting particle systems in each situation. We also mention that the prediction formulation of the optimal filter described in (2.4) leads to a genetic type particle scheme with a pair mutation-selection transition associated with the pair transition-potential function $(\widehat{M}_n, \widehat{G}_n)$ given in (2.5). This strategy is based on observation depended explorations and the resulting stochastic grid is often more accurate than the one based on free mutations. Nevertheless, expect in some particular situations the sampling of a transition $\widehat{\xi}_{n-1}^i \rightsquigarrow \xi_n^i$ according to the distribution $\widehat{M}_n(\widehat{\xi}_{n-1}^i, dx_n)$ is often difficult. One idea is to introduce an auxiliary empirical particle approximation

$$M_n^{N'}(\widehat{\xi}_{n-1}^i, dx_n) := \frac{1}{N'} \sum_{j=1}^{N'} \delta_{\zeta_n^{i,j}}$$

of the Markov transition $M_n(\widehat{\xi}_{n-1}^i, dx_n)$, and based on sampling N'-independent random transitions $\widehat{\xi}_{n-1}^i \rightsquigarrow \zeta_n^{i,j}$, $1 \leq k \leq N'$ with common distribution given by $M_n(\widehat{\xi}_{n-1}^i, dx_n)$. The second step is to replace the pair transition-potential function $(\widehat{M}_n, \widehat{G}_n)$ given in (2.5) by their N'-particle approximations $(\widehat{M}_n^{N'}, \widehat{G}_n^{N'})$ defined by

$$\widehat{M}_n^{N'}(\widehat{\xi}_{n-1}^i, dx_n) := \frac{M_n^{N'}(\widehat{\xi}_{n-1}^i, dx_n)\ G_n(x_n)}{M_n^{N'}(G_n)(\widehat{\xi}_{n-1}^i)}$$

and

$$\widehat{G}_{n-1}^{N'}(\widehat{\xi}_{n-1}^i) := M_n^{N'}(G_n)(\widehat{\xi}_{n-1}^i)$$

These particle exploration models can be combined without further work with the path space enlargement techniques presented in (2.8). Roughly speaking, the corresponding path particle model is based on exploring the state space using the conditional transitions of signal sequences $(X_q)_{p_{n-1}\leq q<p_n} \rightsquigarrow (X_q)_{p_n\leq q<p_{n+1}}$ based on $(p_n - p_{n-1})$ observations data $(Y_q)_{p_n\leq q<p_{n+1}}$. The auxiliary N'-particle approximation described above only provides a practical way of sampling these explorations on an auxiliarly pool of N' sampled sequences.

These conditional exploration strategies have been originally developed in [13, 21]. We also refer the interested reader to chapter 11 of the book [11] for further stochastic particle recipes including a variety of branching type selection variants.

3 Stability Analysis

This section is concerned with the regularity and the stability properties of the evolution semigroup of the measure valued process introduced in (2.12). This analysis is motivated by tow essential problems. From the signal processing perspective, the first one is to ensure that the underlying filtering problem is well posed, in the sense that it corrects automatically any erroneous initial data. This stability property is fundamental in most of the filtering problems encountered in practical situations where the initial distribution of the signal is unknown. The second motivation is well-known in the numerical analysis of dynamical systems. Indeed, it is generally useless to approximate an unstable and chaotic dynamical system that propagates any local perturbation. In our context, the mean field particle model can be interpreted as a stochastic perturbation of the nonlinear dynamical system (2.12). In this situation, the regularity properties of the evolution semigroup ensure that these local perturbations will not propagate. These ideas will be made clear in the forthcoming development of section 4.

3.1 Feynman-Kac evolution semigroups

We let $Q_{p,n}$, with $0 \leq p \leq n$, be the Feynman-Kac semi-group associated with the flow of unnormalized Feynman-Kac measures $\gamma_n = \gamma_p Q_{p,n}$ defined in (2.3). For $p = n$, we use the convention that $Q_{n,n} = Id$, the identity operator. Using the Markov property, it is not difficult to check that $Q_{p,n}$ has the following functional representation

$$Q_{p,n}(f_n)(x_p) = \mathbb{E}\left[f_n(X_n)\ \prod_{p\leq k<n} G_k(X_k) \mid X_p = x_p\right] \tag{3.1}$$

for any test function $f_n \in \mathcal{B}(E_n)$, and any state $x_p \in E_p$. We denote by $\Phi_{p,n}$, $0 \leq p \leq n$, the nonlinear semigroup associated with the normalized measures η_n introduced in (2.2)

$$\Phi_{p,n} = \Phi_n \circ \Phi_{n-1} \circ \ldots \circ \Phi_{p+1} \tag{3.2}$$

As usual we use the convention $\Phi_{n,n} = Id$, for $p = n$. It is important to observe that this semigroup is alternatively defined by the formulae

$$\Phi_{p,n}(\eta_p)(f_n) = \frac{\eta_p(Q_{p,n}(f_n))}{\eta_p(Q_{p,n}(1))} = \frac{\eta_p(G_{p,n}\ P_{p,n}(f_n))}{\eta_p(G_{p,n})}$$

with the pair potential and Markov transition $(G_{p,n}, P_{p,n})$ defined by

$$G_{p,n} = Q_{p,n}(1) \quad \text{and} \quad P_{p,n}(f_n) = Q_{p,n}(f_n)/Q_{p,n}(1)$$

The next two parameters

$$r_{p,n} = \sup_{x_p,y_p\in E_p} (G_{p,n}(x_p)/G_{p,n}(y_p))$$

and

$$\beta(P_{p,n}) = \sup_{x_p,y_p\in E_p} \|P_{p,n}(x_p,.) - P_{p,n}(y_p,.)\|_{\rm tv} \tag{3.3}$$

measure respectively the relative oscillations of the potential functions $G_{p,n}$ and the contraction properties of the Markov transition $P_{p,n}$. Various asymptotic estimates on particle models derived in the forthcoming sections will be expressed in terms of these parameters. For instance and for further use in several places in this article, we have the following Lipschitz regularity property.

Proposition 3.1 *For any $f_n \in Osc_1(E_n)$ we have*

$$|[\Phi_{p,n}(\eta_p) - \Phi_{p,n}(\mu_p)]\,(f_n)| \leq \; 2\; r_{p,n}\; \beta(P_{p,n}) \quad \left|[\eta_p - \mu_p]\,\overline{P}^{\mu_p}_{p,n}(f_n)\right| \tag{3.4}$$

for some function $\overline{P}^{\mu_p}_{p,n}(f_n) \in Osc_1(E_p)$ that doesn't depends on the measure η_p.

Proof:
We check this inequality using the key decomposition

$$[\Phi_{p,n}(\eta_p) - \Phi_{p,n}(\mu_p)]\,(f_n) = \frac{1}{\eta_p(G^{\mu_p}_{p,n})}\; [(\eta_p - \mu_p)R^{\mu_p}_{p,n}(f_n)]$$

with the function $G^{\mu_p}_{p,n} := G_{p,n}/\mu_p(G_{p,n})$ and the integral operator $R^{\mu_p}_{p,n}$ from $\mathcal{B}(E_n)$ into $\mathcal{B}(E_p)$ defined below :

$$R^{\mu_p}_{p,n}(f_n)(x_p) \quad := \quad G^{\mu_p}_{p,n}(x_p) \;\times\; P_{p,n}\,[f_n - \Phi_{p,n}(\mu_p)(f_n)]\,(x_p)$$

Recalling that $\mathrm{osc}(P_{p,n}(f_n)) \leq \beta(P_{p,n})\ \mathrm{osc}(f_n)$, one readily check that (3.4) is satisfied with the integral operator $\overline{P}^{\mu_p}_{p,n}$ from $\mathcal{B}(E_n)$ into $\mathcal{B}(E_p)$ defined

$$\overline{P}^{\mu_p}_{p,n}(f_n)(x_p)$$

$$:= \tfrac{1}{2}\;\tfrac{1}{r_{p,n}}\,\tfrac{G_{p,n}(x_p)}{\inf G_{p,n}} \int \frac{1}{\beta(P_{p,n})}\,[P_{p,n}(f_n)(x_p) - P_{p,n}(f_n)(y_p)]\;\; G^{\mu_p}_{p,n}(y_p)\; \mu_p(dy_p)$$

■

3.2 Contraction properties

In this section, we present an abstract class of H-entropies like criteria. We also provide a brief introduction a brief introduction to the Lipschitz contraction properties of Markov integral operators. For a more detailed discussion on this subject, we refer the reader to chapter 4 of the book [11], and to a joint work of the first author with M. Ledoux and L. Miclo [24]. The contraction functional inequalities described in this section are extended to the nonlinear Feynman-Kac semigroup introduced in (3.2).

Let $h : \mathbb{R}^2_+ \to \mathbb{R} \cup \{\infty\}$ be a convex function satisfying for any $a, x, y \in \mathbb{R}_+$ such that $h(ax, ay) = ah(x, y)$, and $h(1, 1) = 0$. We associate with this homogeneous function the H-relative entropy on $\mathcal{M}_+(E)$ defined symbolically as

$$H(\mu, \nu) = \int\; h\,(d\mu, d\nu)$$

By homogeneity arguments, the above entropy is better defined in terms of any measure $\lambda \in \mathcal{M}(E)$ dominating μ and ν by the formula

$$H(\mu, \nu) = \int\; h\left(\frac{d\mu}{d\lambda}, \frac{d\nu}{d\lambda}\right)\; d\lambda$$

To illustrate this rather abstract definition, we provide hereafter a collection of classical h-relative entropies arising in the literature. Before to proceed, we first we come back to the definition of h-entropy, and we denote by $h' : \mathbb{R}_+ \to \mathbb{R} \sqcup \{+\infty\}$ the convex function given for any $x \in \mathbb{R}_+$ by $h'(x) = h(x, 1)$. By homogeneity arguments, we note that h is almost equivalent to h'. More precisely, only the specification of the value $h(1, 0)$ is missing. In most applications, the natural convention is $h(1, 0) = \infty$. The next lemma connects the h-relative entropy with the h'-divergence in the sense of Csiszar [10].

