The Annals of Applied Probability
2011, Vol. 21, No. 3, 1017–1052
DOI: 10.1214/10-AAP716
© Institute of Mathematical Statistics, 2011

CONCENTRATION INEQUALITIES FOR MEAN FIELD PARTICLE MODELS

BY PIERRE DEL MORAL AND EMMANUEL RIO

INRIA Bordeaux-Sud-Ouest, and INRIA Bordeaux-Sud-Ouest and Université de Versailles

This article is concerned with the fluctuations and the concentration properties of a general class of discrete generation and mean field particle interpretations of nonlinear 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.

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 $\Phi_{n+1}:\mathcal{P}(E_n)\to\mathcal{P}(E_{n+1})$, $n\geq 0$, 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:

$$\eta_{n+1}=\Phi_{n+1}(\eta_n). \tag{1.1}$$

These discrete time versions of conservative and nonlinear integro-differential type equations in distribution spaces arise in a variety of scientific disciplines including in physics, biology, information theory and engineering sciences. To motivate the article, before describing their mean field particle interpretations, we illustrate these rather abstract evolution models working out explicitly some of these equa-

Received April 2010.

MSC2010 subject classifications. Primary 60E15, 60K35; secondary 60F99, 60F10, 82C22.

Key words and phrases. Concentration inequalities, mean field particle models, measure valued processes, Feynman–Kac semigroups, McKean–Vlasov models.

tions in a series of concrete examples. The first one is related to nonlinear filtering problems arising in signal processing. Suppose we are given a pair signal-observation Markov chain $(X_n, Y_n)_{n\geq 0}$ on some product space $(\mathbb{R}^{d_1} \times \mathbb{R}^{d_2})$, with some initial distribution and Markov transition of the following form:

$$\mathbb{P}\big((X_0, Y_0) \in d(x, y)\big) = \eta_0(dx) g_0(x, y) \lambda_0(dy),$$
$$\mathbb{P}\big((X_{n+1}, Y_{n+1}) \in d(x, y) | (X_n, Y_n)\big) = M_{n+1}(X_n, dx) g_{n+1}(x, y) \lambda_{n+1}(dy).$$

In the above display, λ_n stands for some reference probability measures on $\mathbb{R}^{d_2}$, g_n is a sequence of positive functions, M_{n+1} are Markov transitions from $\mathbb{R}^{d_1}$ into itself and finally η_0 stands for some initial probability measure on $\mathbb{R}^{d_1}$. For a given sequence of observations $Y = y$ delivered by some sensor, the filtering problem consists of computing sequentially the flow of conditional distributions defined by

$$\widehat{\eta}_n = \mathrm{Law}(X_n | Y_0 = y_0, \ldots, Y_n = y_n)$$

and

$$\eta_{n+1} = \mathrm{Law}(X_{n+1} | Y_0 = y_0, \ldots, Y_n = y_n).$$

These distributions satisfy a nonlinear evolution equation of the form (1.1) with the transformations

$$\Phi_{n+1}(\eta_n)(dx') = \int \widehat{\eta}_n(dx) M_{n+1}(x, dx') \quad \text{and} \tag{1.2}$$
$$\widehat{\eta}_n(dx) = \frac{G_n(x)}{\int \eta_n(dx') G_n(x')} \eta_n(dx)$$

for some collection of likelihood functions $G_n = g_n(\cdot, y_n)$. Replacing these functions by some $]0, 1]$-valued potential function G_n on $\mathbb{R}^{d_1}$, we obtain the conditional distributions of a particle absorption model X'_n with free evolution transitions M_n and killing rate $(1 - G_n)$. More precisely, if T stands for the killing time of the process, we have that

$$\widehat{\eta}_n = \mathrm{Law}(X'_n | T > n) \quad \text{and} \quad \eta_{n+1} = \mathrm{Law}(X'_{n+1} | T > n). \tag{1.3}$$

These nonabsorption conditional distributions arise in the analysis of confinement processes, as well as in computational physics with the numerical solving of Schrödinger ground state energies (see, e.g., [11, 13, 28]).

