Concentration Inequalities for Mean Field Particle Models

Pierre del Moral, Emmanuel Rio

INSTITUT NATIONAL DE RECHERCHE EN INFORMATIQUE ET EN AUTOMATIQUE

Concentration Inequalities for Mean Field Particle Models

Pierre Del Moral — Emmanuel Rio

N° 6901

April 2009

Thème NUM

Rapport de recherche

ISSN 0249-6399 ISRN INRIA/RR--6901--FR+ENG

INSTITUT NATIONAL DE RECHERCHE EN INFORMATIQUE ET EN AUTOMATIQUE

INRIA

centre de recherche BORDEAUX - SUD OUEST

Concentration Inequalities for Mean Field Particle Models

Pierre Del Moral*, Emmanuel Rio†

Thème NUM — Systèmes numériques
Équipes-Projets ALEA

Rapport de recherche n° 6901 — April 2009 — 29 pages

Abstract: This article is concerned with the fluctuations and the concentration properties of a general class of discrete generation and mean field particle interpretations of non linear measure valued processes.

We combine an original stochastic perturbation analysis with a concentration analysis for triangular arrays of conditionally independent random sequences, which may be of independent interest. Under some additional stability properties of the limiting measure valued processes, uniform concentration properties with respect to the time parameter are also derived. The concentration inequalities presented here generalize the classical Hoeffding, Bernstein and Bennett inequalities for independent random sequences to interacting particle systems, yielding very new results for this class of models.

We illustrate these results in the context of McKean Vlasov type diffusion models, McKean collision type models of gases, and of a class of Feynman-Kac distribution flows arising in stochastic engineering sciences and in molecular chemistry.

Key-words: Concentration inequalities, mean field particle models, measure valued processes, Feynman-Kac semigroups, McKean Vlasov models.

* 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

† Centre INRIA Bordeaux et Sud-Ouest & LMV Université de Versailles - Bâtiment Fermat, 45 Av. des Etats-Unis, 78035 Versailles Cedex France, rio@math.uvsq.fr

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

Inégalités de concentration pour des modèles particulaires de champ moyen

Résumé : Nous analysons dans cet article les fluctuations et les propriétés de concentration d'une classe générale de systèmes de particules en interaction de type champ moyen et à temps discret. Ces modèles probabilistes sont liés à des interprétations particulaires de processus à valeurs mesures non linéaires.

Nous développons une analyse originale fondée sur des techniques de perturbation stochastique de semigroupes non linéaires et sur un théorème de fluctuations de tableaux triangulaires de variables conditionnellement indépendantes. Dans certaines conditions de stabilité des semigroupes associés au processus limite, nous présentons des inégalités de concentration uniformes par rapport au paramètre temporel. Les inégalités de concentration développées dans cette étude sont des extensions des inégalités classiques de Hoeffding, Bernstein et de Bennett dans le cadre des sequence de variables indépendantes, à des systèmes de particules en interaction. Ces résultats semblent être les premiers de ce type pour ces classes de processus en interaction.

Nous illustrons ces propriétés de concentration dans le cadre de modèles diffusifs de type McKean Vlasov, pour des modèles de collisions de type McKean issus de la mécanique des fluides, ainsi que pour une classe de modèles de Feynman-Kac utilisés en ingénierie stochastique et en chimie moléculaire.

Mots-clés : Inégalités de concentration, modèles particulaires de champ moyen, processus à valeurs mesures, semigroupes de Feynman-Kac, modèles de McKean Vlasov

1 Introduction

1.1 Mean Field Particle Models

Let $(E_n)_{n\geq 0}$ be a sequence of measurable spaces equipped with some σ-fields $(\mathcal{E}_n)_{n\geq 0}$, and we let $\mathcal{P}(E_n)$ be the set of all probability measures over the set E_n, with $n \geq 0$. We consider a collection of transformations $\Phi_n : \mathcal{P}(E_{n-1}) \to \mathcal{P}(E_n)$ and we denote by $(\eta_n)_{n\geq 0}$ a sequence of probability measures on E_n satisfying a nonlinear equation of the following form

$$\eta_{n+1} = \Phi_n(\eta_{n-1}) \,. \tag{1.1}$$

The mean field type interacting particle system associated with the equation (1.1) relies on the fact that the one step mappings can be rewritten in the following form

$$\Phi_n(\eta_{n-1}) = \eta_n K_{n+1,\eta_n} \tag{1.2}$$

for some collection of Markov kernels K_{n+1,μ_n} indexed by the time parameter n and the set of measures μ_n on the space E_n. We already mention that the choice of the Markov transitions $K_{n,\eta}$ is not unique. In the literature on mean field particle models, $K_{n,\eta}$ are called a choice of McKean transitions. These models provide a natural interpretation of the distribution laws η_n as the laws of a non linear Markov chain whose elementary transitions depends on the current occupation measure. For a thorough description of these discrete generation and non linear McKean type models, we refer the reader to [2]. In the further development of the article, we always assume that a the mappings

$$\left(x_n^i\right)_{1\leq i\leq N} \in E_n^N \mapsto K_{n+1,\frac{1}{N}\sum_{j=1}^N \delta_{x_n^j}}\left(x_n^i, A_{n+1}\right)$$

are $\mathcal{E}_n^{\otimes N}$-measurable, for any $n \geq 0$, $N \geq 1$, and $1 \leq i \leq N$, and any measurable subset $A_{n+1} \subset E_{n+1}$. In this situation, 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

$$\mathbb{P}\left(\xi_{n+1}^{(N)} \in dx \;\middle|\; \mathcal{F}_n^{(N)}\right) = \prod_{i=1}^N K_{n+1,\eta_n^N}(\xi_n^{(N,i)}, dx^i) \quad \text{with} \quad \eta_n^N := \frac{1}{N}\sum_{j=1}^N \delta_{\xi_n^{(N,j)}} \,. \tag{1.3}$$

In the above displayed formula, $\mathcal{F}_n^N$ stands for the σ-field generated by the random sequence $(\xi_p^{(N)})_{0\leq p\leq n}$, and $dx = dx^1 \times \ldots \times dx^N$ stands for an infinitesimal neighborhood of a point $x = (x^1, \ldots, x^N) \in E_n^N$. The initial system $\xi_0^{(N)}$ consists of N independent and identically distributed random variables with common law η_0. As usual, to simplify the presentation, when there is no possible confusion we suppress the parameter N, so that we write ξ_n and ξ_n^i instead of $\xi_n^{(N)}$ and $\xi_n^{(N,i)}$. The state components of this Markov chain are called particles or sometimes walkers in physics to distinguish the stochastic sampling model with the physical particle in molecular models.

The rationale behind this is that η_{n+1}^N is the empirical measure associated with N independent variables with distributions $K_{n+1,\eta_n^N}\left(\xi_n^i, dx\right)$, 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}. Roughly speaking, this induction argument shows that η_n^N tends to η_n, as the population size N tends to infinity.

During the last two decades, the mean field particle interpretations of these discrete generation measure valued equations is increasingly identified as a powerful stochastic simulation algorithm with emerging subjects in physics, biology and engineering sciences. They have led to spectacular results in signal processing processing with the corresponding particle filter technology, in stochastic engineering with interacting type Metropolis and Gibbs sampler methods, as well as in quantum chemistry with quantum and diffusion Monte Carlo algorithms leading to precise estimates of the top eigenvalues and the ground states of Schroedinger operators. For a thorough discussion on these application areas, we refer the reader to [2, 3, 4], and the references therein. To motivate the article, we illustrate the fluctuation and the concentration results presented in this work with three illustrative examples, including Feynman-Kac models, McKean Vlasov diffusion type models, as well as interacting jump type McKean model of gases.

1.2 Description of the main results

The mathematical and numerical analysis of the mean field particle models (1.3) is one of the most attractive research area in pure and applied probability, as well as in advanced stochastic engineering and computational physics. For a rather exhaustive list of pointers, we refer the reader to the pair of books [2, 4]. The fluctuation analysis of these discrete generation particle models around their limiting distributions is often restricted to Feynman-Kac type models (see for instance [2], and references therein) or specific continuous time mean field models including McKean-Vlasov diffusions and Boltzmann type collision model of gases [6, 9].

In the present article, firstly we design an original stochastic perturbation analysis that applies to a large class of models satisfying a rather weak first order regularity property. To describe with some precision this first main result we observe that the local sampling errors associated with the corresponding mean field particle modelare expressed in terms of the centered random fields W_n^N, given by the following stochastic perturbation formulae:

$$\eta_n^N = \eta_{n-1}^N K_{n,\eta_{n-1}^N} + \frac{1}{\sqrt{N}}\, W_n^N\,. \tag{1.4}$$

To analyze the propagation properties of these local sampling errors, *up to a second order remainder measure*, we further assume that the one step mappings Φ_n governing the equation (1.1) have a first order decomposition

$$\Phi_n(\eta) - \Phi_n(\mu) \simeq (\eta - \mu) D_\mu \Phi_n \tag{1.5}$$

with a first order integral operator $D_\mu\Phi_n$ from $\mathcal{B}(E_n)$ into $\mathcal{B}(E_{n-1})$. Our first main result is a functional central limit theorem for the random fields

$$V_n^N := \sqrt{N}\ \left[\eta_n^N - \eta_n\right]\,. \tag{1.6}$$

This fluctuation theorem takes basically the following form.

Theorem 1.1 • *The sequence $(W_n^N)_{n\geq 0}$ converges in law, as N tends to infinity, to the sequence of n independent, Gaussian and centered random fields $(W_n)_{n\geq 0}$ with a covariance function that can be explicitly defined in terms of the McKean transitions.*

- *For any fixed time horizon $n \geq 0$, the sequence of random fields V_n^N converges in law, as the number of particles N tends to infinity, to a Gaussian and centered random fields $V_n = \sum_{p=0}^{n} W_p \mathcal{D}_{p,n}$ In the above display, $\mathcal{D}_{p,n}$ stands for the semigroup associated with the operator $\mathcal{D}_n = D_{\eta_{n-1}}\Phi_n$.*

The precise definition of the first order regularity property (1.5) and a more precise description of the above fluctuation theorem is provided in Section 3.1.

The second part of this article is concerned with the concentration properties of mean field particle models. These results quantify exponentially small probabilities of deviations events between the occupation measures η_n^N and their limiting values. Besides the fact that the non asymptotic analysis of weakly dependent type variables is rather well developed, the concentration properties of discrete generation and interacting particle systems often resume to asymptotic large deviation type results, or to non asymptotic rough exponential estimates (see for instance [2], and references therein). Our main result in this subject is an original concentration theorem that includes Hoeffding, Bennett and Bernstein exponential inequalities for mean field particle models. This result takes basically the following form.

Theorem 1.2 *For any $N \geq 1$, any $n \geq 0$, we set $S_n^N = N\left[\eta_n^N - \eta_n\right]$. Then, for any $x \geq 0$ the probability of each of the following pair of events*

$$S_n^N(f) \leq r_n\left(1+\epsilon_0^{-1}(x)\right) + N d_n \epsilon_1^{-1}\left(\frac{x}{N d_n}\right)$$

and

$$S_n^N(f) \leq r_n\ \left(1+\epsilon_0^{-1}(x)\right) + d_n'\sqrt{2xN}$$

is greater than $1-e^{-x}$, with the pair of functions (ϵ_0, ϵ_1) defined below:

$$\epsilon_0(\lambda) = \frac{1}{2}\left(\lambda - \log\left(1+\lambda\right)\right)\ , \quad \epsilon_1(\lambda) = \left(1+\lambda\right)\log\left(1+\lambda\right) - \lambda \tag{1.7}$$

and with some parameters (d_n, d_n', r_n) whose values depend respectively on the amplitude of the first and second order terms in the decompotions (1.5). Under additional stability properties of the semigroup associated with the limiting model (1.2), the parameters (d_n, d_n', r_n) are uniformly bounded w.r.t. the time parameter.