Lemma 3.2 ([11]) *Assume that $h(1, 0) = +\infty$. Then, for any μ and $\nu \in \mathcal{M}_+(E)$, we have*

$$H(\mu, \nu) = \int h'\left(\frac{d\mu}{d\nu}\right) d\nu \qquad \text{if } \mu \ll \nu, \text{ and } H(\mu, \nu) = \infty \text{ otherwise.} \tag{3.5}$$

As we promised above, let us present some traditional H-entropies commonly used probability theory. If we take $h'(t) = |t - 1|^p$, $p \geq 1$, we find the $\mathbb{L}_p$-norm given for any $\mu, \nu \in \mathcal{P}(E)$ by $H(\mu, \nu) = \|1 - d\mu/d\nu\|_{p,\nu}^p$ if $\mu \ll \nu$, and ∞ otherwise.The case $h'(t) = t \log(t)$ corresponds to the Boltzmann entropy or Shannon-Kullback information given by the formula $H(\mu, \nu) = \int \ln\left(\frac{d\mu}{d\nu}\right) d\mu$, if $\mu \ll \nu$ and ∞ otherwise. The Havrda-Charvat entropy of order $p > 1$ corresponds to the choice $h'(t) = \frac{1}{p-1}(t^p - 1)$; that is we have for any $\mu \ll \nu$, $H(\mu, \nu) = \frac{1}{p-1}\left[\int \left(\frac{d\mu}{d\nu}\right)^p d\nu - 1\right]$. The Hellinger and Kakutani-Hellinger integrals of order $\alpha \in (0, 1)$ correspond to the choice $h'(t) = t - t^\alpha$, and with some obvious abusive notation we have $H(\mu, \nu) = 1 - \int (d\mu)^\alpha (d\nu)^{1-\alpha}$. In the special case $\alpha = 1/2$, this relative entropy coincides with the Kakutani-Hellinger distance; and finally, the case $h'(t) = |t - 1|/2$ corresponds to the total variation distance. In the study of regularity properties of $\Phi_{p,n}$, the following notion will play a major role.

Definition 3.3 *Let $(E, \mathcal{E})$ and $(F, \mathcal{F})$ be a pair of measurable spaces. We consider an h-relative entropy criterion H on the sets $\mathcal{P}(E)$ and $\mathcal{P}(F)$. The contraction or Lipschitz coefficient $\beta_H(\Phi) \in \mathbb{R}_+ \cup \{\infty\}$ of a mapping $\Phi : \mathcal{P}(E) \to \mathcal{P}(F)$ with respect to H is the best constant such that for any pair of measures $\mu, \nu \in \mathcal{P}(E)$ we have*

$$H(\Phi(\mu), \Phi(\nu)) \;\leq\; \beta_H(\Phi)\; H(\mu, \nu)$$

When H represents the total variation distance, we simplify notation and sometimes we write $\beta(\Phi)$ instead of $\beta_H(\Phi)$.

When H is the total variation distance, the parameter $\beta(\Phi)$ coincides with the traditional notion of a Lipschitz constant of a mapping between two metric spaces. In [24], we prove the following regularity property. This functional inequality will be pivotal in the contraction analysis of Feynman-Kac semigroups developed in section 3.3.

Theorem 3.4 *For any pair of probability measures μ and $\nu \in \mathcal{P}(E)$ and for any Markov kernel M from E into another measurable space F, we have the contraction estimate*

$$H(\mu M, \nu M) \leq \beta(M)\; H(\mu, \nu) \tag{3.6}$$

3.3 Functional inequalities

The main objective of this section will be to estimate the contraction coefficients $\beta_H(\Phi_{p,n})$ of the nonlinear Feynman-Kac transformations $\Phi_{p,n}$ in terms of the Dobrushin coefficient $\beta(P_{p,n})$ and the relative oscillations of the potential functions $G_{p,n}$.

To describe precisely these functional inequalities precisely, it is convenient to introduce some additional notation. When H is the h'-divergence associated with a differentiable $h' \in C^1(\mathbb{R}_+)$, we denote by Δh the function on $\mathbb{R}_+^2$ defined by

$$\Delta h(t,s) = h'(t) - h'(s) - \partial h'(s)\ (t-s) \quad (\geq 0)$$

where $\partial h'(s)$ stands for the derivative of h' at $s \in \mathbb{R}_+$. We further assume that the following growth condition is satisfied

$$\forall (r,s,t) \in \mathbb{R}_+^3 \text{ we have } \quad \Delta h(rt,s) \ \leq \ a(r)\ \Delta h(t, \theta(r,s)) \tag{3.7}$$

for some nondecreasing function a on $\mathbb{R}_+$ and some mapping θ on $\mathbb{R}_+^2$ such that $\theta(r, \mathbb{R}_+) = \mathbb{R}_+$, for any $r \in \mathbb{R}_+$.

For instance, the Boltzmann entropy corresponds to the situation where $h'(t) = t \log t$. In this case, (3.7) is met with $a(r) = r$ and $\theta(r,s) = s/r$. For the Havrda-Charvat entropy of order $\alpha > 1$, we have $h'(t) = (t^\alpha - 1)/(\alpha - 1)$, and the growth condition (3.7) is now met with $a(r) = r^\alpha$ and $\theta(r,s) = s/r$. The Hellinger integrals of order $\alpha \in (0,1)$ correspond to the choice $h'(t) = t - t^\alpha$, and the growth condition is met with the same parameters. The $\mathbb{L}_2$-relative entropy corresponds to the case $h'(t) = (t-1)^2$, and we find that (3.7) is met with $a(r) = r^2$, and again $\theta(r,s) = s/r$.

The following theorem is s slightly weaker version of a theorem proved in [11].

Theorem 3.5 *For any $0 \leq p \leq n$ and any pair of measures $\mu_p, \nu_p \in \mathcal{P}(E_p)$, we have the Lipschitz contraction inequality*

$$\|\Phi_{p,n}(\mu_p) - \Phi_{p,n}(\nu_p)\|_{\rm tv} \leq \ r_{p,n}\ \beta(P_{p,n})\ \|\mu_p - \nu_p\|_{\rm tv} \tag{3.8}$$

In addition, if we set $a_H(r) := ra(r)$ then we have

$$\beta_H(\Phi_{p,n}) \leq \ a_H(r_{p,n})\ \beta(P_{p,n})$$

and

$$\beta(P_{p,n}) = \sup_{\mu_p,\nu_p \in \mathcal{P}(E_p)} \|\Phi_{p,n}(\mu_p) - \Phi_{p,n}(\nu_p)\|_{\rm tv} \tag{3.9}$$

Our next objective is to estimate the the contraction coefficient $\beta_H(\Phi_{p,n})$ in terms of the mixing type properties of the semigroup

$$M_{p,n}(x_p, dx_n) := M_{p+1}M_{p+2}\ldots M_n(x_p, dx_n)$$

associated with the Markov operators M_n. We introduce the following regularity condition.

$(M)_m$ *There exists an integer $m \geq 1$ and a sequence $(\epsilon_p(M))_{p\geq 0} \in (0,1)^{\mathbb{N}}$ such that*

$$\forall p \geq 0 \quad \forall (x_p, x'_p) \in E_p^2 \qquad M_{p,p+m}(x_p, \centerdot) \geq \epsilon_p(M) \quad M_{p,p+m}(x'_p, \centerdot)$$

It is well known that the above condition is satisfied for any aperiodic and irreducible Markov chains on finite spaces. Loosely speaking, for non compact spaces this condition is related to the tails of the transition distributions on the boundaries of the state space. For instance, let us suppose that $E_n = \mathbb{R}$ and M_n is the bi-Laplace transition given by

$$M_n(x, dy) = \frac{c(n)}{2}\ e^{-c(n)\,|y-A_n(x)|}\ dy$$

for some $c(n) > 0$ and some drift function A_n with bounded oscillations $\text{osc}(A_n) < \infty$. In this case, it is readily checked that condition $(M)_m$ holds true for $m = 1$ with the parameter $\epsilon_{n-1}(M) = \exp\left(-c(n)\ \text{osc}(A_n)\right)$.

Under the mixing type condition $(M)_m$ we have for any $n \geq m \geq 1$, and $p \geq 1$

$$r_{p,p+n} \leq \epsilon_p(M)^{-1} \prod_{0 \leq k < m} \epsilon_{p+k}(G)^{-1}$$

and

$$\beta(P_{p,p+n}) \leq \prod_{k=0}^{\lfloor n/m \rfloor - 1} \left(1 - \epsilon^{(m)}_{p+km}(G, M)\right)$$

with the sequence of parameters $\epsilon^{(m)}_p(G, M)$ given by

$$\epsilon^{(m)}_p(G, M) = \epsilon^2_p(M) \prod_{p+1 \leq k < p+m} \epsilon_k(G)$$

Several contraction inequalities can be deduced from these estimates (see for instance chapter 4 of the book [11]). To give a flavor of these results, we further assume that $(M)_m$ is satisfied with $m = 1$, and we have $\epsilon(M) = \inf_n \epsilon_n(M) > 0$. In this case, we can check that

$$r_{p,p+n} \leq (\epsilon(M)\epsilon_p(G))^{-1} \quad \text{and} \quad \beta(P_{p,p+n}) \leq \left(1 - \epsilon(M)^2\right)^n$$

By (3.9) we conclude that

$$\beta_H(\Phi_{p,p+n}) \leq\ a_H((\epsilon(M)\epsilon_p(G))^{-1})\ \left(1 - \epsilon(M)^2\right)^n$$

In addition, using the same line of reasoning we also prove the following potential free estimates

$$\beta_H(\widehat{\Phi}_{p,p+n}) \quad \leq \quad a_H(\epsilon^{-1}(M))\ \left(1 - \epsilon^2(M)\right)^n$$

Another strategy consists in combining the Markov contraction inequality (3.6) with some entropy inequalities obtained by Ocone in [41] (see also [4]). Using this strategy, we proved in [11] the following annealed contraction inequalities.