Another important class of measures arising in particle physics and stochastic optimization problems is the class of Boltzmann distributions, also known as the Gibbs measure, defined by

$$\eta_n(dx) = \frac{1}{Z_n} e^{-\beta_n V(x)} \lambda(dx), \tag{1.4}$$

with some reference probability measure λ, some inverse temperature parameter and some nonnegative potential energy function V on some state space E. The normalizing constant Z_n is sometimes called "partition function" or the "free energy."

In sequential Monte Carlo methodology, as well as in operation research literature, these multiplicative formulae are often used to compute rare events probabilities, as well as the cardinality or the volume of some complex state spaces. Further details on these stochastic techniques can be found in the series of articles [5, 29, 30]. To fix the ideas we can consider the uniform measure on some finite set E. Surprisingly, this flow of measures also satisfies the above nonlinear evolution equation, as soon as η_n is an invariant measure of M_n, for each $n \geq 1$, and the potential functions G_n are chosen of the following form: $G_n = \exp((\beta_{n+1} - \beta_n)V)$. The partition functions can also be computed in terms of the flow of measures $(\eta_p)_{0 \leq p < n}$ using the easy-to-check multiplicative formula,

$$Z_n = \prod_{0 \leq p < n} \int \eta_p(dx) G_p(x),$$

as soon as $\beta_0 = 0$. In statistical mechanics literature, the above formula is sometimes called the Jarzynski or the Crooks equality [8, 19, 20]. Notice that the stochastic models discussed above remain valid if we replace $e^{-\beta_{n+1}V}$ by any collection of functions g_{n+1} s.t. $g_{n+1} = g_n \times G_n = \prod_{0 \leq p \leq n} G_p$, for some potential functions G_n, with $n \geq 0$. Further details on this model, with several worked-out applications on concrete hidden Markov chain problems and Bayesian inference can be found in the article [10] dedicated to sequential Monte Carlo technology. All of the models discussed above can be abstracted in a single probabilistic model. The latter is often called a Feynman–Kac model, and it will be presented in some details in Section 2.1.

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

$$\Phi_{n+1}(\eta_n) = \eta_n K_{n+1,\eta_n} \tag{1.5}$$

for some collection of Markov kernels $K_{n+1,\mu}$ indexed by the time parameter $n \geq 0$, and the set of measures μ on the space E_n. We already mention that the choice of the Markov transitions $K_{n,\eta}$ is not unique. Several examples are presented in Section 2 in the context of Feynman–Kac semigroups or McKean–Vlasov diffusion-type models. To fix the ideas, we can choose elementary transitions $K_{n,\eta}(x, dx') = \Phi_n(\eta)(dx')$ that do not depend on the state variable x.

In the literature on mean field particle models, the transitions $K_{n,\eta}$ are called a choice of McKean transitions. These models provide a natural interpretation of the flow of measures η_n as the laws of the time inhomogeneous Markov chain $\overline{X}_n$ with elementary transitions

$$\mathbb{P}(\overline{X}_n \in dx | \overline{X}_{n-1}) = K_{n,\eta_{n-1}}(\overline{X}_{n-1}, dx) \qquad \text{with } \eta_{n-1} = \mathrm{Law}(\overline{X}_{n-1})$$

and starting with some initial random variable with distribution $\eta_0 = \mathrm{Law}(\overline{X}_0)$. The Markov chain $\overline{X}_n$ can be thought of as a perfect sampling algorithm. For a thorough description of these discrete generation and nonlinear McKean-type models, we refer the reader to [9]. In the further development of the article, we always assume that the mappings

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

are $\mathcal{E}_n^{\otimes N}$-measurable, for any $n \ge 0$, $N \ge 1$, and $1 \le i \le N$, 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 $\xi_n^{(N)} = (\xi_n^{(N,i)})_{1\le i\le N}$, with elementary transitions defined as

$$\mathbb{P}(\xi_{n+1}^{(N)} \in dx | \mathcal{F}_n^{(N)}) = \prod_{i=1}^N K_{n+1,\eta_n^N}(\xi_n^{(N,i)}, dx^i) \tag{1.6}$$
$$\text{with } \eta_n^N := \frac{1}{N}\sum_{j=1}^N \delta_{\xi_n^{(N,j)}}.$$

In the above displayed formula, $\mathcal{F}_n^N$ stands for the σ-field generated by the random sequence $(\xi_p^{(N)})_{0\le p\le n}$, and $dx = dx^1 \times \cdots \times dx^N$ stands for an infinitesimal neighborhood of a point $x = (x^1, \dots, x^N) \in E_n^N$. 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 $K_{n+1,\eta_n^N}(\xi_n^i, dx)$, so as soon as η_n^N is a good approximation of η_n then, in view of (1.6), 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.