A precise description of the concentration inequalities stated in Theorem 1.2 and some of their consequences is provided in Section 3.3.

The outline of the rest of the article is as follows. To motivate the present article, we have collected in Section 2 three different classes of mean field particle models that can be studied using the fluctuation and the concentration analysis developed in this article. Section 3 is mainly concerned with the precise statement of the two main results of this article. In a first section, Section 3.1, we discuss the main regularity properties used in our analysis. Sections 3.2 and 3.3 provide a precise description of the fluctuation and the concentration theorems stated above. Section 4 is mainly concerned with the detailed proofs of the theorems stated above. We combine a natural stochastic perturbation analysis with non linear semigroup techniques that allow to describe both the fluctuations and the concentration of the mean field measures in terms of the local error random field models introduced in (1.4). The functional central limit theorem is proved in Section 4.1. In Section 5.6, we provide a preliminary convex analysis including estimates of inverses of Legendre-Fenchel transformations of classical convex functions needed in this article. In Section 4.2, we prove a technical concentration lemma for triangular arrays of conditionally independent random variables. In Section 4.3, we apply this lemma to prove concentration inequalities for mean field models.

We end this introduction with some more or less traditional notation used in the present article. We denote respectively by $\mathcal{M}(E)$, $\mathcal{M}_0(E)$, and $\mathcal{B}(E)$, the set of all finite signed measures on some measurable space $(E, \mathcal{E})$, the convex subset of measures with null mass, and the Banach space of all bounded and measurable functions f equipped with the uniform norm $\|f\|$. We also denote by $\mathrm{Osc}_1(E)$, the convex set of $\mathcal{E}$-measurable functions f with oscillations $\mathrm{osc}(f) \leq 1$. We let $\mu(f) = \int \mu(dx)\ f(x)$, be the Lebesgue integral of a function $f \in \mathcal{B}(E)$, with respect to a measure $\mu \in \mathcal{M}(E)$. We recall that a bounded integral operator M from a measurable space $(E, \mathcal{E})$ into an auxiliary measurable space $(F, \mathcal{F})$ is an operator $f \mapsto M(f)$ from $\mathcal{B}(F)$ into $\mathcal{B}(E)$ such that the functions $M(f)(x) := \int_F M(x, dy)f(y)$ are $\mathcal{E}$-measurable and bounded, for any $f \in \mathcal{B}(F)$. A Markov kernel is a positive and bounded integral operator M with $M(1) = 1$. Given a pair of bounded integral operators (M_1, M_2), we let (M_1M_2) the composition operator defined by $(M_1M_2)(f) = M_1(M_2(f))$. For time homogenous state spaces, we denote by $M^m = M^{m-1}M = MM^{m-1}$ the m-th composition of a given bounded integral operator M, with $m \geq 1$. A bounded integral operator M from a measurable space $(E, \mathcal{E})$ into an auxiliary measurable space $(F, \mathcal{F})$ also generates a dual operator $\mu \mapsto \mu M$ from $\mathcal{M}(E)$ into $\mathcal{M}(F)$ defined by $(\mu M)(f) := \mu(M(f))$. We also used the notation

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

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

2 Some illustrative examples

2.1 Feynman-Kac models

The first prototype model we have in mind is a class of Feynman-Kac distribution flow equation arising in a variety of application areas including in stochastic engineering, physics, biology and Bayesian statistics. For a thorough discussion on these application domains, we refer the reader to the book [2] and references therein. These models are defined in terms of a series of bounded and positive integral operators Q_n from E_{n-1} into E_n with the following dynamical equation

$$\forall f_n \in \mathcal{B}(E_n), \qquad \eta_n(f_n) = \eta_{n-1}(Q_n(f_n))/\eta_{n-1}(Q_n(1)) \tag{2.1}$$

with a given initial distribution $\eta_0 \in \mathcal{P}(E_0)$. To avoid unnecessary technical discussions we simplify the analysis and we assume that

$$\forall n \geq 0, \qquad 0 < \inf_{x \in E_n} G_n(x) \leq \sup_{x \in E_n} G_n(x) < \infty \quad \text{with} \quad G_n(x) := Q_{n+1}(1)(x).$$

Rewritten in a slightly different way, we have

$$\eta_n = \Phi_n(\eta_{n-1}) := \Psi_{n-1}(\eta_{n-1}) M_n \quad \text{with} \quad M_n(f_n) = Q_n(f_n)/Q_n(1)$$

and the Boltzmann-Gibbs transformation Ψ_n from $\mathcal{P}(E_n)$ into itself given by

$$\forall f_n \in \mathcal{B}(E_n), \qquad \Psi_n(\eta_n)(f_n) = \eta_n(G_n f_n)/\eta_n(G_n)\,.$$

We leave the reader to check that this flow of measures satisfy the recursive equation (1.1) for any choice of Markov transitions given below

$$K_{n+1,\eta_n}(x, dy) = \epsilon_n G_n(x)\ M_n(x, dy) + (1 - \epsilon_n G_n(x))\ \Phi_{n+1}(\eta_n)(dy)\,. \tag{2.2}$$

In the above displayed formula ϵ_n stands for some $[0,1]$-valued parameters that may depend on the current measure η_n and such that $\|\epsilon_n G_n\| \leq 1$. In this situation, the mean field N-particle model associated with the collection of Markov transitions (2.2) is a combination of simple selection/mutation genetic transition $\xi_n \rightsquigarrow \widehat{\xi}_n = (\widehat{\xi}^i_n)_{1 \leq i \leq N} \rightsquigarrow \xi_{n+1}$. During the selection stage, with probability $\epsilon_n G_n(\xi^i_n)$, we set $\widehat{\xi}^i_n = \xi^i_n$; otherwise, the particle jumps to a new location, randomly drawn from the discrete distribution $\Psi_n(\eta^N_n)$. During the mutation stage, each of the selected particles $\widehat{\xi}^i_n \rightsquigarrow \xi^i_{n+1}$ evolves according to the transition M_{n+1}.

2.2 Gaussian mean field models

The concentration analysis presented in this article is not restricted to Feynman-Kac type models. It also applies to McKean type models associated with a collection of multivariate Gaussian type Markov transitions on $E_n = \mathbb{R}^d$, defined by

$$\begin{aligned} &K_{n,\eta}(x, dy) \\ &= \frac{1}{\sqrt{(2\pi)^d \det(Q_n)}}\ \exp\left\{-\tfrac{1}{2}\,(y - d_n(x,\eta))'\ Q_n^{-1}\,(y - d_n(x,\eta))\right\}\ dy\,, \end{aligned} \tag{2.3}$$

with a non singular, positive and semi-definite covariance matrix Q_n and some sufficiently regular drift mapping $d_n\ :\ (x,\eta) \in \mathbb{R}^d \times \mathcal{P}(\mathbb{R}^d) \mapsto d(x,\eta) \in \mathbb{R}^d$. In this context, the N-mean field particle model is given by the following recursion:

$$\forall 1 \leq i \leq N \qquad \xi^i_n = d_n\left(\xi^i_{n-1}, \eta^N_{n-1}\right) + W^i_n\,,$$

where $(W^i_n)_{i \geq 0}$ is a collection of independent and identically distributed d-valued Gaussian random variables with covariance matrix Q_n.

2.3 A McKean model of gases

We end this Section with a mean field particle model arising in fluid mechanics. We consider a measurable state space $(S_n, \mathcal{S}_n)$ with a countably generated σ-field and an $(\mathcal{S}_n \otimes \mathcal{E}_n)$-measurable mapping a_n be a from $(S_n \times E_n)$ into $\mathbb{R}_+$ such that $\int \nu_n(ds) a_n(s,x) = 1$, for any $x \in E_n$, and some bounded positive measure $\nu_n \in \mathcal{M}(S_n)$. To illustrate this model, we can take a partition of the state $E_n = \cup_{s \in S_n} A_s$ associated with a countable set S_n equipped with the counting measure $\nu_n(s) = 1$, and set $a_n(s,x) = 1_{A_s}(x)$. We let $K_{n+1,\eta}$ be the McKean transition defined by

$$K_{n+1,\eta}(x, dy) = \int \nu_n(ds)\ \eta(du)\ a_n(s,u)\ M_{n+1}((s,x), dy)\,. \tag{2.4}$$

In the above displayed formula, M_n stands for some Markov transition from $(S_n \times E_n)$ into E_{n+1}. The discrete time version f the McKean's 2-velocities model for Maxwellian gases correspond to the time homogenous model on $E_n = S_n = \{-1, +1\}$ associated with the counting measure ν_n and the pair of parameters

$$a_n(s,x) = 1_s(x) \quad \text{and} \quad M_{n+1}((s,x), dy) = \delta_{sx}(dy)\,.$$

In this situation, the measure valued equation (1.1) takes the following quadratic form:

$$\eta_{n+1}(+1) = \eta_n(+1)^2 + (1 - \eta_n(+1))^2\ .$$

The leave the reader to write out the mean field particle interpretation of this model. For more details on this model, we refer to [9].

3 Fluctuations and concentration properties

3.1 Some weak regularity properties

To describe precisely the concentration inequalities developed in the article, we need to introduce a first round of notation.

Definition 3.1 *When the bounded integral operator M has a constant mass, that is, when $M(1)(x) = M(1)(y)$ for any $(x,y) \in E^2$, the operator $\mu \mapsto \mu M$ maps $\mathcal{M}_0(E)$ into $\mathcal{M}_0(F)$. In this situation, we let $\beta(M)$ be the Dobrushin coefficient of a bounded integral operator M defined by the formula* $\beta(M) := \sup\ \{\mathrm{osc}(M(f))\ ;\ \ f \in \mathrm{Osc}_1(F)\}$.

Definition 3.2 *We let $\Upsilon(E,F)$ be the set of mappings*

$$\Phi\ :\ \mu \in \mathcal{P}(E) \mapsto \Phi(\mu) \in \mathcal{P}(F)$$

satisfying the first order decomposition

$$\Phi(\mu) - \Phi(\eta) = (\mu - \eta) D_\eta \Phi + \mathcal{R}^\Phi(\mu, \eta)\,. \tag{3.1}$$

In the above displayed formula, the first order operators $(\mathcal{D}_\eta \Phi)_{\eta \in \mathcal{P}(E)}$ is some collection of bounded integral operators from E into F such that

$$\begin{aligned} &\forall \eta \in \mathcal{P}(E) \quad \forall x \in E \\ &(D_\eta \Phi)(1)(x) = 0 \quad \text{ and } \quad \beta(\mathcal{D}\Phi) := \sup_{\eta \in \mathcal{P}(E)} \beta(D_\eta \Phi) < \infty\,. \end{aligned} \tag{3.2}$$

The collection of second order remainder signed mesures $(\mathcal{R}^{\Phi}(\mu,\eta))_{(\mu,\eta)\in\mathcal{P}(E^2)}$ *on* F *are such that*

$$\left|\mathcal{R}^{\Phi}(\mu,\eta)(f)\right| \leq \int \left|(\mu-\eta)^{\otimes 2}(g)\right| \; R^{\Phi}_{\eta}(f,dg)\,, \tag{3.3}$$

for some collection of integral operators R^{Φ}_{η} *from* $\mathcal{B}(F)$ *into the set* $\mathrm{Osc}_1(E)^2$ *such that*

$$\sup_{\eta\in\mathcal{P}(E)} \int \mathrm{osc}(g_1)\;\mathrm{osc}(g_2)\; R^{\Phi}_{\eta}(f,d(g_1\otimes g_2)) \leq \;\mathrm{osc}(f)\;\delta\left(R^{\Phi}\right) \tag{3.4}$$

with $\delta\left(R^{\Phi}\right)<\infty$.