Theorem 3.6 *Let η'_n and $\widehat{\eta}'_n := \Psi_n(\eta'_n)$ be an auxiliary model defined with the same random equation as the pair η_n and $\widehat{\eta}_n = \Psi_n(\eta_n)$, but starting at some possibly different $\eta'_0 \in \mathcal{P}(E_0)$. For any $n \in \mathbb{N}$, and any η'_0 we have*

$$\mathbb{E}(\text{Ent}(\widehat{\eta}_n \mid \widehat{\eta}'_n)) \ \leq\ \mathbb{E}(\text{Ent}(\eta_n \mid \eta'_n)) \ \leq\ \left[\prod_{p=1}^{n} \beta(M_p)\right]\ \text{Ent}(\eta_0 \mid \eta'_0) \qquad (3.10)$$

4 Asymptotic analysis

This section is concerned with the convergence analysis of the particle approximation measures introduced in section 2.4. In the first subsection, a stochastic perturbation methodology that allows to enter the stability of the limiting Feynman-Kac semigroup into the convergence analysis of the particle models. The convergence of empirical processes including exponential estimates and $\mathbb{L}_p$-mean error bounds for the McKean particle measures $\mathbb{K}_n^N$ and their density profiles η_n^N are discussed in subsection 4.2. Central limit theorems and large deviation principles are presented respectively in subsection 4.3 and subsection 4.4.

4.1 A stochastic perturbation model

We provide in this section a brief introduction to the asymptotic analysis of the particle approximation models (1.3) as the size of the systems and/or the time horizon tends to infinity. Firstly, we observe that the local sampling errors are expressed in terms of the random fields W_n^N, given for any $f_n \in \mathcal{B}(E_n)$ by the formula

$$W_n^N(f_n) := \sqrt{N}\ [\eta_n^N - \Phi_n(\eta_{n-1}^N)](f_n) = \frac{1}{\sqrt{N}} \sum_{i=1}^N [f_n(\xi_n^i) - K_{n,\eta_{n-1}^N}(f_n)(\xi_{n-1}^i)].$$

Rewritten in a slightly different way, we have the stochastic perturbation formulae

$$\eta_n^N = \Phi_n(\eta_{n-1}^N) + \frac{1}{\sqrt{N}}\ W_n^N$$

with the centered random fields W_n^N with conditional variance functions given by

$$\mathbb{E}(W_n^N(f_n)^2\,|\xi_{n-1}) = \eta_{n-1}^N \left[K_{n,\eta_{n-1}^N}\left((f_n - K_{n,\eta_{n-1}^N}(f_n))^2\right)\right] \tag{4.1}$$

In section 4.3 we shall see that the random fields $(W_n^N)_{n\geq 0}$ behave asymptotically as a sequence of independent Gaussian and centered random fields $(W_n)_{n\geq 0}$ with conditional variance functions given as in (4.1), by replacing the particle empirical measures η_n^N by their limiting values η_n. These fluctuation covariance functions depend on the choice of the McKean interpretation model. To underline the role of the parameter $\epsilon_n(\eta_n)$ given in (2.14), we observe that for any $\mu \in \mathcal{P}(E_{n-1})$ we have that

$$\mu\left[K_{n,\mu}\left((f_n - K_{n,\mu}(f_n))^2\right)\right]$$

$$= \Phi_n(\mu)\left[(f_n - \Phi_n(\mu)(f_n))^2\right] - \epsilon_{n-1}(\mu)^2\ \mu\left[[G_{n-1}(Id - \Psi_{n-1}(\mu))(M_n(f_n))]^2\right]$$

This formula shows that the simple genetic particle model associated with a null parameter $\epsilon_n(\mu) = 0$ is in this sense the less accurate. The following picture gives a sound basis to the main questions related to the convergence analysis of the mean field particle model.

$$\begin{array}{ccccccccc}
\eta_0 & \to & \eta_1=\Phi_1(\eta_0) & \to & \eta_2=\Phi_{0,2}(\eta_0) & \to & \cdots & \to & \Phi_{0,n}(\eta_0) \\
\Downarrow & & & & & & & & \\
\eta_0^N & \to & \Phi_1(\eta_0^N) & \to & \Phi_{0,2}(\eta_0^N) & \to & \cdots & \to & \Phi_{0,n}(\eta_0^N) \\
 & & \Downarrow & & & & & & \\
 & & \eta_1^N & \to & \Phi_2(\eta_1^N) & \to & \cdots & \to & \Phi_{1,n}(\eta_1^N) \\
 & & & & \Downarrow & & & & \\
 & & & & \eta_2^N & \to & \cdots & \to & \Phi_{2,n}(\eta_2^N) \\
 & & & & & & \Downarrow & & \vdots \\
 & & & & & & \eta_{n-1}^N & \to & \Phi_n(\eta_{n-1}^N) \\
 & & & & & & & & \Downarrow \\
 & & & & & & & & \eta_n^N
\end{array}$$

In the above display, the local perturbations W_n^N are represented by the implication sign "$\Downarrow$". This picture yields the following pivotal formula

$$\eta_n^N - \eta_n = \sum_{q=0}^{n} \left[\Phi_{q,n}(\eta_q^N) - \Phi_{q,n}(\Phi_q(\eta_{q-1}^N))\right]$$

with the convention $\Phi_0(\eta_{-1}^N) = \eta_0$ for $p = 0$.

Speaking somewhat loosely, a first order development of the semigroup $\Phi_{q,n}$ around the measure $\Phi_q(\eta_{q-1}^N)$ shows that the "small errors" induced by local perturbations do not propagate, as soon as the semigroup $\Phi_{p,n}$ is sufficiently stable. Using the same line of arguments, the fluctuation analysis of the particle measures η_n^N around their limiting values η_n results from a second order approximation of the semigroup $\Phi_{p,n}$. These two observations will be made clear in section 4.2 and section 4.3, devoted respectively to non asymptotic $\mathbb{L}_p$-mean error bounds and to central limit theorems.

4.2 Convergence of empirical processes

Using the Lipschitz regularity property (3.4) we obtain the following first order estimate :

$$\left|\left[\Phi_{q,n}(\eta_q^N) - \Phi_{q,n}(\Phi_q(\eta_{q-1}^N))\right](f_n)\right| \leq \ 2\, r_{q,n}\, \beta(P_{q,n})\, \frac{1}{\sqrt{N}} \left| W_q^N \left(\overline{P}_{q,n}^{\Phi_q(\eta_{q-1}^N)}(f_n)\right)\right|$$

From these estimates and using the refined version of the Marcinkiewicz-Zygmund lemma presented in [11] (see lemma 7.3.3), we prove the following theorem.

Theorem 4.1 ([11]) *For any $n \geq 0$, $p \geq 1$, any tensor product function $F_n = (f_0 \otimes \ldots \otimes f_n)$, with any functions $f_q \in \mathrm{Osc}_1(E_q)$ for all $q \leq n$, we have*

$$\sup_{N\geq 1} \sqrt{N}\ \mathbb{E}\left(\left|\left[\mathbb{K}_n^N - \mathbb{K}_n\right](F_n)\right|^p\right)^{\frac{1}{p}} < \infty$$

and

$$\sqrt{N}\ \mathbb{E}\left(|[\eta_n^N - \eta_n](f_n)|^p\right)^{\frac{1}{p}} \leq \ c(n)\ d(p)^{1/p} \tag{4.2}$$

with $c(n) \leq 2 \sum_{q=0}^{n} r_{q,n} \beta(P_{q,n})$*, and the sequence of parameters* $d(p)$*, with* $p \geq 1$*, given by*

$$d(2p) = (2p)_p \; 2^{-p} \quad \textit{and} \quad d(2p-1) = \frac{(2p-1)_p}{\sqrt{p-1/2}} \; 2^{-(p-1/2)}$$

The $\mathbb{L}_p$ mean error estimates in the right and side of (4.2) can be used to derive the following exponential estimate

$$\forall \epsilon > 0 \qquad \mathbb{P}(|\eta_n^N(f_n) - \eta_n(f_n)| > \epsilon) \leq (1 + \epsilon\sqrt{N/2}) \; \exp\left(-\frac{N \; \epsilon^2}{2c(n)^2}\right) \tag{4.3}$$

We associate with a collection of measurable functions $f : E \to \mathbb{R}$, with $\|f\| \leq 1$, the Zolotarev seminorm on $\mathcal{P}(E)$ defined by

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

(see for instance [42]). To avoid some unnecessary technical measurability questions, we further suppose that $\mathcal{F}$ is separable in the sense that it contains a countable and dense subset. To measure the size of a given class $\mathcal{F}$, one considers the covering numbers $N(\epsilon, \mathcal{F}, L_p(\mu))$ defined as the minimal number of $L_p(\mu)$-balls of radius $\epsilon > 0$ needed to cover $\mathcal{F}$. By $\mathcal{N}(\epsilon, \mathcal{F})$, $\epsilon > 0$, and by $I(\mathcal{F})$ we denote the uniform covering numbers and entropy integral given by

$$\mathcal{N}(\epsilon, \mathcal{F}) = \sup\big\{\mathcal{N}(\epsilon, \mathcal{F}, L_2(\eta)); \; \eta \in \mathcal{P}(E)\big\}$$

and

$$I(\mathcal{F}) = \int_0^1 \sqrt{\log\left(1 + \mathcal{N}(\epsilon, \mathcal{F})\right)} \, d\epsilon$$

Various examples of classes of functions with finite covering and entropy integral are given in the book of Van der Vaart and Wellner [44]. The estimation of the quantities introduced above depends on several deep results on combinatorics that are not discussed here.

Let $\mathcal{F}_n$ be a countable collection of functions f_n with $\|f_n\| \leq 1$ and finite entropy $I(\mathcal{F}_n) < \infty$. Suppose that the Markov transitions M_n have the form $M_n(u, dv) = m_n(u, v) \; p_n(dv)$ for some measurable function m_n on $(E_{n-1} \times E_n)$ and some $p_n \in \mathcal{P}(E_n)$. Also assume that we have for some collection of mappings θ_n on E_n

$$\sup_{u \in E_{n-1}} |\log m_n(u, v)| \leq \theta_n(v) \quad \text{with} \quad p_n(e^{3\theta_n}) < \infty \tag{4.4}$$

In this situation, we obtain the following fluctuation result.