These stochastic particle algorithms can be thought of in various ways: from the physical view point, they can be seen as microscopic particle interpretations of physical nonlinear measure valued equations. From the pure mathematical point of view, they can also be interpreted as natural stochastic linearizations of nonlinear evolution semigroups. From the probabilistic point of view, they can be interpreted as interacting recycling acceptance-rejection sampling techniques. In this case, they can be seen as a sequential and interacting importance sampling technique.

For instance, in the context of the nonlinear filtering equation (1.2), the mean field particle model associated with the flow of optimal one-step predictors η_n, with the McKean transitions $K_{n,\eta}(x, dx') = \Phi_n(\eta)(dx')$, is the $(\mathbb{R}^{d_1})^N$-valued Markov chain defined by sampling N conditionally independent random variables $\xi_{n+1} =$

$(\xi_{n+1}^i)_{1\leq i\leq N}$, with common distribution given by

$$\Phi_{n+1}\left(\frac{1}{N}\sum_{j=1}^{N}\delta_{\xi_n^j}\right)(dx)=\sum_{i=1}^{N}\frac{G_n(\xi_n^i)}{\sum_{j=1}^{N}G_n(\xi_n^j)}M_{n+1}(\xi_n^i,dx). \tag{1.7}$$

By construction, the resulting particle model is a simple genetic-type stochastic algorithm: the mutation and the selection transitions are dictated by the prediction and the updating transitions defined in (1.2). During the selection transition, one updates the positions of the particles in accordance with the fitness likelihood functions G_n. This mechanism is called the selection-updating transition as the more likely particles with high G_n-potential value are selected for reproduction. In other words, this transition allows particles to give birth to some particles at the expense of light particles which die. The second mechanism is called the mutation-prediction transition since at this step each particle evolves randomly according to the transition kernels M_n. Another important feature of genetic-type particle models is that their ancestral or their complete genealogical tree structure can be used to approximate the smoothing problem, including the computation of the distribution of the signal trajectories given the observations. Further details on this subject can be found in [9, 13].

The same genetic-type particle algorithm applies for the particle absorption model (1.3) and the Boltzmann–Gibbs model (1.4), by replacing, respectively, the likelihood functions by the nonabsorption rates G_n and the fitness functions $G_n=\exp((\beta_{n+1}-\beta_n)V)$.

In the reverse angle, the occupation measures of a given genetic-type particle mean field model converge, as the size of the population tends to infinity, to the solution of an evolution equation of the form (1.1), with the one-step transformations (1.2). These limiting models are often called the infinite population models. For a recent treatment on these genetic models, we refer the reader to [16].

The origins of genetic-type particle methods can be traced back in physics and molecular chemistry in the 1950s with the pioneering works of Harris and Kahn [17] and Rosenbluth and Rosenbluth [27]. During the last two decades, the mean field particle interpretations of these discrete generation measure valued equations were 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 [9, 10, 15], and the references therein. To motivate the article, we illustrate the fluctuation and the concentration results presented in this work with three additional illustrative examples, including Feynman–Kac models, McKean–Vlasov diffusion-type models, as well as interacting jump type McKean model of gases.

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 $(M_1 M_2)$ the composition operator defined by $(M_1 M_2)(f) = M_1(M_2(f))$. For time homogenous state spaces, we denote by $M^m = M^{m-1} M = M M^{m-1}$ the mth 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

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

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

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 $\beta(M) := \sup\{\mathrm{osc}(M(f));\, f \in \mathrm{Osc}_1(F)\}$.

1.2. *Description of the main results.* The mathematical and numerical analysis of the mean field particle models (1.6) is one of the most attractive research areas in pure and applied probability, as well as in advanced stochastic engineering and computational physics.