This rather weak first order regularity property is satisfied for a large class of one step transformation Φ_n associated with a non linear measure valued process (1.1). For instance, in Section 4.3 we shall prove that the Feynman-Kac transformations Φ_n introduced in (2.1) belong to the set $\Upsilon(E_{n-1},E_n)$. The latter is also met for the Gaussian transitions introduced in (2.3) as soon as the drift function $d(x,\eta)$ is sufficiently regular. For instance, this condition is met for $d(x,\eta)=b(x)\eta(a)$, as well as for $d(x,\eta)=b(x)-\eta(a)$, with any pair (a,b) of bounded functions. This condition is also met for the McKean type model of gases (2.4) presented in Section 2.3. The proof of this assertion is rather technical and it is postponed in Section 5.4.

3.2 A functional central limit theorem

We assume that the one step mappings

$$\Phi_n \;:\; \mu\in\mathcal{P}(E_{n-1}) \longrightarrow \Phi_n(\mu):=\mu K_{n,\mu}\in\mathcal{P}(E_n)$$

governing the equation (1.1) are chosen so that $\Phi_n\in\Upsilon(E_{n-1},E_n)$, for any $n\geq 1$. We also let $\Phi_{p,n}$, $0\leq p\leq n$, be the semigroup associated with the measure valued equation defined in (1.1)

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

For $p=n$, we use the convention $\Phi_{n,n}=Id$, the identity operator. The main advantage of the regularity condition comes from the fact that $\Phi_{p,n}\in\Upsilon(E_p,E_n)$ with the first order decomposition type formula

$$\Phi_{p,n}(\eta)-\Phi_{p,n}(\mu)=[\eta-\mu]D_{\mu}\Phi_{p,n}+\mathcal{R}^{\Phi_{p,n}}(\eta,\mu)$$

for some collection of bounded integral operators $D_{\mu}\Phi_{p,n}$ from E_p into E_n and some second order remainder signed mesures $\mathcal{R}^{\Phi_{p,n}}(\eta,\mu)$. A proof of this assertion can be found in the appendix.

Under some appropriate regularity properties on the McKean transitions $K_{n,\eta}$, the fluctuation of the occupation measures η^N_n around their limiting values η_n are described by the following theorem.

Theorem 3.3 *For any fixed time horizon* $n\geq 0$*, the sequence of random fields* V^N_n *introduced in (1.6) converges in law, as the number of particles* N *tends to infinity, to a Gaussian and centered random field* V_n *given by*

$$V_n=\sum_{p=0}^{n} W_p D_{\eta_p}\Phi_{p,n}\,. \tag{3.5}$$

In the above display, $(W_n)_{n\geq 0}$ stands for a sequence of independent, Gaussian and centered random fields ; with, for any $f, g \in \mathcal{B}(E_n)$, and $n \geq 0$,

$$\mathbb{E}(W_n(f)W_n(g)) = \eta_{n-1}K_{n,\eta_{n-1}}([f - K_{n,\eta_{n-1}}(f)][g - K_{n,\eta_{n-1}}(g)])\,. \tag{3.6}$$

A complete detailed proof of the functional central limit theorem stated above is provided in Section 4, dedicated to a stochastic perturbation analysis of mean field particle models. To get one step further, we examine the variance of the the limiting random fields (3.5). Firstly, sor any $f_n \in \mathrm{Osc}_1(E_n)$, one observes that

$$\mathbb{E}(V_n(f_n)^2) \;=\; \sum_{p=0}^{n} \mathbb{E}\left(\left(W_p\left[D_{\eta_p}\Phi_{p,n}(f_n)\right]\right)^2\right) \leq \sum_{p=0}^{n} \sigma_p^2\; \beta(D\Phi_{p,n})^2 \tag{3.7}$$

with the uniform local variance parameters:

$$\sigma_n^2 := \sup_{f_n\in\mathrm{Osc}_1(E_n)}\; \sup_{\mu\in\mathcal{P}(E_{n-1})} \left|\mu\left(K_{n,\mu}\left[f_n - K_{n,\mu}(f_n)\right]^2\right)\right| \;(\leq 1)\;.$$

3.3 Concentration inequalities

The concentration inequalities discussed in this article are expressed in terms of the pair of expansion parameters defined below.

Definition 3.4 *We introduce the first order expansion parameters $(\overline{\sigma}_n, \beta_n, b_n^\star)$ are given by*

$$\overline{\sigma}_n^2 = \sum_{p=0}^{n} \sigma_p^2\; \beta(D\Phi_{p,n})^2 \leq \beta_n^2 = \sum_{p=0}^{n} \beta(D\Phi_{p,n})^2 \quad \textit{and} \quad b_n^\star = \sup_{0\leq p\leq n} \beta(D\Phi_{p,n})\,.$$

Finally, we denote by r_n the second order parameter $r_n = \sum_{p=0}^{n} \delta(R^{\Phi_{p,n}})$.

Let us briefly examine some interpretations of these parameters. Firstly, we observe that the first order expansion parameter ($\overline{\sigma}_n$ is related to the variance of the limiting Gaussian field, while the parameter r_n can be thought as a second order stochastic perturbation term related to the quadratic remainder measures $R^{\Phi_{p,n}}$.

When the Markov kernels $K_{n,\mu} = K_n$ do not depend on the measure μ, the N-particle model reduce to a collection of independent copies of the Markov chain with elementary transitions $P_n = K_n$. In this special case, the second order parameters vanish (i.e. $r_n = 0$), while the first order expansion parameters $(\overline{\sigma}_n, \beta_n)$ are related to the mixing properties of the semigroup of the underlying Markov chain, that is we have that

$$\overline{\sigma}_n^2 = \sum_{p=0}^{n} \sigma_p^2\; \beta(P_{p,n})^2 \leq \beta_n^2 = \sum_{p=0}^{n} \beta(P_{p,n})^2 \quad \text{with} \quad P_{p,n} = K_{p+1}\ldots K_{n-1}K_n\,,$$

with the Dobrushin ergodic coefficient $\beta(P_{p,n})$ associated with $P_{p,n}$. When the chain is asymptotically stable in the sense that $\sup_{n\geq 0}\sum_{p=0}^{n}\beta(P_{p,n}) < \infty$, the first order expansion parameters given above are uniformly bounded with respect to the time parameter.

In more general situations, the analysis of these parameters depends on the model at hand. For instance, for time homogeneous Feynman-Kac models (i.e. $E_n = E$, and $(G_n, M_n) = (G, M)$) these parameters can be related to the mixing properties of the Markov chain associated with the transitions M. To be more precise, let us suppose that the following condition is met

$$(M)_m \quad \exists m \geq 1 \quad \exists \epsilon_m > 0 \quad \text{s.t.} \quad \forall (x,y) \in E^2 \qquad M^m(x,\cdot) \geq \epsilon_m \; M^m(y,\cdot). \tag{3.8}$$

It is well known that the mixing type condition $(M)_m$ is satisfied for any aperiodic and irreducible Markov chains on finite spaces, as well as for bi-Laplace exponential transitions associated with a bounded drift function and for Gaussian transitions with a mean drift function that is constant outside some compact domain. To go one step further, we introduce the following quantities:

$$\delta_m := \sup \prod_{0 \leq p < m} \left(G(x_p)/G(y_p) \right). \tag{3.9}$$

In the above displayed formula, the supremum is taken over all admissible pair of paths with elementary transitions M. In this situation, we can check that

$$r_n \leq 4 \; \varpi_{3,1}(m), \; b_n^\star \leq 2 \; \delta_m/\epsilon_m$$

as well as

$$\overline{\sigma}_n^2 \leq 4 \; \varpi_{2,2}(m) \; \sigma^2 \quad \text{and} \quad \beta_n^2 \leq 4 \; \varpi_{2,2}(m)$$

with the uniform local variance parameter σ^2 and a collection of parameters $\varpi_{k,l}(m)$ such that $\varpi_{k,l}(m) \leq m \; \delta_{m-1} \; \delta_m^k / \epsilon_m^{k+2}$. The detailed proof of these estimates can be found in Section 5.3. The precise statement of Theorem 1.2 is given below.

Theorem 3.5 *For any $N \geq 1$, any $n \geq 0$, and any $x \geq 0$ the probability of each of the following pair of events is greater than $1 - e^{-x}$*

$$[\eta_n^N - \eta_n](f_n) \leq \frac{r_n}{N} \; \left(1 + \epsilon_0^{-1}(x)\right) + \overline{\sigma}_n^2 \; b_n^\star \; \epsilon_1^{-1} \left(\frac{x}{N \overline{\sigma}_n^2} \right)$$

and

$$[\eta_n^N - \eta_n](f_n) \leq \frac{r_n}{N} \; \left(1 + \epsilon_0^{-1}(x)\right) + \sqrt{\frac{2x}{N}} \; \beta_n$$

with the pair of functions (ϵ_0, ϵ_1) defined in (1.7)

Let us examine some direct consequences of these concentration inequalities. As we mentioned above, in the special case where the Markov kernels $K_{n,\mu} = K_n$ do not depend on the measure μ, the random measures η_n^N coincide with the occupation measure associated with N independent and identically distributed random variables with common law η_n. In this situation, the pair of events described in Theorem 3.5 resumes to the following Bennett and Hoeffding type concentration events respectively given by

$$[\eta_n^N - \eta_n](f_n) \leq \overline{\sigma}_n^2 \; b_n^\star \; \epsilon_1^{-1} \left(\frac{x}{N \overline{\sigma}_n^2} \right) \quad \text{and} \quad [\eta_n^N - \eta_n](f_n) \leq \sqrt{\frac{2x}{N}} \; \beta_n \, .$$

The first inequality can be described more explicitly using the analytic estimates:

$$\epsilon_1^{-1}(x) \leq \frac{\sqrt{2x} + (4x/3) - \log(1 + (x/3) + \sqrt{2x})}{\log(1 + (x/3) + \sqrt{2x})} \leq (x/3) + \sqrt{2x}.$$

In the context of Feynman-Kac models, the second order terms can be estimated more explicitly using the upper bounds

$$\epsilon_0^{-1}(x) \leq 2x + \log(1 + 2x + 2\sqrt{x}) + \frac{\log(1 + 2x + 2\sqrt{x}) - 2\sqrt{x}}{2x + 2\sqrt{x}} \leq 2x + 2\sqrt{x}.$$

A detailed proof of the upper bounds given above is detailed in Section 5.6, dedicated to the convex analysis of the Legendre-Fenchel transformations used in this article. The second rough estimate in the r.h.s. of the above displayed formulae leads to Bernstein type concentration inequalities.

Corollary 3.6 *For any $N \geq 1$ and any $n \geq 0$, we have the following Bernstein type concentration inequalities*

$$\begin{aligned} & -\frac{1}{N} \log \mathbb{P}\left([\eta_n^N - \eta_n](f_n) \geq \frac{r_n}{N} + \lambda\right) \\ & \geq \frac{\lambda^2}{2} \left(\left(b_n^\star\ \overline{\sigma}_n + \frac{\sqrt{2} r_n}{\sqrt{N}} \right)^2 + \lambda \left(2r_n + \frac{b_n^\star}{3} \right) \right)^{-1} \end{aligned}$$

and

$$-\frac{1}{N} \log \mathbb{P}\left([\eta_n^N - \eta_n](f_n) \geq \frac{r_n}{N} + \lambda\right) \geq \frac{\lambda^2}{2} \left(\left(\beta_n + \frac{\sqrt{2} r_n}{\sqrt{N}} \right)^2 + 2 r_n \lambda \right)^{-1}.$$

In terms of the random fields V_n^N, the first concentration inequality stated in Corollary 3.6 takes the following form

$$\begin{aligned} & -\log \mathbb{P}\left(V_n^N(f_n) \geq \frac{r_n}{\sqrt{N}} + \lambda\right) \\ & \geq \frac{\lambda^2}{2} \left(\left(b_n^\star\ \overline{\sigma}_n + \frac{\sqrt{2} r_n}{\sqrt{N}} \right)^2 + \frac{\lambda}{\sqrt{N}} \left(2r_n + \frac{b_n^\star}{3} \right) \right)^{-1} \\ & \longrightarrow_{N \to \infty} \frac{\lambda^2}{2\,(b_n^\star\ \overline{\sigma}_n)^2}\ . \end{aligned}$$

This observation shows that this concentration inequality is "almost" asymptotically sharp, with a variance type term whose values are pretty close to the exact limiting variances presented in (3.7). A more precise asymptotic estimate would require a refined moderate deviation analysis. We hope to discuss these properties in a forthcoming study.