Theorem 4.2 *For any* $n \geq 0$ *and* $p \geq 1$ *we have*

$$\sqrt{N} \; \mathbb{E}\left(\|\eta_n^N - \eta_n\|_{\mathcal{F}_n}^p\right)^{\frac{1}{p}} \leq a \; [p/2]! \; [I(\mathcal{F}_n) + c(n-1) \; r_{n-1,n} \; p_n(e^{3\theta_n})] \tag{4.5}$$

with a collection of constants $a < \infty$ *and* $c(n) \leq \sum_{q=0}^{n} r_{q,n} \beta(P_{q,n})$*.*

4.3 Fluctuation analysis

The fluctuation analysis of the particle measures η_n^N around their limiting values η_n is essentially based on the asymptotic analysis of the local sampling errors W_n^N associated with the particle approximation sampling steps. The next central limit theorem for random fields is pivotal. Its complete proof can be found in [11] (see theorem 9.3.1 and corollary 9.3.1 on pages 295-298).

Theorem 4.3 *For any fixed time horizon $n \geq 1$, the sequence $(W_p^N)_{1\leq p\leq n}$ converges in law, as N tends to infinity, to a sequence of n independent, Gaussian and centered random fields $(W_p)_{1\leq p\leq n}$; with, for any $f_p, g_p \in \mathcal{B}(E_p)$, and $1 \leq p \leq n$,*

$$\mathbb{E}(W_p(f_p)W_p(g_p)) = \eta_{p-1}K_{p,\eta_{p-1}}([f_p - K_{p,\eta_{p-1}}(f_p)][g_p - K_{p,\eta_{p-1}}(g_p)]) \quad (4.6)$$

Using the pair of decompositions

$$\begin{aligned}\gamma_p^N Q_{p,n} - \gamma_p Q_{p,n} &= \sum_{q=0}^{p} \gamma_q^N(1)\left[\eta_q^N - \Phi_q\left(\eta_{q-1}^N\right)\right] Q_{q,n} \\ &= \frac{1}{\sqrt{N}} \sum_{q=0}^{p} \gamma_q^N(1)\ W_q^N Q_{q,n}\end{aligned}$$

and

$$[\eta_n^N - \eta_n](f_n) = \frac{\gamma_n(1)}{\gamma_n^N(1)}\ [\gamma_n^N - \gamma_n]\left(\frac{1}{\gamma_n(1)}(f_n - \eta_n(f_n))\right)$$

we readily deduce the following corollary :

Corollary 4.4 *For any fixed time horizon $n \geq 1$, the sequence of random fields*

$$W_n^{N,\gamma} := \sqrt{N}\ [\gamma_n^N - \gamma_n] \quad \textit{and} \quad W_n^{N,\eta} := \sqrt{N}\ [\eta_n^N - \eta_n]$$

converges in law, as N tends to infinity, to a sequence of n independent, Gaussian and centered random fields W_n^γ and W_n^η; with, for any $f_n \in \mathcal{B}(E_n)$

$$W_n^\eta(f_n) = W_n^\gamma\left(\frac{1}{\gamma_n(1)}(f_n - \eta_n(f_n))\right)$$

and

$$W_n^\gamma(f_n) = \sum_{q=0}^{n} \gamma_q(1)\ W_q[Q_{q,n}(f_n)]$$

The random fields $W_n^{N,\gamma}$ and $W_n^{\eta,N}$ can also be regarded as empirical processes indexed by the collection of bounded measurable functions $\mathcal{F}_n \subset \mathcal{B}(E_n)$. If $\mathcal{F}_n$ is a countable collection of functions f_n, with $\|f_n\| \leq 1$ and $I(\mathcal{F}_n) < \infty$, then the $\mathcal{F}_n$-indexed process $W_n^N(f_n)$, $f_n \in \mathcal{F}_n$, is asymptotically tight (see for instance lemma 9.6.1 in [11]). In this situation, the random fields $W_n^{N,\gamma}$ and $W_n^{\eta,N}$ also converge in law in $l^\infty(\mathcal{F}_n)$ to the Gaussian processes W_n^γ and W_n^η.

The fluctuation analysis of the random fields $\mathbb{K}_n^N$ around the McKean measure $\mathbb{K}_n$ is technically much more involved. For completeness, we have chosen to present this result in a rather simple form. The complete proof can be found in section 9.7 of the book [11]. We only examine the situation with a null acceptance rate $\epsilon_n(\mu) = 0$, and we further suppose that condition (4.4) is satisfied with $p_n(e^{r\theta_n})$, for any $r \geq 1$.

For any $x = (x_0, \ldots, x_n)$ and $z = (z_0, \ldots, z_n) \in \Omega_n := (E_0 \times \ldots \times E_n)$ we set

$$a_n(x,z) := b_n(x,z) - \int_{\Omega_n} \mathbb{K}_n(dx')\ b_n(x',z) \qquad \text{and} \qquad b_n(x,z) := \sum_{k=1}^{n} c_k(x,z)$$

In the above displayed formulae c_k stands for the collection of functions on Ω_n^2 given by

$$c_k(x,z) := \frac{G_{k-1}(z_{k-1})\ m_k(z_{k-1},x_k)}{\int_{E_{k-1}} G_{k-1}(u_{k-1})\ m_k(u_{k-1},x_k)\ \eta_{k-1}(du)}$$

Under our assumptions,we can prove that the integral operator A_n given by for any $F_n \in \mathbb{L}_2(\mathbb{K}_n)$ by

$$\mathcal{A}_n(F_n)(x) \quad = \quad \int_{\Omega_n} a_n(z,x)\ F_n(z)\ \mathbb{K}_n(dz)$$

is a Hilbert-Schmidt operator on $\mathbb{L}_2(\mathbb{K}_n)$ and the operator $(I - A_n)$ is invertible.

Extending the fluctuation analysis developed by T. Shiga and H. Tanaka in [43] to nonlinear mean field interaction models, we prove in [20] the following central limit theorem.

Theorem 4.5 *The random field $\mathcal{W}_n^N := \sqrt{N}\ [\mathbb{K}_n^N - \mathbb{K}_n]$ converges as $N \to \infty$ to a centered Gaussian field $\mathcal{W}_n$ satisfying*

$$\mathbb{E}\left(\mathcal{W}_n(F_n)W(F'_n)\right) = \langle (I-\mathcal{A}_n)^{-1}(F_n - \mathbb{K}_n(F_n)), (I-\mathcal{A}_n)^{-1}(F'_n - \mathbb{K}_n(F'_n))\rangle_{\mathbb{L}_2(\mathbb{K}_n)}$$

for any $F_n, F'_n \in \mathbb{L}_2(\mathbb{K}_n)$, in the sense of convergence of finite-dimensional distributions.

4.4 Large deviations principles

In this section, we introduce the reader to the large deviation analysis of mean field particle models associated with a given McKean interpretation of a measure valued equation. We start with the derivation of the large deviation principles (*abbreviated LDP*) combining Sanov's theorem with the Laplace-Varadhan integral lemma. This rather elementary result relies on an appropriate regularity condition on the McKean transitions under which the law of the N-particle model is absolutely continuous w.r.t. the law of N independent copies of the Markov chain associated with the McKean interpretation model.

We equip the space of all finite and signed measures $\mathcal{M}(E)$ on a Polish space $(E, \mathcal{E})$ with the weak topology generated by the open open neighborhoods

$$\mathcal{V}_{f,\epsilon}(\mu) = \{\eta \in \mathcal{P}(E)\ :\ |\eta(f) - \mu(f)| < \epsilon\}$$

where f is a bounded continuous function on E, $\mu \in \mathcal{P}(E)$, and $\epsilon \in (0, \infty)$. Using this topological framework, if we take $E = E_n$ and $f = f_n$, then the deviant event presented in (4.3) is equivalently expressed in terms of a basis neighborhood of the McKean measure

$$\mathbb{P}(|\eta_n^N(f_n) - \eta_n(f_n)| > \epsilon) = \mathbb{P}(\eta_n^N \not\in \mathcal{V}_{f_n,\epsilon}(\eta_n))$$

as soon as the function f_n is continuous. One objective of the forthcoming analysis is to estimate the exact exponential deviation of the particle measures η_n^N and more generally $\mathbb{K}_n^N$ around there limiting values η_n and $\mathbb{K}_n$. For instance, we would like is to compute the following quantities for any continuous bounded functions f_n or F_n on the state space E_n or on the path space $\Omega_n := (E_0 \times \ldots \times E_n)$

$$\lim_{N\to\infty} \frac{1}{N} \log \mathbb{P}(|\eta_n^N(f_n) - \eta_n(f_n)| > \epsilon)$$

and

$$\lim_{N\to\infty} \frac{1}{N} \log \mathbb{P}(|\mathbb{K}_n^N(F_n) - \mathbb{K}_n(F_n)| > \epsilon)$$

To describe precisely this result we need to introduce another round of notation. We let $P_n^N \in \mathcal{P}(\Omega_n^N)$ be the distribution of the mean field N-interacting particle model $\left(\xi_0^i, \ldots \xi_n^i\right)_{1\leq i\leq N}$ introduced in (1.3), and we denote by $\overline{P}_n^N = P_n^N \circ (\pi_n^N)^{-1}$ the image distribution of the empirical measures $\mathbb{K}_n^N$, with the empirical projection mapping π_n^N given by

$$\pi_n^N \;:\; \omega = (\omega_0^i, \ldots, \omega_n^i)_{1\leq i\leq N} \in \Omega_n^N \longrightarrow \pi_n^N(\omega) = \frac{1}{N}\sum_{i=1}^N \delta_{(\omega_0^i,\ldots,\omega_n^i)} \in \mathcal{P}(\Omega_n)$$

We recall that the sequence $\overline{P}_n^N$ satisfies the LDP with the good rate function I_n as soon as we find a mapping lower semicontinuous mapping I_n from $\mathcal{P}(\Omega_n)$ into $[0,\infty]$ and such that

$$-\inf_{\overset{\circ}{A}} I_n \leq \liminf_{N\to\infty} \log \frac{1}{N}\overline{P}_n^N(A) \leq \limsup_{N\to\infty} \log \frac{1}{N}\overline{P}_n^N(A) \leq -\inf_{\overline{A}} I_n \tag{4.7}$$

for any Borel subset $A \subset \mathcal{P}(\Omega_n)$, where $\overset{\circ}{A}$ and $\overline{A}$ denote respectively the interior and the closure of the set A.