The fluctuation analysis of these discrete generation particle models around their limiting distributions is often restricted to Feynman–Kac-type models (see, e.g., [7, 9, 12, 14] and references therein) or specific continuous time mean field models including McKean–Vlasov diffusions and Boltzmann-type collision models of gases [24, 31].

In the present article, we design an original and natural stochastic perturbation analysis that applies to a rather large class of models satisfying a rather weak first-order regularity property. 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.

To describe with some precision this first main result we observe that the local sampling errors associated with the corresponding mean field particle model are expressed in terms of the centered random fields W_n^N, given by the following stochastic perturbation formulae:

$$\eta_n^N = \eta_{n-1}^N K_{n,\eta_{n-1}^N} + \frac{1}{\sqrt{N}} W_n^N. \tag{1.8}$$

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 equation (1.1) have a first-order decomposition

$$\Phi_n(\eta) - \Phi_n(\mu) \simeq (\eta - \mu) D_\mu \Phi_n \tag{1.9}$$

with a first-order integral operator $D_\mu \Phi_n$ from $\mathcal{B}(E_n)$ into $\mathcal{B}(E_{n-1})$, s.t. $D_\mu \Phi_n(1) = 0$. The precise definition of the first-order regularity property (1.9) is provided in Definition 3.1.

Our first main result is a functional central limit theorem for the random fields

$$V_n^N := \sqrt{N}[\eta_n^N - \eta_n]. \tag{1.10}$$

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 given for any $f, g \in \mathcal{B}(E_n)$, and any $n \geq 1$, by*

$$\mathbb{E}(W_n(f) W_n(g)) = \eta_{n-1} K_{n,\eta_{n-1}}\bigl([f - K_{n,\eta_{n-1}}(f)][g - K_{n,\eta_{n-1}}(g)]\bigr) \tag{1.11}$$

  and, for $n = 0$, by

$$\mathbb{E}(W_0(f) W_0(g)) = \eta_0\bigl[\bigl(f - \eta_0(f)\bigr)\bigl(g - \eta_0(g)\bigr)\bigr].$$

- *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 $\mathcal{D}_n = D_{\eta_{n-1}} \Phi_n$.

A complete detailed proof of the functional central limit theorem stated above is provided in Section 5, dedicated to a stochastic perturbation analysis of mean field particle models.

We let $\Phi_{p,n} = \Phi_{p+1,n} \circ \Phi_{p+1}$, $0 \le p \le n$, be the semigroup associated with the measure valued equation defined in (1.1). For $p = n$, we use the convention $\Phi_{n,n} = \mathrm{Id}$, the identity operator. By construction, we have

$$\Phi_{p,n}(\eta) - \Phi_{p,n}(\mu) \simeq (\eta - \mu) D_\mu \Phi_{p,n} \quad \text{and} \quad D_{\eta_p} \Phi_{p,n} = \mathcal{D}_{p,n}.$$

The fluctuation theorem stated above shows that the fluctuations of η_n^N around the limiting measure η_n is precisely dictated by first-order differential-type operators $D_{\eta_p}\Phi_{p,n}$ of the semigroup $\Phi_{p,n}$ around the flow of measures η_p, with $p \le n$. Furthermore, for any $f_n \in \mathrm{Osc}_1(E_n)$, one observes that

$$\begin{aligned} \mathbb{E}(V_n(f_n)^2) &= \sum_{p=0}^{n} \mathbb{E}((W_p[D_{\eta_p}\Phi_{p,n}(f_n)])^2) \\ &\le \sum_{p=0}^{n} \sigma_p^2 \beta(D\Phi_{p,n})^2 := \overline{\sigma}_n^2 \end{aligned} \tag{1.12}$$

with the uniform local variance parameters

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

and

$$\beta(D\Phi_{p,n}) := \sup_{\eta \in \mathcal{P}(E_p)} \beta(D_\eta \Phi_{p,n}).$$

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. The exponential deviation events discussed in this article are described in terms of the parameters

$$\overline{\sigma}_n^2 \le \beta_n^2 := \sum_{p=0}^{n} \beta(D\Phi_{p,n})^2 \quad \text{and} \quad b_n^\star := \sup_{0 \le p \le n} \beta(D\Phi_{p,n}).$$