Last, but not least, without further work, Theorem 3.5 leads to uniform concentration inequalities for mean field particle interpretations of Feynman-Kac semigroups.

Corollary 3.7 *In the context of Feynman-Kac models, under the mixing type condition $(M)_m$ introduced in (3.8), for any $N \geq 1$, any $n \geq 0$, and any $x \geq 0$ the probability of each of the following pair of events*

$$[\eta_n^N - \eta_n](f_n)$$

$$\leq \frac{4}{N}\varpi_{3,1}(m)\left(1+\epsilon_0^{-1}(x)\right) + \frac{8\delta_m}{\epsilon_m}\ \varpi_{2,2}(m)\ \sigma^2\ \ \epsilon_1^{-1}\left(\frac{x}{4\sigma^2\varpi_{2,2}(m)\ N}\right)$$

and

$$[\eta_n^N - \eta_n](f_n) \leq \frac{4}{N}\varpi_{3,1}(m)\left(1+\epsilon_0^{-1}(x)\right) + 2\sqrt{\frac{2\varpi_{2,2}(m)x}{N}}$$

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

4 A stochastic perturbation analysis

4.1 Proof of the functional central limit theorem

Definition 4.1 *We say that a collection of Markov transitions K_η from a measurable space $(E,\mathcal{E})$ into another $(F,\mathcal{F})$ satisfies condition (K) as soon as the following Lipschitz type inequality is met for every $f \in \mathrm{Osc}_1(F)$:*

$$(K) \qquad \|\,[K_\mu - K_\eta]\,(f)\| \leq \int\ |(\mu-\eta)(h)|\ T_\eta^K(f,dh)\,. \tag{4.1}$$

In the above display, T_η^K stands for some collection of bounded integral operators from $\mathcal{B}(F)$ into $\mathcal{B}(E)$ such that

$$\sup_{\eta\in\mathcal{P}(E)}\int\ \mathrm{osc}(h)\ T_\eta^K(f,dh) \leq\ \mathrm{osc}(f)\ \delta\left(T^K\right)\ , \tag{4.2}$$

for some finite constant $\delta\left(T^\Phi\right) < \infty$. In the special case where $K_\eta(x,dy) = \Phi(\eta)(dy)$, for some mapping $\Phi\ :\ \eta\in\mathcal{P}(E)\mapsto\Phi(\eta)\in\mathcal{P}(F)$, condition (4.1) is a simple Lipschitz type condition on the mapping Φ. In this situation, we denote by (Φ) the corresponding condition; and whenever it is met, we says that the mapping Φ satisfy condition (Φ).

We further assume that we are given a collection of McKean transitions $K_{n,\eta}$ satisfying the weak Lipschitz type condition stated in (4.1). In this situation, we already mention that the corresponding one step mappings $\Phi_n(\eta) = \eta K_{n,\eta}$ and the corresponding semigroup $\Phi_{p,n}$ satisfy condition $(\Phi_{p,n})$ for some collection of bounded integral operators $T_\eta^{\Phi_{p,n}}$.

In the context of Feynman-Kac type mdels, it is not difficult to check that condition (Φ_n) is equivalent to the fact that the McKean transitions $K_{n,\eta}$ given in (2.2) satisfy the Lipschitz condition (4.1). The latter is also met for the Gaussian transitions introduced in (2.3) as soon as the drift function $d(x,\eta)$ is sufficiently regular. A before, this condition is met for $d(x,\eta) = b(x)\eta(a)$, as well as for $d(x,\eta) = b(x)-\eta(a)$, with any pair (a,b) of bounded functions. It is again satisfied for the McKean type model of gases (2.4) presented in Section 2.3. For a more detailed discussion on these stability properties, we refer the reader to the appendix, on page 20.

Notice that the centered random fields W_n^N introduced in (1.4) have conditional variance functions given by

$$\mathbb{E}(W_n^N(f_n)^2 \mid \mathcal{F}_{n-1}^N\ ) = \eta_{n-1}^N \left[K_{n,\eta_{n-1}^N}\left((f_n - K_{n,\eta_{n-1}^N}(f_n))^2\right)\right]. \qquad (4.3)$$

Using Kintchine's inequality, for every $f \in \mathrm{Osc}_1(E_n)$, $N \geq 1$ and any $n \geq 0$ and $m \geq 1$ we have the $\mathbb{L}_{2m}$ almost sure estimates

$$\mathbb{E}\left(\left|W_n^N(f_n)\right|^{2m}\left|\mathcal{F}_{n-1}^{(N)}\right.\right)^{\frac{1}{2m}} \leq b(2m) \quad \text{with} \quad b(2m)^{2m} := 2^{-m}(2m)!/m!\,. \quad (4.4)$$

We can also prove the following theorem.

Theorem 4.2 *The sequence $(W_n^N)_{n\geq 0}$ converges in law, as N tends to infinity, to the sequence of n independent, Gaussian and centered random fields $(W_n)_{n\geq 0}$ described in Theorem 3.3.*

The proof of this theorem follows the same line of arguments as those we used in [2] in the context of Feynman-Kac models. For completeness and for the convenience of the reader, the complete proof of this result is housed in Section 5.2, in the appendix.

Let us examine some direct consequences of this result. Combining the Lipschitz property $(\Phi_{p,n})$ of the semigroup $\Phi_{p,n}$ with the decomposition

$$\left[\eta_n^N - \eta_n\right] = \sum_{p=0}^{n}\ \left[\Phi_{p,n}(\eta_p^N) - \Phi_{p,n}\left(\Phi_p(\eta_{p-1}^N)\right)\right]\,,$$

we find that

$$\sqrt{N}\ \left|\left[\eta_n^N - \eta_n\right](f_n)\right| = \sum_{p=0}^{n}\int\ \left|W_p^N(h)\right|\ T^{\Phi_{p,n}}_{\Phi_p(\eta_{p-1}^N)}(f,dh)\,.$$

In the above displayed formulae, we have used the convention $\Phi_0(\eta_{-1}^N) = \eta_0$, for $p = 0$. From the previous $\mathbb{L}_{2m}$ almost sure estimates, we readily conclude that

$$\sup_{N\geq 1}\sqrt{N}\ \mathbb{E}\left(\left|\left[\eta_n^N - \eta_n\right](f_n)\right|^{2m}\right)^{\frac{1}{2m}} \leq b(2m)\ \sum_{p=0}^{n}\delta(T^{\Phi_{p,n}})\,.$$

We are now in position to prove the fluctuation Theorem 3.3. Using the decomposition

$$V_n^N \quad = \quad W_n^N + V_{n-1}^N\mathcal{D}_n + \sqrt{N}\ R^{\Phi_n}\left(\eta_{n-1}^N, \eta_{n-1}\right)\,,$$

we readily prove that

$$V_n^N = \sum_{p=0}^{n} W_p^N \mathcal{D}_{p,n} + \frac{1}{\sqrt{N}}\ \mathcal{R}_n^N\,, \qquad (4.5)$$

with the remainder second order measure

$$\mathcal{R}_n^N := N\ \sum_{p=0}^{n-1} R_{p+1}^{\Phi_{p+1}}\left(\eta_p^N, \eta_p\right) D_{p+1,n}\,.$$

In the above display, $\mathcal{D}_{p,n} = \mathcal{D}_{p+1}\ldots\mathcal{D}_{n-1}\mathcal{D}_n$ stands for the semigroup associated with the integral operators $\mathcal{D}_n := D_{\eta_{n-1}}\Phi_n$, with the usual convention $\mathcal{D}_{n,n} = Id$, for $p = n$. Using a first order derivation formula for the semi-group $\Phi_{p,n}$ (cf. for instance Lemma 5.1 on page 20), it is readily checked that

$$D_{\eta_p}\Phi_{p,n} = (D_{\eta_p}\Phi_{p+1})(D_{\eta_{p+1}}\Phi_{p+1,n}) = \mathcal{D}_{p+1}(D_{\eta_p}\Phi_{p,n}) = \mathcal{D}_{p,n}\,.$$

Using the fact that

$$\left|\mathcal{R}_n^N(f_n)\right| \leq \sum_{p=0}^{n-1}\int \left|\left(V_p^N\right)^{\otimes 2}(g)\right| \; R_{\eta_p}^{\Phi_{p+1}}(f,dg),$$

we conclude that, for any $m \geq 1$ we have

$$\mathbb{E}\left(\left|\mathcal{R}_n^N(f_n)\right|^m\right)^{1/m} \leq b(2m)^2 \sum_{p=0}^{n-1}\beta(\mathcal{D}_{p+1,n})\left(\sum_{q=0}^{p}\delta(T^{\Phi_{q,p}})\right)^2 \delta\left(R^{\Phi_{p+1}}\right)\,.$$

This clearly implies that $\frac{1}{\sqrt{N}}\,\mathcal{R}_n^N$ converge in law to the null measure, in the sense that $\frac{1}{\sqrt{N}}\,\mathcal{R}_n^N(f_n)$ converge in law to zero, for any bounded test function f_n on E_n. Using the fact that W_n^N converges in law to the sequence of n independent, random fields W_n, the proposition is now a direct consequence of the decomposition formula (4.5). This ends the proof of Theorem 3.3. ∎

4.2 A concentration lemma for triangular arrays

For every $n \geq 0$ and $N \geq 1$, we let $X_n^{(N)} := (X_n^{(N,i)})_{1\leq i\leq N}$ be a triangular array of random variables defined on some filtered probability space $(\Omega, \mathcal{F}_n^N)$ associated with a collection of increasing σ-fields $(\mathcal{F}_n^N)_{n\geq 0}$. We assume that $(X_n^{(N,i)})_{1\leq i\leq N}$ are $\mathcal{F}_{n-1}^N$-conditionally independent and centered random variables. Suppose furthermore that

$$\forall n\geq 0 \quad a_n \leq X_n^{(N,i)} \leq b_n \quad \text{and} \quad \mathbb{E}\left(\left(X_n^{(N,i)}\right)^2 \mid \mathcal{F}_{n-1}^N\right) \leq c_n^2$$

for some collection of finite constants (a_n, b_n, c_n), with the convention $\mathcal{F}_{-1}^N = \{\emptyset, \Omega\}$ for $n = 0$. For any $n \geq 0$, let

$$T_n^N := S_n^N + R_n^N, \quad \text{where} \quad \Delta S_n^N := S_n^N - S_{n-1}^N = \sum_{i=1}^N X_n^{(N,i)}$$

and R_n^N is a random perturbation term such that

$$\forall m\geq 1 \qquad \mathbb{E}\left(\left|R_n^N\right|^m\right)^{\frac{1}{m}} \leq b(2m)^2\; d_n$$

for some finite constant d_n. We use the convention $S_{-1}^N = 0$, for $n = 0$. We set

$$\overline{c}_n^2 := (b_n^\star)^{-2}\sum_{p=0}^n c_p^2 \quad \text{and} \quad \overline{\delta}_n^2 := \sum_{p=0}^n \delta_p^2 \quad \text{with the middle point} \quad \delta_n := \frac{b_n - a_n}{2}\,.$$

Lemma 4.3 *For any $N \geq 1$ and any $n \geq 0$, the probability of each of the following pair of events*

$$T_n^N \leq d_n \left(1 + (\alpha_0^\star)^{-1}(x)\right) + N\ \overline{c}_n^2\ b_n^\star\ (\alpha_1^\star)^{-1}\left(\frac{x}{N\overline{c}_n^2}\right) \tag{4.6}$$

and

$$T_n^N \leq d_n \left(1 + (\alpha_0^\star)^{-1}(x)\right) + \overline{\delta}_n\sqrt{2xN} \tag{4.7}$$

is greater than $1 - e^{-x}$, for any $x \geq 0$.