By a direct application of Sanov's theorem, under the tensor product measure $\mathbb{K}_n^{\otimes N}$ the laws $\overline{Q}_n^N := \mathbb{K}_n^{\otimes N} \circ (\pi_n^N)^{-1}$ of the empirical measures $\mathbb{K}_n^N$ associated with N independent path particles with common distribution $\mathbb{K}_n$ satisfy the LDP with a good rate function H_n given by the Boltzmann relative entropy defined by

$$H_n \;:\; \mu \in \mathcal{P}(\Omega_n) \longrightarrow H(\mu) = \mathrm{Ent}(\mu \mid \mathbb{K}_n) \in [0,\infty]$$

To get one step further, we suppose that condition (4.4) is satisfied with $p_n(e^{r\theta_n})$, for any $r \geq 1$. In this condition, the measures $\overline{P}_n^N$ and $\overline{Q}_n^N$ are absolutely continuous with a Radon-Nikodim derivative defined by

$$d\overline{P}_n^N / d\overline{Q}_n^N = \exp\left(NV_n\right) \qquad \overline{Q}_n^N\text{-a.e.}$$

In the above display, the function V_n is given for any probability measure $\mu \in \mathcal{P}(\Omega_n)$ by te following formula

$$
\begin{aligned}
V_n(\mu) &:= \int \log \frac{d\mathbb{M}_n(\mu)}{d\mathbb{M}_n(\mathbb{K}_n)}\, d\mu \\
&= \sum_{p=1}^{n} \int_{E_{p-1}\times E_p} \mu_{p-1,p}(d(u_{p-1},u_p))\, \log\left[\frac{dK_{p,\mu_{p-1}}(u_{p-1},.)}{dK_{p,\eta_{p-1}}(u_{p-1},.)}(u_p)\right]
\end{aligned}
$$

with the mapping

$$\mathbb{M}_n : \mu \in \mathcal{P}(\Omega_n) \to \mathbb{M}_n(\mu) \in \mathcal{P}(\Omega_n)$$

defined by

$$\mathbb{M}_n(\mu)(d(x_0,\ldots,x_n)) = \eta_0(dx_0)\, K_{1,\mu_0}(x_0,dx_1)\ldots K_{n,\mu_{n-1}}(x_{n-1},dx_n)$$

and where $\mu_p \in \mathcal{P}(E_p)$ stands for the pth time marginal of μ with $0 \leq p \leq n$.

Under our assumptions, using an extended version of the Laplace-Varadhan integral lemma (see for instance lemma 10.4.1 in [11]) we prove the following theorem.

Theorem 4.6 *The sequence of measures $\overline{\mathbb{P}}_n^N$ satisfies the LDP with the good rate function*

$$I_n(\mu) := [H_n - V_n](\mu) = \mathrm{Ent}(\mu \mid \mathbb{M}_n(\mu)) \quad \text{and} \quad I_n(\mu) = 0 \Leftrightarrow \mu = \mathbb{K}_n$$

As we mentioned above, the large deviation analysis we have described is not restricted to mean field interpretations of the nonlinear filtering equations. A more general result for an abstract class of McKean kernels $(K_{n,\eta})_{n,\eta}$ is provided in theorem 10.1.1 in [11].

Working a little harder, we can also derive a LDP for the τ-topology for the distribution of the particle density profiles $(\eta_p^N)_{0\leq p\leq n}$ associated with the McKean transitions $K_{n,\eta}(x,.) := \Phi_n(\eta)$. For the convenience of the reader, we recall that the τ-topology on a set of probability measures is generated by the sets $\mathcal{V}_{f,\epsilon}(\mu)$, with non necessarily continuous functions $f \in \mathcal{B}(E)$. The following theorem is taken from a joint work of the first author with D. Dawson [16].

Theorem 4.7 *Assume that for any $n \geq 1$, there exists some reference measure $\lambda_n \in \mathcal{P}(E_n)$ and some parameters $\rho_n > 0$ such that*

$$\forall \mu \in (\mathcal{P}(E_{n-1}) \qquad \rho_n\ \lambda_n \leq \Phi_n(\mu) \ll \lambda_n \tag{4.8}$$

In this situation, the law of the flow of particle density profiles $(\eta_p^N)_{0\leq p\leq n}$ satisfies the LDP for the product τ-topology with the good rate function J_n on $\prod_{p=0}^n \mathcal{P}(E_p)$ given by

$$J_n((\mu_p)_{0\leq p\leq n}) = \sum_{p=0}^{n} \mathrm{Ent}(\mu_p \mid \Phi_p(\mu_{p-1}))$$

In the context of Feynman-Kac models, it can be checked that the one-step mappings Φ_n are τ-continuous and condition (4.8) is met as soon as for any $x_{n-1} \in E_{n-1}$ we have

$$\rho_n\ \lambda_n \leq M_n(x_{n-1},.) \ll \lambda_n$$

For instance, if we take $E_n = \mathbb{R}$ and $M_n(x, dy) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(y-a_n(x))^2} dy$, where a_n is a bounded measurable drift function on $\mathbb{R}$, then condition (4.8) is met with the reference measure given by $\lambda_n(dy) = p_n(dy)/p_n(\mathbb{R})$ and with

$$p_n(dy) =_{\text{def.}} \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2} - |y| \, \|a_n\|} dy \quad \text{and} \quad \rho_n = e^{-\|a_n\|^2/2} \, p_n(\mathbb{R})$$

5 Propagations of chaos properties

The initial configuration of the N-particle mean field model (1.3) consists in N independent and identically distributed random variables. Their common law is dictated by the initial distribution of the underlying nonlinear equation (1.2). In this sense, the initial system is in "complete chaos" and the law of each particle perfectly fits the initial distribution.

During their time evolution the particles interact one another. The nature of the interactions depends on the McKean interpretation model. Nevertheless, in all interesting cases the independence property and the adequacy of the laws of the particles with the desired distributions are broken. The propagation of chaos properties of mean field particle models presented in this section can be seen as a way to measure these properties. In the first subsection, we estimate the relative entropy of the laws of the first q paths of a complete genealogical tree model with respect to the q-tensor product of the McKean measure. In the second part of this section, we present an original Laurent type and algebraic tree-based integral representations of particle block distributions. In contrast to the first entropy approach, this technique applies to any genetic type particle model without any regularity condition on the mutation transition.

5.1 Relative entropy estimates

In this section, we provide strong propagation estimates with respect to the relative entropy criterion for the interacting particle system associated with the McKean interpretation model defined in (1.3). We further assume that $\epsilon_n(\mu)$ does not depend on the measure μ, and the Markov transitions M_n satisfy the regularity condition (4.4), with $p_n = \eta_n$. We let P_n^N be the distribution of the mean field N-interacting particle model $\left(\xi_0^i, \ldots \xi_n^i\right)_{1 \le i \le N}$. By the deviation analysis provided in section 4.4, we have that

$$\frac{dP_n^N}{d\mathbb{K}_n^{\otimes N}} = \exp\left(N(V_n \circ \pi_n^N)\right) \qquad \mathbb{K}_n^{\otimes N}\text{-a.e.}$$

We let $\mathbb{P}_n^{(N,q)}$ be the distribution of the first q path particles $\left(\xi_0^i, \ldots \xi_n^i\right)_{1 \le i \le q}$, with $q \le N$. Using an elegant lemma of Csiszar on the entropy of the first q-coordinates of an exchangeable measure (see [9], or lemma 8.5.1 in [11]), we find that

$$\text{Ent}\left(\mathbb{P}_n^{(N,q)} \,\middle|\, \mathbb{K}_n^{\otimes q}\right) \quad \le \quad 2 \, \frac{q}{N} \, \text{Ent}\left(\mathbb{P}_n^N \mid \mathbb{K}_n^{\otimes N}\right) = 2q \, \mathbb{E}\left(V_n(\mathbb{K}_n^N)\right) \quad (5.1)$$

By the exchangeability property of the particle model, we also have that

$$\mathbb{E}(V_n(\mathbb{K}_n^N)) = \sum_{p=0}^{n-1} \mathbb{E}\left[\text{Ent}\left(K_{p+1,\eta_p^N}(\xi_p^1, .) \mid K_{p+1,\eta_p}(\xi_p^1, .)\right)\right]$$

Using the fact that $\mathrm{Ent}(\mu|\eta) \leq \|1 - d\mu/d\eta\|^2_{2,\eta}$, and estimating the local Lipschitz continuity coefficient of the mappings $\mu \mapsto \frac{dK_{n,\mu}(u,\centerdot)}{dK_{n,\eta_{n-1}}(u,\centerdot)}(v)$ at $\mu = \eta_{n-1}$, we prove the following theorem.

Theorem 5.1 *For any $q \leq N$ and any time horizon n, we have the relative entropy estimate*

$$\mathrm{Ent}\left(\mathbb{P}_n^{(N,q)} \left| \mathbb{K}_n^{\otimes q}\right.\right) \;\leq\; \frac{q}{N}\; c(n)$$

with

$$c(n) \leq c' \sum_{p=0}^{n-1} r_p^2 \; (1 + \eta_{p+1}(e^{2\theta_{p+1}}) \sum_{q=0}^{p} r_{q,p} \beta(P_{q,p})$$

and some universal finite constant c'.

5.2 Polynomial tree-based expansions

This section is concerned with propagation of chaos properties of the mean field particle model associated with the null acceptance rate parameter $\epsilon_n(\mu) = 0$. Without any regularity conditions on the Markov transitions M_n it is more all less well known that we have the following strong propagation of chaos estimate

$$\|\mathbb{P}_n^{(N,q)} - \mathbb{K}_n^{\otimes q}\|_{\mathrm{tv}} \leq c(n)\; q^2/N \tag{5.2}$$

for some finite constant, which only depends on the time parameter n. The complete proof of this result can be founded in [11] (see for instance theorem 8.3.3 and theorem 8.3.4).