Remark 4.4 *Notice that (4.7) gives always a better concentration inequality when $\sum_{p=0}^n c_p^2 \geq \sum_{p=0}^n \delta_p^2$. In the opposite situation, if $\sum_{p=0}^n c_p^2 < \sum_{p=0}^n \delta_p^2$, inequality (4.6) gives better concentration estimates for sufficiently small values of the precision parameter x.*

Before getting into the details of the proof of the above lemma, we examine some direct consequences. Firstly, combining (5.15) with (5.14) we observe that, with probability greater than $1 - e^{-x}$,

$$T_n^N \leq d_n\left(1 + 2\sqrt{x} + \theta_0(x)\right) + b_n^\star\left(\overline{c}_n\sqrt{N}\ \sqrt{2x} + N\overline{c}_n^2\theta_1\left(\frac{x}{N\overline{c}_n^2}\right)\right)$$

with the pair of functions

$$\theta_0(x) := 2x + \log(1 + 2\sqrt{x} + 2x) - 2\sqrt{x} + \frac{\log(1 + 2\sqrt{x} + 2x)) - 2\sqrt{x}}{2x + 2\sqrt{x}} \leq 2x$$

and

$$\theta_1(x) := \frac{\sqrt{2x} + (4x/3)}{\log(1 + (x/3) + \sqrt{2x})} - 1 - \sqrt{2x} \leq \frac{x}{3}\,.$$

The upper bounds given above together with (5.7) imply that, with probability greater than $1 - e^{-x}$,

$$T_n^N \leq d_n + A_n x + \sqrt{2xB_n^N},$$

where

$$A_n := \left(2d_n + \frac{b_n^\star}{3}\right) \quad \text{and} \quad B_n^N := \left(\sqrt{2}d_n + b_n^\star\ \overline{c}_n\sqrt{N}\right)^2.$$

Using these successive upper bounds, we arrive at the following Bernstein's type inequality:

$$-\frac{1}{N}\ \log\mathbb{P}\left(\frac{T_n^N}{N} \geq \frac{d_n}{N} + \lambda\right) \geq \frac{\lambda^2}{2}\ \left(\left(b_n^\star\ \overline{c}_n + \frac{\sqrt{2}d_n}{\sqrt{N}}\right)^2 + \lambda\left(2d_n + \frac{b_n^\star}{3}\right)\right)^{-1}. \tag{4.8}$$

In much the same way, starting from (4.6), we have, with probability greater than $1 - e^{-x}$,

$$T_n^N \leq d_n\left(1 + 2(x + \sqrt{x})\right) + \overline{\delta}_n\sqrt{2xN} = d_n + A_n\ x + \sqrt{2xB_n^N}\,, \tag{4.9}$$

with the pair of constants

$$A_n := 2d_n \quad \text{and} \quad B_n^N := \left(\sqrt{2}d_n + \overline{\delta}_n\sqrt{N}\right)^2.$$

Using these successive upper bounds, we arrive at the following Bernstein's type inequality:

$$-\frac{1}{N}\ \log \mathbb{P}\left(\frac{T_n^N}{N} \geq \frac{d_n}{N} + \lambda\right) \geq\ \frac{\lambda^2}{2}\ \left(\left(\overline{\delta}_n + \frac{\sqrt{2}d_n}{\sqrt{N}}\right)^2 + 2d_n\lambda\right)^{-1}. \qquad (4.10)$$

Proof of Lemma 4.3: Firstly, we observe that

$$\forall t \in [0, 1/(2d_n)[\qquad \mathbb{E}\left(e^{tR_n^N}\right) \leq \sum_{m\geq 0} \frac{(td_n)^m}{m!}\ b(2m)^{2m}.$$

To obtain a more explicit form of the r.h.s. term, we recall that $b(2m)^{2m} = \mathbb{E}(X^{2m})$ with a Gaussian centered random variable with $\mathbb{E}(X^2) = 1$ and

$$\forall d \in [0, 1/2[\quad \mathbb{E}(\exp\left\{dX^2\right\}) = \sum_{m\geq 0} \frac{s^m}{m!}\ b(2m)^{2m} = \frac{1}{\sqrt{1-2d}}.$$

From this observation, we readily find that

$$\forall t \in [0, 1/(2d_n)[\qquad L_{0,n}^N(t) := \log \mathbb{E}\left(e^{t(R_n^N - d_n)}\right) \leq \alpha_{0,n}(t) := \alpha_0(td_n).$$

Using (5.6), we obtain the following almost sure inequality

$$\log \mathbb{E}\left(e^{t\Delta S_n^N}\ \mid \mathcal{F}_{n-1}^N\right) \leq N\ \left(\frac{c_n}{b_n}\right)^2\ \alpha_1\left(b_n t\right).$$

It implies that

$$\forall t \geq 0 \qquad L_{1,n}^N(t) := \log \mathbb{E}\left(e^{tS_n^N}\right) \leq N\ \sum_{p=0}^{n} \left(\frac{c_p}{b_p}\right)^2\ \alpha_1\left(b_p t\right) \leq \alpha_{1,n}^N(t),$$

with the increasing and convex function $\alpha_{1,n}^N(t) = N\ \overline{c}_n^2\ \alpha_1\left(b_n^\star t\right)$, which completes the proof of the lemma. ∎

Using (5.7), we now obtain the following Cramer-Chernoff estimate

$$\forall x \geq 0 \qquad \mathbb{P}\left(S_n^N + R_n^N \geq r_n + \left(L_{0,n}^{N\ \star}\right)^{-1}(x) + \left(L_{1,n}^{N\ \star}\right)^{-1}(x)\right) \leq e^{-x}. \quad (4.11)$$

In other words, the probability that

$$S_n^N + R_n^N \leq r_n + \left(L_{0,n}^{N\ \star}\right)^{-1}(x) + \left(L_{1,n}^{N\ \star}\right)^{-1}(x)$$

is greater than $1 - e^{-x}$, which, together with the homogeneity properties of the inverses of Legendre-Fenchel transforms recalled in Section 5.6, gives (4.6).

The proof of (4.7) is based on Hoeffding's inequality

$$8\ \log \mathbb{E}\left(e^{tX_n^{(N,i)}}\ \mid \mathcal{F}_{n-1}^N\right) \leq t^2\ (b_n - a_n)^2$$

From these estimates, we readily find that $L_{1,n}^N(t) \leq \alpha_{2,n}^N(t) := N\ \overline{\delta}_n^2\ t^2/2$. Arguing as before, we find that

$$\left(L_{1,n}^{N\ \star}\right)^{-1}(x) \leq \left(\alpha_{2,n}^{N\ \star}\right)^{-1}(x) = \sqrt{2xN\overline{\delta}_n^2}$$

We end the proof of the second assertion using (4.11). This ends the proof of the lemma. ■

4.3 Concentration properties of mean field models

Theorem 4.5 *For any $N \geq 1$ and any $n \geq 0$, the probability of each of the following pair of events*

$$N\ [\eta_n^N - \eta_n](f_n) \leq r_n\ \left(1 + (\alpha_0^\star)^{-1}(x)\right) + N\ \overline{\sigma}_n^2\ b_n^\star\ (\alpha_1^\star)^{-1}\left(\frac{x}{N\overline{\sigma}_n^2}\right) \tag{4.12}$$

and

$$N\ [\eta_n^N - \eta_n](f_n) \leq r_n\ \left(1 + (\alpha_0^\star)^{-1}(x)\right) + \beta_n\sqrt{2xN} \tag{4.13}$$

is greater than $1 - e^{-x}$, for any $x \geq 0$.

Proof:
To simplify the presentation, we set

$$\mathcal{D}_{p,n}^{(N)} := \mathcal{D}_{\Phi_p(\eta_{p-1}^N)}\Phi_{p,n} \quad \text{and} \quad \mathcal{R}_{p,n} = \mathcal{R}^{\Phi_{p,n}}\ .$$

Under our assumptions, we have the almost sure estimates

$$\sup_{N\geq 1}\beta\left(\mathcal{D}_{p,n}^{(N)}\right) \leq \beta\left(\mathcal{D}\Phi_{p,n}\right) := \sup_{\eta\in\mathcal{P}(E_p)}\beta\left(\mathcal{D}_\eta\Phi_{p,n}\right)\ .$$

In this notation, one important consequence of the above lemma is the following decomposition

$$\begin{aligned} V_n^N &:= \sqrt{N}\ \left[\eta_n^N - \eta_n\right] \\ &= \sqrt{N}\ \sum_{p=0}^n\ \left[\Phi_{p,n}(\eta_p^N) - \Phi_{p,n}\left(\Phi_p(\eta_{p-1}^N)\right)\right] = I_n^N + J_n^N \end{aligned}$$

with the pair of random measures (I_n^N, J_n^N) given by

$$I_n^N\ :=\ \sum_{p=0}^n W_p^N\mathcal{D}_{p,n}^{(N)} \quad \text{and} \quad J_n^N := \sqrt{N}\ \sum_{p=0}^n \mathcal{R}_{p,n}\left(\eta_p^N, \Phi_p(\eta_{p-1}^N)\right).$$

In what follows f_n stands for some test function $f_n \in \mathrm{Osc}_1(E_n)$. Combining (4.4) with the generalized Minkowski integral inequality we find that

$$N\ \mathbb{E}\left(\left|\mathcal{R}_{p,n}\left(\eta_p^N, \Phi_p(\eta_{p-1}^N)\right)(f_n)\right|^m\ \middle|\ \mathcal{F}_{p-1}^{(N)}\right)^{\frac{1}{m}} \leq b(2m)^2\ \delta(R^{\Phi_{p,n}}),$$

from which we readily conclude that

$$\begin{aligned}\mathbb{E}\left(\left|\sqrt{N}J_n^N(f_n)\right|^m\right)^{\frac{1}{m}} &= N\ \mathbb{E}\left(\left|\sum_{p=0}^n \mathcal{R}_{p,n}\left(\eta_p^N,\Phi_p(\eta_{p-1}^N)\right)(f_n)\right|^m\right)^{\frac{1}{m}}\\ &\leq b(2m)^2\ \sum_{p=0}^n \delta(R^{\Phi_{p,n}})).\end{aligned}$$

Notice that

$$\sqrt{N}\ I_n^N = \sum_{p=0}^n\sum_{i=1}^N \mathcal{X}_{p,n}^{(N,i)}(f_n) \quad \text{where} \quad \mathcal{X}_{p,n}^{(N,i)}(f_n) = U_p^{(N,i)}(\mathcal{D}_{p,n}^{(N)}(f_n))$$

and the random measures $U_p^{(N,i)}$ are given, for any $g_p \in \mathrm{Osc}_1(E_p)$, by

$$U_p^{(N,i)}(g_p) := g_p\left(\xi_p^{(N,i)}\right) - K_{p,\eta_{p-1}^N}(g_p)\left(\xi_{p-1}^{(N,i)}\right).$$

In the further development of this section, we fix the final time horizon n and the the function $f_n \in \mathrm{Osc}_1(E_n)$. To clarify the presentation, we omit the final time index and the test function f_n, and we set, for any p in $[0,n]$,

$$X_p^{(N,i)} = \mathcal{X}_{p,n}^{(N,i)}(f_n), \quad S_p^N = \sum_{q=0}^p\sum_{i=1}^N X_q^{(N,i)}$$

and

$$R_p^N := N\ \sum_{k=0}^p \mathcal{R}_{q,n}\left(\eta_q^N,\Phi_q(\eta_{q-1}^N)\right).$$

At the final time horizon, we have

$$p = n \Longrightarrow S_n^N = \sqrt{N}\ I_n^N \quad \text{and} \quad R_n^N = \sqrt{N}J_n^N\ .$$

By construction, these variables form a triangular array of $\mathcal{F}_{p-1}^N$-conditionally independent random variables and

$$\mathbb{E}\left(X_p^{(N,i)})^2 \mid \mathcal{F}_{p-1}^N\ \right) = 0.$$

In addition, we readily check the following almost sure estimates

$$\left|X_p^{(N,i)}\right| \leq \beta\left(\mathcal{D}\Phi_{p,n}\right) \quad \text{and} \quad \mathbb{E}\left((X_p^{(N,i)})^2 \mid \mathcal{F}_{p-1}^N\ \right)^{\frac{1}{2}} \leq \sigma_p\ \beta\left(\mathcal{D}\Phi_{p,n}\right).$$

for any $0 \leq p \leq n$.The proof of the theorem is now a direct consequence of Lemma 4.3. ∎

5 Appendix

5.1 A first order composition lemma

Lemma 5.1 *For any pair of mappings* $\Phi_1 \in \Upsilon(E_0, E_1)$ *and* $\Phi_2 \in \Upsilon(E_1, E_2)$ *the composition mapping* $(\Phi_2 \circ \Phi_1) \in \Upsilon(E_0, E_2)$ *and we have the first order derivation type formula*

$$\mathcal{D}_\eta \left(\Phi_2 \circ \Phi_1\right) = \mathcal{D}_\eta \Phi_1 \; \mathcal{D}_{\Phi_1(\eta)} \Phi_2 \,. \tag{5.1}$$

To check this property, we first observe that under this condition, we clearly have the Lipschitz property

$$(\Phi) \qquad \left|\left[\Phi(\mu) - \Phi(\eta)\right](f)\right| \leq \int \; |(\mu - \eta)(h)| \; T_\eta^\Phi(f, dh),$$

for some collection of integral operators T_η^Φ from $\mathcal{B}(F)$ into the set $\mathrm{Osc}_1(E)$ such that

$$\sup_{\eta \in \mathcal{P}(E)} \int \; \mathrm{osc}(h) \; T_\eta^\Phi(f, dh) \leq \; \mathrm{osc}(f) \; \delta\left(T^\Phi\right) \tag{5.2}$$

for some finite constant $\delta\left(T^\Phi\right) < \infty$. Using this property, we easily check that (5.1) is met with

$$\beta\left(\mathcal{D}\left(\Phi_2 \circ \Phi_1\right)\right) \leq \beta\left(\mathcal{D}\Phi_2\right) \; \beta\left(\mathcal{D}\Phi_1\right)$$

and

$$\delta\left(R^{\Phi_2 \circ \Phi_1}\right) \leq \delta\left(T^{\Phi_1}\right) + \delta\left(T^{\Phi_1}\right)^2 \; \delta\left(R^{\Phi_2}\right).$$

This ends the proof of the lemma.

We also mention that for any pair of mappings $\Phi_1 \; : \; \eta \in \mathcal{P}(E_0) \mapsto \Phi_1 \in \mathcal{P}(E_1)$ and $\Phi_2 \; : \; \eta \in \mathcal{P}(E_1) \mapsto \Phi_1 \in \mathcal{P}(E_2)$, the composition mapping $\Phi = \Phi_2 \circ \Phi_1$ satisfies condition (Φ) as soon as this condition is met for each mapping. In this case, we also notice that

$$\delta\left(T^{\Phi_2 \circ \Phi_1}\right) \leq \delta\left(T^{\Phi_2}\right) \times \delta\left(T^{\Phi_1}\right).$$

Suppose we are given a mapping Φ defined in terms of a non linear transport formula

$$\Phi(\eta) = \eta K_\eta,$$

with a collection of Markov transitions K_η from a measurable space $(E, \mathcal{E})$ into another $(F, \mathcal{F})$ satisfying condition (K). Using the decomposition

$$\Phi(\mu) - \Phi(\eta) = \left[\eta - \mu\right] K_\eta + \mu \left[K_\mu - K_\eta\right],$$

we readily check that

$$(K) \implies (\Phi) \quad \text{with} \quad T_\eta^\Phi(f, dh) = \delta_{K_\eta(f)}(dh) + T_\eta^K(f, dh).$$

5.2 Proof of theorem 4.2

Let $\mathcal{F}^N = \{\mathcal{F}_n^N \ ; \ n \geq 0\}$ be the natural filtration associated with the N-particle system $\xi_n^{(N)}$. The first class of martingales that arises naturally in our context is the $\mathbb{R}^d$-valued and $\mathcal{F}^N$-martingale $M_n^N(f)$ defined by

$$M_n^N(f) = \sum_{p=0}^{n} \left[\eta_p^N(f_p) - \Phi_p(\eta_{p-1}^N)(f_p)\right], \tag{5.3}$$

where $f_p : x_p \in E_p \mapsto f_p(x_p) = (f_p^u(x_p))_{u=1,\ldots,d} \in \mathbb{R}^d$ is a d-dimensional and bounded measurable function. By direct inspection, we see that the vth component of the martingale $M_n^N(f) = (M_n^N(f^u))_{u=1,\ldots,d}$ is the d-dimensional and F^N-martingale defined for any $u = 1, \ldots, d$ by the formula

$$M_n^N(f^u) = \sum_{p=0}^{n} \left[\eta_p^N(f_p^u) - \Phi_p(\eta_{p-1}^N)(f_p^u)\right] = \sum_{p=0}^{n} \left[\eta_p^N(f_p^u) - \eta_{p-1}^N K_{p,\eta_{p-1}^N}(f_p^u)\right],$$

with the usual convention $K_{0,\eta_{-1}^N} = \eta_0 = \Phi_0(\eta_{-1}^N)$ for $p = 0$. The idea of the proof consists in using the CLT for triangular arrays of $\mathbb{R}^d$-valued random variables (Theorem 3.33, p. 437 in [5]). We first rewrite the martingale $\sqrt{N}\, M_n^N(f)$ in the following form:

$$\sqrt{N}\, M_n^N(f) = \sum_{i=1}^{N} \sum_{p=0}^{n} \frac{1}{\sqrt{N}} \left(f_p(\xi_p^{(N,i)}) - K_{p,\eta_{p-1}^N}(f_p)(\xi_{p-1}^{(N,i)}) \right).$$

This readily yields $\sqrt{N}\ M_n^N(f) = \sum_{k=1}^{(n+1)N}\ U_k^N(f)$ where for any $1 \leq k \leq (n+1)N$ with $k = pN + i$ for some $i = 1, \ldots, N$ and $p = 0, \ldots, n$

$$U_k^N(f) = \frac{1}{\sqrt{N}} \left(f_p(\xi_p^{(N,i)}) - K_{p,\eta_{p-1}^N}(f_p)(\xi_{p-1}^{(N,i)}) \right).$$

We further denote by $\mathcal{G}_k^N$ the σ-algebra generated by the random variables ξ_p^j for any pair index (j,p) such that $pN + j \leq k$. It can be checked that, for any $1 \leq u < v \leq d$ and for any $1 \leq k \leq (n+1)N$ with $k = pN + i$ for some $i = 1, \ldots, N$ and $p = 0, \ldots, n$, we have $\mathbb{E}(U_k^N(f^u) \mid \mathcal{G}_{k-1}^N) = 0$ and

$$\mathbb{E}(U_k^N(f^u) U_k^N(f^v) \mid \mathcal{G}_{k-1}^N)$$

$$= \tfrac{1}{N} K_{p,\eta_{p-1}^N}[(f_p^u - K_{p,\eta_{p-1}^N} f_p^u)\ (f_p^v - K_{p,\eta_{p-1}^N} f_p^v)](X_{p-1}^{(N,i)}).$$

This also yields that

$$\textstyle\sum_{k=pN+1}^{pN+N} \mathbb{E}(U_k^N(f^u) U_k^N(f^v) \mid \mathcal{F}_{k-1}^N)$$

$$= \eta_{p-1}^N[K_{p,\eta_{p-1}^N}[(f_p^u - K_{p,\eta_{p-1}^N} f_p^u)\ (f_p^v - K_{p,\eta_{p-1}^N} f_p^v)].$$

Our aim is now to describe the limiting behavior of the martingale $\sqrt{N}\ M_n^N(f)$ in terms of the process $X_t^N(f) \overset{\text{def.}}{=} \sum_{k=1}^{[Nt]+N}\ U_k^N(f)$. By the definition of the particle model associated with a given mapping Φ_n and using the fact that $\left[\frac{[Nt]}{N}\right] = [t]$, one gets that for any $1 \leq u, v \leq d$

$$\sum_{k=1}^{[Nt]+N} E\left(U_k^N(f^u)U_k^N(f^v)\,|\mathcal{F}_{k-1}^N\right)$$

$$= C_{[t]}^N\left(f^u, f^v\right) + \frac{[Nt]-N[t]}{N}\,\left(C_{[t]+1}^N\left(f^u, f^v\right) - C_{[t]}^N\left(f^u, f^v\right)\right)$$

where, for any $n \geq 0$ and $1 \leq u, v \leq d$,

$$C_n^N\left(f^u, f^v\right) = \sum_{p=0}^{n} \eta_{p-1}^N\left[K_{p,\eta_{p-1}^N}\left(\left(f_p^u - K_{p,\eta_{p-1}^N}f_p^u\right)\left(f_p^v - K_{p,\eta_{p-1}^N}f_p^v\right)\right)\right].$$

Under our regularity conditions on the McKean transitions, this implies that for any $1 \leq i, j \leq d$,

$$\sum_{k=1}^{[Nt]+N} E\left(U_k^N(f^u)U_k^N(f^v)\,|\mathcal{F}_{k-1}^N\right) \;\xrightarrow[N\to\infty]{P}\; C_t(f^u, f^v),$$

with

$$C_n\left(f^u, f^v\right) = \sum_{p=0}^{n} \eta_{p-1}[K_{p,\eta_{p-1}}\left(\left(f_p^u - K_{p,\eta_{p-1}}f_p^u\right)\left(f_p^v - K_{p,\eta_{p-1}}f_p^v\right)\right)]$$

and, for any $t \in \mathbb{R}_+$,

$$C_t(f^u, f^v) \quad = \quad C_{[t]}\left(f^u, f^v\right) + \{t\}\,\left(C_{[t]+1}\left(f^u, f^v\right) - C_{[t]}\left(f^u, f^v\right)\right).$$

Since $\left\|U_k^N(f)\right\| \leq \frac{2}{\sqrt{N}}\,(\vee_{p\leq n}\|f_p\|)$, for any $1 \leq k \leq [Nt]+N$, the conditional Lindeberg condition is clearly satisfied and therefore one concludes that the $\mathbb{R}^d$-valued martingale $\{X_t^N(f)\ ;\ t \in \mathbb{R}_+\}$ converges in law to a continuous Gaussian martingale $\{X_t(f)\ ;\ t \in \mathbb{R}_+\}$ such that, for any $1 \leq u, v \leq d$ and $t \in \mathbb{R}_+$, $\langle X(f^u), X(f^v)\rangle_t = C_t(f^u, f^v)$. Recalling that $X_{[t]}^N(f) = \sqrt{N}\ M_{[t]}^N(f)$, we conclude that the $\mathbb{R}^d$-valued and $\mathcal{F}^N$-martingale $\sqrt{N}\ M_n^N(f)$ converges in law to an $\mathbb{R}^d$-valued and Gaussian martingale $M_n(f) = (M_n(f^u))_{u=1,\ldots,d}$ such that for any $n \geq 0$ and $1 \leq u, v \leq d$

$$\langle M(f^u), M(f^v)\rangle_n = \sum_{p=0}^{n} \eta_{p-1}[K_{p,\eta_{p-1}}\left(\left(f_p^u - K_{p,\eta_{p-1}}f_p^u\right)\left(f_p^v - K_{p,\eta_{p-1}}f_p^v\right)\right)],$$

with the convention $K_{0,\eta_{-1}} = \eta_0$ for $p = 0$.