The main object of this section is to provide an explicit asymptotic functional expansion of the law of the first q particles at a given time n with respect to the precision parameter N. These expansions at any order extend the sharp but first order propagation of chaos estimates developed by the first author with A. Doucet, and G.W. Peters in a recent article [19].

Next, for any $q \leq N$ we consider the q-tensor product occupation measures, resp. the restricted q-tensor product occupation measures, on E_n^q defined by

$$(\eta_n^N)^{\otimes q} = \frac{1}{N^q} \sum_{a \in [N]^{[q]}} \delta_{(\xi_n^{(a(1),N)},\ldots,\xi_n^{(a(q),N)})}$$

and

$$(\eta_n^N)^{\odot q} = \frac{1}{(N)_q} \sum_{a \in \langle q,N\rangle} \delta_{(\xi_n^{(a(1),N)},\ldots,\xi_n^{(a(q),N)})} \tag{5.3}$$

In the above display, $[N]^{[q]}$, resp. $\langle q, N\rangle$, stands for the set of all N^q mappings, resp. all $(N)_q := N!/(N-q)!$ one-to-one mappings, from the set $[q] := \{1,\ldots,q\}$ into $[N] := \{1,\ldots,N\}$. The unnormalized versions of these measures are simply defined by

$$(\gamma_n^N)^{\otimes q} := (\gamma_n^N(1))^q \times (\eta_n^N)^{\otimes q} \quad \text{and} \quad (\gamma_n^N)^{\odot q} := (\gamma_n^N(1))^q \times (\eta_n^N)^{\odot q}$$

One central problem is to obtain functional expansions, with respect to the precision parameter N, of the pair of particle block distributions defined for any $F \in \mathcal{B}(E_n^q)$ by the formulae

$$\mathbb{P}_{n,q}^N(F) := \mathbb{E}((\eta_n^N)^{\odot q}(F)) = \mathbb{E}(F(\xi_n^{(1,N)},\ldots,\xi_n^{(q,N)}))$$

and

$$\mathbb{Q}_{n,q}^N(F) := \mathbb{E}((\gamma_n^N)^{\otimes q}(F)) \tag{5.4}$$

In this Note, we design a coalescent tree-based representation of the unnormalized particle distributions $\mathbb{Q}_{n,q}^N$ with the following polynomial form

$$\mathbb{Q}_{n,q}^N = \gamma_n^{\otimes q} + \sum_{1\leq k\leq (q-1)(n+1)} \frac{1}{N^k}\, \partial^k \mathbb{Q}_{n,q} \tag{5.5}$$

In the above display, $\partial^k \mathbb{Q}_{n,q}$ stands for a sum of signed, and weak derivative measures, whose values can be explicitly described in terms of a class of forests with maximal coalescent degree k.

These Laurent type expansions reflect the complete interaction structure of the particle model, the k-th order terms represent the $1/N^k$-contributions of mean field particle scenarios with an interaction degree k.

The analysis of the distributions $\mathbb{P}_{n,q}^N$ is a little more involved, combining judicious renormalization techniques on path spaces, with colored tree-based combinatorial expansions (cf. [31]).

$$\mathbb{P}_{n,q}^N \simeq \eta_n^{\otimes q} + \sum_{1l\leq l\leq k} \frac{1}{N^l}\, \partial^l \mathbb{P}_{n,q} + \frac{1}{N^{k+1}}\, \partial^{k+1}\mathbb{P}_{n,q}^N$$

with

$$\sup_{N\geq 1} \|\partial^{k+1}\mathbb{P}_{n,q}^N\|_{\rm tv} < \infty$$

5.3 Coalescent tree based representations

Let $\mathcal{A}_{n,q}$ be the set of $(n+1)$-sequences $\mathbf{a} = (a_p)_{0\leq p\leq n}$ of mappings a_p from $[q]$ into itself. By $|a|$, we denote the cardinality of the set $a([q])$; and for $\mathbf{a} \in \mathcal{A}_{n,q}$, we write $|\mathbf{a}|$ the integer sequence $(|a_i|)_{0\leq i\leq n}$. For any pair of integers $q \leq N$, we use the the multi index notation $(\mathbf{q})_{|\mathbf{a}|} = \prod_{0\leq k\leq n} (q)_{|a_k|}$. For $b \in [q]^{[q]}$ and $F \in \mathcal{B}(E_n^q)$, we define :

$$D_b(F)(x^1,\ldots,x^q) = F(x^{b(1)},\ldots,x^{b(q)}) \ , \qquad (F)_{sym} := \frac{1}{p!}\sum_{\sigma\in\mathcal{G}_q} D_\sigma F \ . \tag{5.6}$$

For any air of mappings $a, b \in [q]^{[q]}$ we have the composition formula $D_a D_b = D_{ab}$. We notice that

$$(\eta_n^N)^{\otimes q}(F) = (\eta_n^N)^{\otimes q}((F)_{sym})$$

and

$$(\eta_n^N)^{\odot q}(F) = (\eta_n^N)^{\odot q}((F)_{sym})$$

So we may assume in our forthcoming computations on q-tensor products occupation measures that F is in $\mathcal{B}^{sym}(E_n^q)$, the subset of $\mathcal{B}(E_n^q)$ of symmetric functions. We also observe that

$$(\eta_n^N)^{\otimes q} = (\eta_n^N)^{\odot q} \left(\frac{1}{N^q} \sum_{b\in[q]^{[q]}} \frac{(N)_{|b|}}{(q)_{|b|}}\, D_b \right) \tag{5.7}$$

To check this formula, we first notice that for any $c \in [N]^{[q]}$, there are $(N - |c|)_{q-|c|} \times (q)_{|c|}$ different ways to write $c = ab \in \langle q, N\rangle \circ [q]^{[q]}$. On the other hand, if $a \in \langle q, N\rangle$, then we have that $|b| = |c|$ and $\frac{(N)_{|c|}}{(q)_{|c|}} \times \frac{(N-|c|)_{q-|c|} \times (q)_{|c|}}{(N)_q} = 1$.

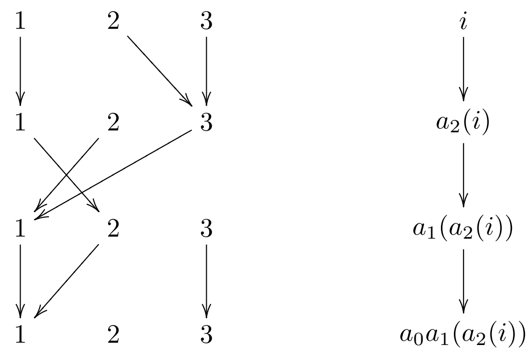
The partial derivative measures $\partial^k \mathbb{Q}_{n,q}$ involved in the Laurent type expansions (5.5) are intimately related to the following measure-valued functional

$$\Delta_{n,q} \ : \ \mathbf{a} \in \mathcal{A}_{n,q} \mapsto \Delta_{n,q}^{\mathbf{a}} = \left(\eta_0^{\otimes q} D_{a_0} Q_1^{\otimes q} D_{a_1} \ldots Q_n^{\otimes q} D_{a_n}\right) \in \mathcal{M}(E_n^q) \tag{5.8}$$

For instance, for $q = 3$ and $n = 2$ and the sequence of mappings

$$\mathbf{a} = (a_0, a_1, a_2)$$

given below



we have the formula

$$\begin{aligned}\Delta_{n,q}^{\mathbf{a}}(F) &= \int \eta_0(dx^1)\eta_0(dx^2)\eta_0(dx^3)\ Q_1(x^1, dy^1)Q_1(x^1, dy^2)Q_1(x^3, dy^3) \\ &\qquad \times\ Q_2(y^1, dz^2)Q_2(y^1, dz^3)Q_2(y^2, dz^1)F(z^1, z^3, z^3)\end{aligned}$$

The first steps to these integral expansions are given by the following fundamental lemma.

Lemma 5.1 *For any particle block size $q \leq N$, any time horizon $n \in \mathbb{N}$, and any $F \in \mathcal{B}(E_n^q)$, we have*

$$\begin{aligned}\mathbb{Q}_{n,q}^N(F) &= \frac{1}{N^q} \sum_{a \in [q]^{[q]}} \frac{(N)_{|a|}}{(q)_{|a|}}\ \mathbb{Q}_{n-1,q}^N(Q_n^{\otimes q} D_a F) \\ &= \frac{1}{N^{q(n+1)}} \sum_{\mathbf{a} \in \mathcal{A}_{n,q}} \frac{(\mathbf{N})_{|\mathbf{a}|}}{(\mathbf{q})_{|\mathbf{a}|}}\ \Delta_{n,q}^{\mathbf{a}}(F)\end{aligned} \tag{5.9}$$

Proof:

We let A_q^N be an $[q]^{[q]}$-valued random variable (independent of the particle model) with distribution $\frac{1}{N^q}\ \sum_{a \in [q]^{[q]}} \frac{(N)_{|a|}}{(q)_{|a|}} \delta_a$. Indeed, combining the definition of the particle model with (5.7), we first find that

$$\begin{aligned}
\mathbb{E}\left(\left(\gamma_n^N\right)^{\otimes q}(F)\ \mid \xi_{n-1}\right) &= \gamma_n^N(1)^q \times \mathbb{E}\left(\left(\eta_n^N\right)^{\otimes q}(F)\ \mid \xi_{n-1}\right) \\
&= \left(\gamma_n^N(1)\right)^q \times \mathbb{E}\left(\Phi_n(\eta_{n-1}^N)^{\otimes q} D_{A_q^N}(F)\ \mid \xi_{n-1}\right) \\
&= \mathbb{E}\left(\left(\gamma_{n-1}^N\right)^{\otimes q}(Q_n^{\otimes q} D_{A_q^N} F)\ \mid \xi_{n-1}\right)
\end{aligned}$$

Integrating over the past, the end of the proof is straightforward. ■

Next, we turn our attention to a coalescent tree based formulation of the integral expansion (5.9). We start with recalling some more or less classical terminology on trees and forests. A tree (respectively a planar tree) is a (isomorphism class of) finite non-empty oriented connected (and respectively planar) graph $\mathbf{t}$ without loops such that any vertex of $\mathbf{t}$ has at most one outgoing edge. Paths are oriented from the vertices to the root. A forest $\mathbf{f}$ is a multiset of trees, that is a set of trees, with repetitions of the same tree allowed, or equivalently an element of the commutative monoid $\langle \mathcal{T} \rangle$ on $\mathcal{T}$, with the empty graph $T_0 = \emptyset$ as a unit. Since the algebraic notation is the most convenient, we write $\mathbf{f} = \mathbf{t}_1^{m_1} ... \mathbf{t}_k^{m_k}$, for the forest with the tree $\mathbf{t}_i$ appearing with multiplicity m_i, $i \leq k$. When $\mathbf{t}_i \neq \mathbf{t}_j$ for $i \neq j$, we say that $\mathbf{f}$ is written in normal form. A planar forest $\mathbf{f}'$ is an ordered sequence of planar trees. Planar forests can be represented by noncommutative monomials (or words) on the set of planar trees. The sets of forests and planar forests with height $(n+1)$, and with q vertices at each level set are written $\mathcal{F}_{q,n}$ and $\mathcal{PF}_{q,n}$.