To take the final step, we let $(\varphi_n)_{n\geq 0}$ be a sequence of bounded measurable functions respectively in $\mathcal{B}(E_n)^{d_n}$. We associate with $\varphi = (\varphi_n)_n$ the sequence of functions $f = (f_p)_{0\leq p\leq n}$ defined for any $0 \leq p \leq n$ by the following formula

$$f_p = (f_p^u)_{u=0,\ldots,n} = (0, \ldots, 0, \varphi_p, 0, \ldots, 0) \in \mathcal{B}(E_p)^{d_0+\ldots+d_p+\ldots+d_n}$$

In the above display, 0 stands for the null function in $\mathcal{B}(E_p)^{d_q}$ (for $q \neq p$). By construction, we have, $f_u^u = \varphi_u$ and for any $0 \leq u \leq n$, we have that

$$f^u = (f_p^u)_{0\leq p\leq n} = (0, \ldots, 0, \varphi_u, 0, \ldots, 0) \in \mathcal{B}(E_0)^{d_0}\times\ldots\times\mathcal{B}(E_u)^{d_u}\times\ldots\mathcal{B}(E_n)^{d_n}$$

so that

$$\sqrt{N}\ M_n^N(f^u) = \sqrt{N}\ [\eta_u^N(\varphi_u) - \eta_{u-1}^N K_{u,\eta_{u-1}^N}(\varphi_u)] = V_u^N(\varphi_u)$$

and therefore

$$\sqrt{N}\ M_n^N(f) := (\sqrt{N}\ M_n^N(f^u))_{0\leq u\leq n} = (V_u^N(\varphi_u))_{0\leq u\leq n} := \mathcal{V}_n^N(\varphi).$$

We conclude that $\mathcal{V}_n^N(\varphi)$ converges in law to an $(n+1)$-dimensional and centered Gaussian random field $\mathcal{V}_n(\varphi) = (V_u(\varphi_u))_{0\leq u\leq n}$ with, for any $0 \leq u, v \leq n$,

$$\begin{aligned}&\mathbb{E}(V_u(\varphi_u^1)V_v(\varphi_v^2))\\ &= 1_u(v)\ \eta_{u-1}[K_{u,\eta_{u-1}}\left(\varphi_u^1 - K_{u,\eta_{u-1}}\varphi_u^1\right) K_{u,\eta_{u-1}}\left(\varphi_u^2 - K_{u,\eta_{u-1}}\varphi_u^2\right)].\end{aligned}$$

This ends the proof of the theorem. ∎

5.3 Feynman-Kac semigroups

In the context of Feynman-Kac flows (2.1) discussed in the introduction, the semigroup $\Phi_{p,n}$ is given by the following formula

$$\eta_n(f) = \frac{\eta_p(Q_{p,n}(f))}{\eta_p(Q_{p,n}(1))} \quad \text{with} \quad Q_{p,n} = Q_{p+1}\dots Q_{n-1}Q_n\,.$$

For $p = n$, we use the convention $Q_{n,n} = Id$, the identity operator. Also observe that

$$[\Phi_{p,n}(\mu) - \Phi_{p,n}(\eta)](f) = \frac{1}{\mu(G_{p,n,\eta})}\ (\mu - \eta)D_\eta\Phi_{p,n}(f),$$

with the first order operator

$$D_\eta\Phi_{p,n}(f) := G_{p,n,\eta}\ P_{p,n}\left(f - \Phi_{p,n}(\eta)(f)\right).$$

In the above display $G_{p,n,\eta}$ and $P_{p,n}$ stand for the potential function and the Markov operator given by

$$G_{p,n,\eta} := Q_{p,n}(1)/\eta(Q_{p,n}(1)) \quad \text{and} \quad P_{p,n}(f) = Q_{p,n}(f)/Q_{p,n}(1)\,.$$

It is now easy to check that

$$\mathcal{R}^{\Phi_{p,n}}(\mu,\eta)(f) := -\frac{1}{\mu(G_{p,n,\eta})}\ [\mu - \eta]^{\otimes 2}(G_{p,n,\eta}\otimes D_{p,n,\eta}(f)).$$

Using the fact that

$$D_\eta\Phi_{p,n}(f)(x) = G_{p,n,\eta}(x)\int\ [P_{p,n}(f)(x) - P_{p,n}(f)(y)]\ G_{p,n,\eta}(y)\ \eta(dy),$$

we find that

$$\forall f\in \text{Osc}_1(E_n) \qquad \|D_\eta\Phi_{p,n}(f)\| \leq q_{p,n}\ \beta(P_{p,n}) \quad \text{with} \quad q_{p,n} = \sup_{x,y}\frac{Q_{p,n}(1)(x)}{Q_{p,n}(1)(y)}\,.$$

This implies that

$$\beta(D\Phi_{p,n}) \leq 2\ q_{p,n}\ \beta(P_{p,n})\,.$$

Finally, we observe that

$$\left|\mathcal{R}^{\Phi_{p,n}}(\mu,\eta)(f)\right| \leq \left(2\ q_{p,n}^2\beta(D_{p,n})\right)\ \left|[\mu-\eta]^{\otimes 2}\left(\frac{G_{p,n,\eta}}{2q_{p,n}}\otimes\frac{D_{p,n,\eta}(f)}{\beta(D_{p,n})}\right)\right|$$

from which, one concludes that

$$\delta(R^{\Phi_{p,n}}) \leq 2\ q_{p,n}^2\ \beta(D\Phi_{p,n}) \leq 4\ q_{p,n}^3\ \beta(P_{p,n})\,.$$

We end this Section with the analysis of these quantities for the time homogeneous models discussed in (3.8) and (3.9). Under the condition $(M)_m$ we have for any $n \geq m \geq 1$, and $p \geq 1$,

$$q_{p,p+n} \leq\ \delta_m/\epsilon_m \quad \text{and} \quad \beta(P_{p,p+n}) \leq \left(1-\epsilon_m^2/\delta_{m-1}\right)^{\lfloor n/m\rfloor}\,. \tag{5.4}$$

The proof of these estimates relies on semi-group techniques; see chapter 4 of [2] for details. Several contraction inequalities can be deduced from these results, given below.

For any $k \geq 0$ and for $l = 1,2$,

$$\sum_{p=0}^{n} q_{p,n}^k\ \beta(P_{p,n})^l \leq\ \varpi_{k,l}(m) :=\ \frac{m\ \left(\delta_m/\epsilon_m\right)^k}{1-\left((1-\epsilon_m^2/\delta_{m-1})\right)^l}\,. \tag{5.5}$$

Notice that

$$\varpi_{k,l}(m) \leq\ m\ \delta_{m-1}\ \frac{\delta_m^k/\epsilon_m^{k+2}}{(2-(\delta_{m-1}/\epsilon_m^2))^{l-1}} \leq m\ \delta_{m-1}\ \delta_m^k/\epsilon_m^{k+2}\,,$$

$$\text{and that} \quad r_n \leq 4\ \varpi_{3,1}(m) \quad \text{and} \quad b_n^\star \leq 2\ \delta_m/\epsilon_m\,, \quad \text{as well as}$$

$$\overline{\sigma}_n^2 \leq 4\ \varpi_{2,2}(m)\ \sigma^2 \quad \text{and} \quad \beta_n^2 \leq 4\ \varpi_{2,2}(m)\,, \quad \text{with} \quad \sigma^2 := \sup_{n\geq 1}\sigma_n^2\ (\leq 1).$$

5.4 McKean mean field model of gases

We consider McKean type model of gases (2.4) presented in Section 2.3. To simplify the presentation, we consider time homogeneous models and we supress the time index. In this notation, we find that

$$\left[K_\eta - K_\mu\right](f)(x) = \int\ \nu(ds)\ \left[\eta-\mu\right](a(s,.))\ \ M(f)(s,x).$$

Observe that

$$[\eta-\mu](K_\eta-K_\mu)(f)(x) = \int\ \nu(ds)\ \left[\eta-\mu\right](a(s,.))\ \left[\eta-\mu\right](M(f)(s,.))\,.$$

Using the decomposition

$$\Phi(\eta)-\Phi(\mu) = (\eta-\mu)K_\mu + \mu(K_\eta-K_\mu) + [\eta-\mu](K_\eta-K_\mu)$$

we readily check that $\Phi \in \Upsilon(E,E)$ with the first order operator

$$D_\mu\Phi(f)(x) = [K_\mu(f)(x) - \Phi(\mu)(f)] + \int \nu(ds)\ a(s,x)\ \mu\left(M(f)(s,.)\right)$$

and the second order remainder measure

$$\mathcal{R}^\Phi(\mu,\eta)(f) = \int\ [\eta-\mu]^{\otimes 2}\left(g_s\right)\ \nu(ds)\,, \quad \text{with} \quad g_s = a(s,.)\otimes M(f)(s,.).$$

In this situation, we notice that

$$\beta(\mathcal{D}\Phi) \leq \sup_{s\in S}\nu(a(s,.)) + \beta(M)\int \nu(ds)\ \mathrm{osc}(a(s,.))$$

and

$$\delta(R^\Phi) \leq \beta(M)\int \nu(ds)\ \mathrm{osc}(a(s,.)).$$

5.5 Gaussian semigroups

To simplify the presentation, we only discuss one dimensional models. We consider the one dimensional gaussian transitions on $E=\mathbb{R}$ given by

$$K_{1,\eta}(x,dy) = \frac{1}{\sqrt{2\pi}}\ \exp\left\{-\frac{1}{2}\left(y-\eta(a)\ b(x)\right)^2\right\}\ dy$$

and

$$K_{2,\eta}(x,dy) = \frac{1}{\sqrt{2\pi}}\ \exp\left\{-\frac{1}{2}\left(y-(b(x)-\eta(a))\right)^2\right\}\ dy$$

where (a,b) is a pair of bounded functions and $\eta\in\mathcal{P}(\mathbb{R})$. We let Φ be the non linear transport mapping from $\mathcal{P}(\mathbb{R})$ into $\mathcal{P}(\mathbb{R})$ given by $\Phi_i(\eta)=\eta K_{i,\eta}$, with $i=1,2$. In this case, we can check that $\Phi_i\in\Upsilon(\mathbb{R},\mathbb{R})$ with the first order operator

$$\begin{aligned}&\left[D_\mu\Phi_1\right](f)(x)\\ &= [K_{1,\mu}(f)(x)-\Phi_1(\mu)(f)] + a(x)\ \int\ \mu(dy)\ b(y)\ K_\mu(y,dz)\ f(z)\ \ (z-\mu(a)b(y))\end{aligned}$$

and

$$\begin{aligned}&\left[D_\mu\Phi_2\right](f)(x)\\ &= [K_{2,\mu}(f)(x)-\Phi_2(\mu)(f)] - a(x)\ \int\ \mu(dy)\ K_{2,\mu}(y,dz)\ f(z)\ \ (z-[b(y)-\mu(a)])\end{aligned}$$

The proof of this assertion relies on tedious but elementary calculations, thus it is omitted.

5.6 Legendre transform and convex analysis

We associate with any increasing and convex function

$$L\ :\ t\in\mathrm{Dom(L)}\mapsto L(t)\in\mathbb{R}_+$$

defined in some domain Dom(L) $\subset \mathbb{R}_+$, with $L(0) = 0$, the Legendre-Fenchel transform $L^\star$ defined by the variational formula

$$\forall \lambda \geq 0 \qquad L^\star(\lambda) := \sup_{t \in \mathrm{Dom(L)}} (\lambda t - L(t))$$

Note that $L^\star$ is a convex increasing function with $L^\star(0) = 0$ and its inverse $(L^\star)^{-1}$ is a concave increasing function (with $(L^\star)^{-1}(0) = 0$).

For instance, the Legendre-Fenchel transforms $(\alpha_0^\star, \alpha_1^\star)$ of the pair of convex non negative functions (α_0, α_1) given below

$$\forall t \in [0, 1/2[\qquad \alpha_0(t) := -t - \frac{1}{2}\log(1 - 2t) \quad \text{and} \quad \forall t \geq 0 \quad \alpha_1(t) := e^t - 1 - t$$

are simply given by

$$\alpha_0^\star(\lambda) = \frac{1}{2}(\lambda - \log(1 + \lambda)) \quad \text{and} \quad \alpha_1^\star(\lambda) = (1 + \lambda)\log(1 + \lambda) - \lambda\,.$$

Recall that, for any centered random variable Y with values in $]-\infty, 1]$ such that $\mathbb{E}(Y^2) \leq v$, we have

$$\mathbb{E}\left(e^{tY}\right) \leq \frac{ve^t + e^{-vt}}{1 + v} \leq 1 + v\alpha_1(t) \leq \exp(v\alpha_1(t)). \tag{5.6}$$

We refer to [1] for a proof of (5.6) and for more precise results. For any pair of such functions (L_1, L_2), it is readily checked that

$$\forall t \in \mathrm{Dom}(L_2) \quad L_1(t) \leq L_2(t) \quad \text{and} \quad \mathrm{Dom}(L_2) \subset \mathrm{Dom}(L_1)$$

$$\Downarrow$$

$$L_2^\star \leq L_1^\star \quad \text{and} \quad (L_1^\star)^{-1} \leq (L_2^\star)^{-1}\,.$$

For any pair of positive numbers (u, v), We also have that

$$\forall t \in v^{-1}\mathrm{Dom}(L_2) \qquad L_1(t) = u\ L_2(v\ t)$$

$$\Downarrow$$

$$\forall \lambda \geq 0 \quad L_1^\star(\lambda) = u\ L_2^\star\left(\frac{\lambda}{uv}\right) \quad \text{and} \quad \forall x \geq 0 \quad (L_1^\star)^{-1}(x) = uv\ (L_2^\star)^{-1}\left(\frac{x}{u}\right)\,.$$

As a simple consequence of the latter results, let us quote the following property that will be used later in the further development of Section 4.3:

$$u \leq \overline{u} \quad \text{and} \quad v \leq \overline{v} \Longrightarrow uv\ (L_2^\star)^{-1}\left(\frac{x}{u}\right) \leq \overline{u}\ \overline{v}\ (L_2^\star)^{-1}\left(\frac{x}{\overline{u}}\right)\,.$$

Here we want to give upper bounds on the inverse functions of the Legendre transforms. Our motivation is due to the following result, which avoids the loss of a factor 2 when adding exponential inequalities. Let A and B be centered random variables with finite log-Laplace transform, which we denote by α_A and α_B, in a neighborhood of 0. Then, denoting by α_{A+B} the log-Laplace transform of $A + B$,

$$(\alpha_{A+B}^\star)^{-1}(t) \leq (\alpha_A^\star)^{-1}(t) + (\alpha_B^\star)^{-1}(t) \tag{5.7}$$

for any positive t (see Lemma 2.1 in [7]).

In order to obtain analytic approximations of these inverse function, one can use the Newton algorithm: let

$$F(z) = z + \frac{x - \alpha^*(z)}{(\alpha^*)'(z)}$$

and define the sequence (z_n) by $z_n = F(z_{n-1})$. From the properties of the Legendre-Fenchel transform, we also have that

$$F(z) = \left(\frac{\alpha((\alpha')^{-1}(z)) + x}{(\alpha')^{-1}(z)} \right). \tag{5.8}$$

Now recall the variational formulation of the inverse of the Legendre-Fenchel transform:

$$(\alpha^\star)^{-1}(x) = \inf_{t>0} t^{-1}(\alpha(t) + x), \tag{5.9}$$

valid for any $x \geq 0$ (see [8], p. 159 for a proof of this formula). From this formula, assuming that $\alpha''(0) > 0$ and setting $z = \alpha'(t)$, we get that

$$(\alpha^\star)^{-1}(x) = \inf_{z \in \alpha'(\mathrm{Dom}(\alpha))} F(z). \tag{5.10}$$

Let then $f(z) = \alpha((\alpha')^{-1}(z)) + x$ and $g(z) = (\alpha')^{-1}(z)$. From the strict convexity of α, the function $t \to t^{-1}((\alpha(t) + x)$ a an unique minimum t_x and is decreasing with negative derivative for $t < t_x$, increasing with positive derivative for $t > t_x$. It follows that f/g has an unique critical point $z(x)$, which is the unique global strict minimum of F and the unique fixed point of F. Furthermore $z(x) = (\alpha^\star)^{-1}(x)$.

Let $z_0 > 0$ be in the interior of the image by α' of the domain of α. If $z_0 > z(x)$, then (z_n) is a decreasing sequence of numbers bounded from below by $z(x)$. Hence (z_n) decreases to $z(x)$ as n tends to ∞. If $z_0 < z(x)$ and $F(z_0)$ belongs to the interior of $\alpha'(\mathrm{Dom}(\alpha))$, then $z_1 > z(x)$ and $(z_n)_{n>0}$ is decreasing to $z(x)$.

We now recall the convergence properties of the Newton algorithm. Assumethat $z_0 > z(x)$ and let A be a positive real such that $F''(z) \leq 2A$ for any z in $[z(x), z_0]$. Then, by the Taylor formula at order 2.

$$0 \leq z_n - z(x) \leq A^{2^n - 1}(z_0 - z(x))^{(2^n)}, \tag{5.11}$$

which provides a supergeometric rate of convergence if $A(z_0 - z(x)) < 1$.

Since F depends on x, A is a function of x. In order to get estimates of the rate of convergence of z_n to $z(x)$ for small values of x, we now assume that α' is convex. We will prove that

$$A := \frac{1}{2} \sup_{z \geq z(x)} F''(z) \leq \frac{(\alpha^\star)^{-1}(x)}{2x\alpha''(0)}. \tag{5.12}$$

To prove (5.12), we start by computing $F'' = (f/g)''$. Since $f' = zg'$,

$$(f/g)' = g'(zg - f)g^{-2}$$

Now $(zg - f)' = g + (zg' - f') = g$. It follows that

$$(f/g)'' = g'g^{-1} + (zg - f)(g''g^{-2} - 2g'^2 g^{-3}).$$

Next, for $z \geq z(x)$, $zg(z) - f(z) \geq 0$, so that

$$(f/g)''(z) \leq g'g^{-1} + (zg - f)g''g^{-2}.$$

Under the additional assumption that α' is convex, the inverse function $(\alpha')^{-1} = g$ is concave, so that $g'' \leq 0$. In that case, for $z \geq z(x)$,

$$(f/g)''(z) \leq g'(z)/g(z) = (\log g)'(z)$$

Now $\log g$ is the inverse function of $\psi(t) = \alpha'(e^t)$. From the properties of α', the function ψ is convex, so that $\log g$ is concave. Hence $(\log g)'$ is nonincreasing, which implies that

$$F''(z) \leq g'(z(x))/g(z(x)) = z(x)g'(z(x))/f(z(x)) \quad \text{for any} \quad z \geq z(x).$$

Since $f(z) \geq x$ and $g'(z(x)) \leq g'(0) = 1/\alpha''(0)$, we get (5.12), noticing that $z(x) = F(z(x)) = (\alpha^\star)^{-1}(x)$.

We now apply these results to the functions α_0 and α_1. Using the fact that

$$\frac{t^2}{2} \leq \alpha_1(t) := e^t - 1 - t \ \leq \overline{\alpha}_1(t) := \frac{t^2}{2(1 - t/3)}$$

for every $t \in [0, 3[$, and applying (B.5), p. 153 in [8], we get that

$$\sqrt{2x} \leq \left(\alpha_1^\star\right)^{-1}(x) \leq \left(\overline{\alpha}_1^\star\right)^{-1}(x) = \sqrt{2x} + (x/3).$$

Also, by the second part of Theorem B.2 in [8], the function $\overline{\alpha}_1^\star$, which is the inverse function of the above function, satisfies

$$\overline{\alpha}_1^\star(t) \geq \frac{t^2}{2(1 + (t/3))}, \tag{5.13}$$

which is the usual bound in the Bernstein inequality. Now $z = e^t - 1$, and consequently $t = \log(1 + z)$ and

$$F(z) = \frac{x + z - \log(1 + z)}{\log(1 + z)}$$

Set $z_0 = \sqrt{2x} + (x/3)$. Then $z_0 > z(x)$. Hence $z(x) < z_1 < z_0$ (here $z_1 = F(z_0)$). So

$$\left(\alpha_1^\star\right)^{-1}(x) \leq z_1 := \frac{\sqrt{2x} + (4x/3) - \log(1 + (x/3) + \sqrt{2x})}{\log(1 + (x/3) + \sqrt{2x})} \leq (x/3) + \sqrt{2x}. \tag{5.14}$$

Furthermore, from (5.11) and (5.12) and the fact that $z_0 - z(x) \leq x/3$,

$$0 \leq z_1 - \left(\alpha_1^\star\right)^{-1}(x) \leq \frac{x}{18}\left(\alpha_1^\star\right)^{-1}(x),$$

which ensures that

$$18z_1/(18 + x) \leq \left(\alpha_1^\star\right)^{-1}(x) \leq z_1.$$

In the same way, noticing that

$$t^2/(1 - 4t/3) \leq \alpha_0(t) \leq t^2/(1 - 2t) \quad \text{for any } t \in [0, 1/2[$$

we get:

$$2\sqrt{x} + (4x/3) \leq (\alpha_0^\star)^{-1}(x) \leq 2\sqrt{x} + 2x := z_0.$$

By definition of α_0, we have $\alpha_0'(t) = 2t/(1-2t)$. Let $z = 2t/(1-2t)$. Then $t = z/(2+2z)$, so that

$$F(z) = \frac{x + \alpha_0((\alpha_0')^{-1}(t))}{(\alpha_0')^{-1}(t)} = 2x + \log(1+z) + \frac{2x + \log(1+z) - z}{z}.$$

Computing $z_1 = F(z_0)$, we get

$$(\alpha_0^\star)^{-1}(x) \leq z_1 := 2x + \log(1+2x+2\sqrt{x}) + \frac{\log(1+2x+2\sqrt{x}) - 2\sqrt{x}}{2x+2\sqrt{x}} \leq 2x + 2\sqrt{x}\,, \tag{5.15}$$

which improves on the previous upper bound. Furthermore, from (5.11) and (5.12)

$$0 \leq z_1 - (\alpha_0^\star)^{-1}(x) \leq \frac{x}{9}\,(\alpha_0^\star)^{-1}(x),$$

which ensures that

$$9z_1/(9+x) \leq (\alpha_0^\star)^{-1}(x) \leq z_1.$$

References

[1] Bentkus, V., *On Hoeffding's inequalities.* Annals of Probab. vol. 32, no. 2, pp. 1650-1673 (2004).

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

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

[4] Doucet A., de Freitas J.F., Gordon N.J., *Sequential Monte-Carlo Methods in Practice*, Springer Verlag New York (2001).

[5] Jacod, J. & Shiryaev, A., *Limit Theorems for Stochastic Processes.* Springer: New York (1987).

[6] Méléard, S., Asymptotic behaviour of some interacting particle systems; McKean-Vlasov and Boltzmann models. Lecture Notes in Mathematics, Springer Berlin / Heidelberg, Volume 1627, pp. 42-95 (1996).

[7] Rio E., *Local invariance principles and their applications to density estimation.* Probability and related Fields, vol. 98, 21-45 (1994).

[8] Rio E., *Théorie asymptotique des processus aléatoires faiblement dépendants.* Mathématiques et Applications 31. Springer-Verlag Berlin Heidelberg (2000).

[9] Shiga, T. & Tanaka, H., Central limit theorem for a system of Markovian particles with mean field interaction, *Zeitschrift für Wahrscheinlichkeitstheorie verwandte Gebiete*, **69**, 439-459 (1985).

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