A sequence $\mathbf{a} \in \mathcal{A}_{q,n}$ is naturally associated a forest $F(\mathbf{a})$: the one with one vertex for each element of $[q]^{n+2}$, and a edge for each pair $(i, a_k(i)), i \in [q]$. The sequence can also be represented graphically uniquely by a planar graph $J(\mathbf{a})$, where however the edges between vertices at level $k+1$ and k are allowed to cross. We call such a planar graph, where paths between vertices are entangled, *a jungle*. The set of such jungles is written $\mathcal{J}_{q,n}$. In a planar forest $\mathbf{f}$, vertices at the same level $k \geq 0$, are naturally ordered from left to right, and therefore in bijection with $[q]$. Planar forests $\mathbf{f} \in \mathcal{PF}_{q,n}$ of height $(n+1)$ are therefore canonically in bijection with sequences $\mathbf{a}$ of weakly increasing map from $[q]$ into itself.

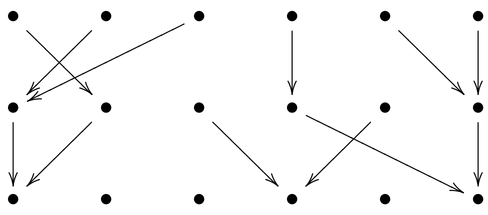


Figure 1: The entangled graph representation of a jungle with the same underlying graph as the planar forest in Fig. 2.

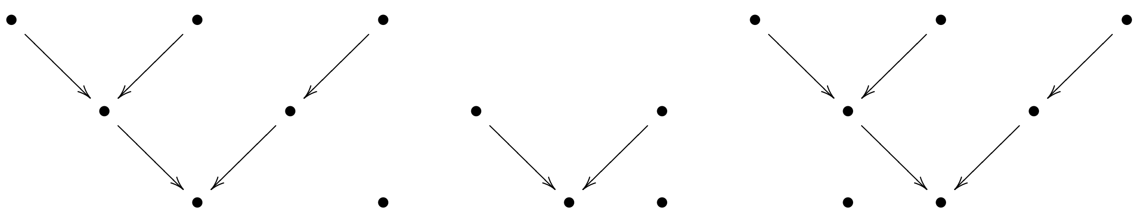


Figure 2: a graphical representation of a planar forest $\mathbf{f} = \mathbf{t}_1\mathbf{t}_3\mathbf{t}_2\mathbf{t}_3\mathbf{t}_3\mathbf{t}_1$.

We let $\mathcal{G}_q$ be the symmetric group of all permutations of $[q]$. The group $\mathcal{G}_q^{n+2}$ acts naturally on sequences of maps $\mathbf{a} \in \mathcal{A}_{q,n}$, and on jungles $J(\mathbf{a}) \in \mathcal{J}_{q,n}$ by permutation of the vertices at each level. More precisely, for all $\mathbf{a} \in \mathcal{A}_{n,q}$ and all $\mathbf{s} = (s_0, ..., s_{n+1}) \in \mathcal{G}_q^{n+2}$ by the pair of formulae

$$\mathbf{s}(\mathbf{a}) := (s_0 a_0 s_1^{-1}, s_1 a_1 s_2^{-1}, ..., s_n a_n s_{n+1}^{-1}) \quad \text{and} \quad \mathbf{s}J(\mathbf{a}) := J(\mathbf{s}(\mathbf{a})) \tag{5.10}$$

Notice that if two sequences $\mathbf{a}$ and $\mathbf{b} \in \mathcal{A}_{\mathbf{q},\mathbf{n}}$ differ only by the order of the vertices in $J(\mathbf{a})$ and $J(\mathbf{b})$, that is by the action of an element of $\mathcal{G}_{\mathbf{p}}$, then the associated forests are identical: $F(\mathbf{a}) = F(\mathbf{b})$. Moreover, the converse is true: if $F(\mathbf{a}) = F(\mathbf{b})$, then $J(\mathbf{a})$ and $J(\mathbf{b})$ differ only by the ordering of the vertices, since they have the same underlying non planar graph. In this situation, $\mathbf{a}$ and $\mathbf{b}$ belong to the same orbit under the action of $\mathcal{G}_{\mathbf{p}}$. In particular, the set of equivalence classes of jungles in $\mathcal{J}_{q,n}$ under the action of the permutation groups $\mathcal{G}_{q,n}$ is in bijection with the set of forests $\mathcal{F}_{q,n}$. We denote by $B(\mathbf{t})$ the forest deduced from cutting the root of tree $\mathbf{t}$; that is, removing its root vertex, and all its incoming edges. In the reverse angle, we denote by $B^{-1}(\mathbf{f})$ the tree deduced from the forest $\mathbf{f}$ by adding a common root to its rooted tree. The symmetry multiset $\mathbf{S}(\mathbf{t})$ of a tree $\mathbf{t} = B^{-1}(\mathbf{t}_1^{m_1} \ldots \mathbf{t}_k^{m_k})$ associated with a forest written in normal form, is defined by $\mathbf{S}(\mathbf{t}) := (m_1, \ldots, m_k)$. The symmetry multiset of a forest is given by

$$\mathbf{S}(\mathbf{t}_1^{m_1} \ldots \mathbf{t}_k^{m_k}) := \left(\underbrace{\mathbf{S}(\mathbf{t}_1), \ldots, \mathbf{S}(\mathbf{t}_1)}_{m_1-\text{terms}}, \ldots, \underbrace{\mathbf{S}(\mathbf{t}_k), \ldots, \mathbf{S}(\mathbf{t}_k)}_{m_k-\text{terms}} \right)$$

Combining the class formula with recursive multiplication principles with respect to the height parameter, one obtains the following closed formula.

Theorem 5.2 *The number $\#(\mathbf{f})$ of jungles in $\mathbf{f} \in \mathcal{F}_{q,n}$, viewed as an equivalence class, is given by*

$$\#(\mathbf{f}) = (q!)^{n+2} / \prod_{i=-1}^{n} \mathbf{S}(B^i(\mathbf{f}))!$$

To get one step further, we notice that the measures $\Delta_{n,q}^{\mathbf{a}}$ have the invariance property $\Delta_{n,q}^{\mathbf{b}}(F) = \Delta_{n,q}^{\mathbf{s}(\mathbf{b})}(F)$, for any symmetric function $F \in \mathcal{B}_b(E_n^q)$. Thus, we can define unambiguously $\Delta_{n,q}^{\mathbf{f}}(F) = \Delta_{n,q}^{\mathbf{a}}(F)$, and $|\mathbf{f}| = |\mathbf{a}|$, for any choice $\mathbf{a}$ of a representative of some $\mathbf{f} \in \mathcal{F}_{q,n}$ in $\mathcal{A}_{q,n} \cong \mathcal{J}_{q,n}$.

The difference $(\mathbf{p} - \mathbf{l})$, of a pair of integer sequences $\mathbf{p} = (p_k)_{0 \leq k \leq n}$ and $\mathbf{l} = (l_k)_{0 \leq k \leq n}$, is the sequence $(\mathbf{p} - \mathbf{l}) := (p_k - l_k)_{0 \leq k \leq n}$. When no confusions can arise, we write $\mathbf{q}$, for the constant sequence $(\mathbf{q})_{0 \leq i \leq n}$. For any multi index $\mathbf{p} = (p_k)_{0 \leq k \leq n} \leq \mathbf{q}$, we let $\mathcal{F}_{n,q}(\mathbf{p}) \subset \mathcal{F}_{n,q}$ be the subset of forets with at least p_k leaves at each level k, with $0 \leq k \leq n$. We also write for any pair of multi indexes $\mathbf{p} \leq \mathbf{l}$

$$|\mathbf{p}| := (p_0 + ... + p_n) \qquad \mathbf{p}! = \prod_{k=0}^{n} p_k! \quad \text{and} \quad s(\mathbf{l}, \mathbf{p}) = \prod_{k=0}^{n} s(l_k, p_k)$$

where the $s(l_k, p_k)$ are Stirling numbers of the first kind. The next theorem is the main result of this section.

Theorem 5.3 *For any $1 \leq q \leq N$, we have the polynomial expansion (5.5), with the collection of signed, and weak derivative measures $\partial^k \mathbb{Q}_{n,q}$ given by the formula*

$$\partial^k \mathbb{Q}_{n,q} = \sum_{\mathbf{r}<\mathbf{q}:|\mathbf{r}|=k} \sum_{\mathbf{f}\in\mathcal{F}_{n,q}(\mathbf{r})} \frac{s(|\mathbf{f}|, \mathbf{q}-\mathbf{r})\ \#(\mathbf{f})}{(\mathbf{q})_{|\mathbf{f}|}}\ \Delta^{\mathbf{f}}_{n,q} \tag{5.11}$$

Let now $\mathcal{B}_0^{sym}(E_n^q)$ be the set of symmetric functions F on E_n^q such that

$$\int F(x_1, \ldots, x_{q-1}, x_q)\ \gamma_n(dx_q) = 0$$

We write $\mathbf{t}_k$ for the unique tree with a single coalescence at level k, and its two leaves at level $(n+1)$; and we write $\mathbf{u}_k$ for the trivial tree of height k.

Corollary 5.4 *For any even integer $q \leq N$ and any symmetric function $F \in \mathcal{B}_0^{sym}(E_n^q)$, we have*

$$\forall k < q/2 \quad \partial^k \mathbb{Q}_{n,q}(F) = 0, \quad \partial^{q/2} \mathbb{Q}_{n,q}(F) = \sum_{\mathbf{r}<\mathbf{q}, |\mathbf{r}|=\frac{q}{2}} \frac{q!}{2^{q/2}\ \mathbf{r}!}\ \Delta^{\mathbf{f_r}}_{n,q} F \tag{5.12}$$

with the forest $\mathbf{f_r} := \mathbf{t}_0^{r_0}\mathbf{u}_0^{r_0} \ldots \mathbf{t}_n^{r_n}\mathbf{u}_n^{r_n}$ associated with a multi index sequence $\mathbf{r} = (r_k)_{0\leq k\leq n} < \mathbf{q}$. For odd integers $q \leq N$, the partial derivatives are the null measure on $\mathcal{B}_0^{sym}(E_n^q)$, up to any order $k \leq \lfloor q/2 \rfloor$.

References

[1] Ch. Brouder. Runge-Kutta methods and renormalization. *Eur. Phys. J. C.* 12, 521-534 (2000).

[2] H. Carvalho, P. Del Moral, A. Monin, and G. Salut. Optimal nonlinear filtering in GPS/INS integration. *IEEE Trans. Aerosp. Electron. Syst.*, 33(3):835–850 (1997).

[3] A. Connes and D. Kreimer. Hopf algebras, renormalization and noncommutative geometry. *Commun. Math. Phys.* 199, No.1, 203-242 (1998).

[4] J.M.C. Clark, D.L. Ocone, and C. Coumarbatch. Relative entropy and error bounds for filtering of Markov processes. *Math. Control Signal Syst.*, 12(4):346–360, 1999.

[5] D. Crisan, P. Del Moral, and T.J. Lyons. Interacting particle systems approximations of the Kushner-Stratonovitch equation. *Adv. in Appl. Probab.*, 31(3):819–838, 1999.

[6] D. Crisan, J. Gaines, and T.J. Lyons. A particle approximation of the solution of the Kushner-Stratonovitch equation. *SIAM J. Appl. Math.*, 58(5):1568–1590, 1998.

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

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

[9] I. Csiszar. Sanov property, generalized i-projection and a conditional limit theorem. *Ann. Probab.*, 12(3):768–793, 1984.

[10] I. Csiszár. Eine informationstheoretische Ungleichung und ihre Anwendung auf den Beweis der Ergodizität von Markoffschen Ketten. *Magyar Tud. Akad. Mat. Kutató Int. Közl.*, 8:85–108, 1963.

[11] P. Del Moral, *Feynman-Kac formulae. Genealogical and interacting particle systems, with applications*, Springer Verlag New York, Series: Probability and its Applications (2004).

[12] P. Del Moral. A Uniform convergence theorem for the numerical solving of the non linear filtering problem. *Journal of Applied Probability*, vol. 35, 873–884 (1998).

[13] P. Del Moral. Measure valued processes and interacting particle systems. Application to non linear filtering problems. *Annals of Applied Probability*, vol. 8, no. 2, 438–495 (1998).

[14] P. Del Moral. Non linear filtering using random particles. *Theory of Probability and its Applications*, vol. 40, no. 4, 690–701 (1996).

[15] P. Del Moral. Non linear filtering: interacting particle solution. Markov Processes and Related Fields, vol. 2, no. 4, 555–580 (1996).

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

[17] P. Del Moral, A. Doucet and A. Jasra. Sequential Monte Carlo for Bayesian Computation. *Bayesian Statistics 8*, Oxford University Press (2006).

[18] P. Del Moral, A. Doucet and A. Jasra. Sequential Monte Carlo Samplers. *Journal of the Royal Statistical Society, Series B*, vol. 68, no. 3, pp. 411-436 (2006).

[19] P. Del Moral, A. Doucet, and G.W. Peters. Sharp Propagation of Chaos Estimates for Feynman-Kac Particle Models. *Theory of Probability and Its Applications*, vol. 51, no. 3 (2007).

[20] P. Del Moral and A. Guionnet. A central limit theorem for nonlinear filtering using interacting particle systems. *Ann. Appl. Probab.*, 9(2):275–297 (1999).

[21] P. Del Moral and A. Guionnet. deviations for interacting particle systems. Applications to non linear filtering problems. *Stochastic Processes and their Applications*, no. 78, 69–95, (1998).

[22] P. Del Moral, J. Jacod. The Monte-Carlo Method for filtering with discrete time observations. Central Limit Theorems. *The Fields Institute Communications*, Numerical Methods and Stochastics, Ed. T.J. Lyons, T.S. Salisbury, American Mathematical Society, 2002.

[23] P. Del Moral, J. Jacod. Interacting particle filtering with discrete-time observations : asymptotic behaviour in the Gaussian case. *Stochastics in finite and infinite dimensions. Trends in mathematics*, Birkhäuser, pp.101-123 (2000).

[24] P. Del Moral, M. Ledoux, and L. Miclo. On contraction properties of Markov kernels. *Probab. Theory Related Fields*, 126:395–420 (2003).

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

[26] P. Del Moral, L. Miclo. Strong Propagations of Chaos in Moran's type Particle interpretations of Feynman-Kac measures. *Stochastic Analysis and Applications*, to appear (2007).

[27] P. Del Moral and L. Miclo. Branching and interacting particle systems approximations of Feynman-Kac formulae with applications to nonlinear 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.

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

[29] P. Del Moral and L. Miclo. Particle approximations of Lyapunov exponents connected to Schrödinger operators and Feynman-Kac semigroups. *ESAIM: Probability and Statistics*, no. 7, pp. 171–208 (2003).

[30] P. Del Moral, J.C. Noyer, and G. Salut. Résolution particulaire et traitement non-linéaire du signal : application radar/sonar. In *Traitement du signal*, (12):4, 287-301 (1995).

[31] P. Del Moral, F. Patras, S. Rubenthaler, Coalescent Tree Based Functional Representations for some Feynman-Kac Particle Models, `https://hal.ccsd.cnrs.fr/ccsd-00086532` (2006).

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

[33] P. Del Moral, G. Salut. Non linear filtering using Monte Carlo particle methods. *C.R. Académie des Sciences, Paris*, t. 320, Série I, 1147–1152 (1995).

[34] A. Doucet, N. de Freitas, and N. Gordon, editors. *Sequential Monte Carlo Methods in Pratice.* Statistics for Engineering and Information Science. Springer New York, 2001.

[35] N.J. Gordon, D.J. Salmon, and C. Ewing. Bayesian state estimation for tracking and guidance using the bootstrap filter. *J. Guidance Control Dyn.*, 18(6):1434–1443, 1995.

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

[37] A. Guionnet and E. Segala Maurel, *Combinatorial aspects of matrix models*, arXiv:math.PR/0503064 v2, June 2005.

[38] G. Kallianpur and C. Striebel. Stochastic differential equations occurring in the estimation of continuous parameter stochastic processes. Tech. Rep. 103, Department of Statistics, University of Minnesota, Minneapolis, 1967.

[39] G. Kitagawa. Monte-Carlo filter and smoother for non-Gaussian nonlinear state space models. *J. Comput. and Graphical Stat.*, 5(1):1–25 (1996).

[40] V.N. Kolokoltsov and V.P. Maslov. *Idempotent Analysis and Its Applications*, volume 401 of Mathematics and its Applications. Kluwer Academic Publishers Group, Dordrecht, 1997. Translation of *Idempotent Analysis and Its Application in optimal control* (Russian), Nauka Moscow, 1994, with an appendix by P. Del Moral.

[41] D. Ocone. Entropy inequalities and entropy dynamics in nonlinear filtering of diffusion processes. In *Stochastic Analysis, Control, Optimization and Applications*, Systems Control Foundations and Applications, pages 477–496. Birkhäuser, Boston, 1999.

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

[43] T. Shiga and H. Tanaka. Central limit theorem for a system of Markovian particles with mean field interaction. *Z. Wahrschein. Verwandte Gebiete*, 69:439–459, 1985.

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

Centre de recherche INRIA Bordeaux – Sud Ouest
Domaine Universitaire - 351, cours de la Libération - 33405 Talence Cedex (France)

Centre de recherche INRIA Grenoble – Rhône-Alpes : 655, avenue de l'Europe - 38334 Montbonnot Saint-Ismier
Centre de recherche INRIA Lille – Nord Europe : Parc Scientifique de la Haute Borne - 40, avenue Halley - 59650 Villeneuve d'Ascq
Centre de recherche INRIA Nancy – Grand Est : LORIA, Technopôle de Nancy-Brabois - Campus scientifique
615, rue du Jardin Botanique - BP 101 - 54602 Villers-lès-Nancy Cedex
Centre de recherche INRIA Paris – Rocquencourt : Domaine de Voluceau - Rocquencourt - BP 105 - 78153 Le Chesnay Cedex
Centre de recherche INRIA Rennes – Bretagne Atlantique : IRISA, Campus universitaire de Beaulieu - 35042 Rennes Cedex
Centre de recherche INRIA Saclay – Île-de-France : Parc Orsay Université - ZAC des Vignes : 4, rue Jacques Monod - 91893 Orsay Cedex
Centre de recherche INRIA Sophia Antipolis – Méditerranée : 2004, route des Lucioles - BP 93 - 06902 Sophia Antipolis Cedex

Éditeur
INRIA - Domaine de Voluceau - Rocquencourt, BP 105 - 78153 Le Chesnay Cedex (France)
http://www.inria.fr
ISSN 0249-6399