Besides the fact that the nonasymptotic analysis of weakly dependent variables is rather well developed, the concentration properties of discrete generation and interacting particle systems often resume to asymptotic large deviation results, or to nonasymptotic rough exponential estimates (see, e.g., [9] and references therein). Our main result on 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 $N \ge 1$, $n \ge 0$, $f_n \in \mathrm{Osc}_1(E_n)$, and any $x \ge 0$ the probability of each of the following pair of events is greater than $1 - e^{-x}$*

$$V_n^N(f_n) \le \frac{r_n}{\sqrt{N}}\left(1 + \varepsilon_0^{-1}(x)\right) + \sqrt{N}\overline{\sigma}_n^2 b_n^\star \varepsilon_1^{-1}\left(\frac{x}{N\overline{\sigma}_n^2}\right)$$

and

$$V_n^N(f_n) \leq \frac{r_n}{\sqrt{N}}\left(1+\varepsilon_0^{-1}(x)\right)+\sqrt{2x}\,\beta_n.$$

In the above display, r_n stands for some parameter whose values only depend on the amplitude of the second-order terms in the development (1.9), *and the pair of functions $(\varepsilon_0, \varepsilon_1)$ are defined by*

$$\varepsilon_0(\lambda) = \tfrac{1}{2}\big(\lambda - \log(1+\lambda)\big), \qquad \varepsilon_1(\lambda) = (1+\lambda)\log(1+\lambda) - \lambda. \tag{1.13}$$

Under additional stability properties of the semigroup associated with the limiting model (1.5), *the parameters $(\overline{\sigma}_n, \beta_n, b_n^\star, r_n)$ are uniformly bounded w.r.t. the time parameter.*

A complete detailed proof of the functional concentration inequalities stated in Theorem 1.2 is provided in Section 5.3. Some of the consequences of the concentration inequalities stated above are provided in Section 4. To give a flavor of these results, using a Bernstein-type concentration inequality we will check that

$$\limsup_{N\to\infty} \log \mathbb{P}\left(V_n^N(f_n) \geq \frac{r_n}{\sqrt{N}} + \lambda\right) \leq -\frac{\lambda^2}{2(b_n^\star \overline{\sigma}_n)^2}. \tag{1.14}$$

The detailed proof of this asymptotic estimate is provided on page 1032. 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 (1.12). A more precise asymptotic estimate would require a refined moderate deviation analysis. We hope to discuss these properties in a forthcoming study.

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 abstract mean field particle models that can be studied using the fluctuation and the concentration analysis developed in this article.

In Section 3, we discuss the main regularity properties used in our analysis. In Section 4, we illustrate the impact of Theorem 1.2 with some more Bennett and Hoeffding-type concentration properties, as well as Bernstein-type concentration inequalities and uniform exponential deviation properties w.r.t. the time parameter. Section 5 is mainly concerned with the detailed proofs of the theorems stated above. We combine a natural stochastic perturbation analysis with nonlinear semigroup techniques that allow us to describe both the fluctuations and the concentration of the mean field measures in terms of the local error random field models introduced in (1.8). The functional central limit theorem is proved in Section 5.1. In Appendix A.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 5.2, we prove a technical concentration lemma for triangular arrays of conditionally independent random variables. In Section 5.3, we apply this lemma to prove concentration inequalities for mean field models.

2. Some illustrative examples.

2.1. *Feynman–Kac models.* As mentioned in the Introduction, 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 stochastic engineering, physics, biology and Bayesian statistics. 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

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

Rewritten in a slightly different way, we have

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

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

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

Using the ratio formulation (2.1) of the semigroup, we will check in Appendix A.3 that the first-order decomposition (1.9) is met with the first-order operator defined by

$$D_\mu \Phi_n(f) := \frac{1}{\mu Q_n(1)} Q_n\bigl(f - \Phi_n(\mu)(f)\bigr).$$

The nonlinear filtering model (1.2), the particle absorption model (1.3) and the Boltzmann–Gibbs distribution flow (1.4) can be abstracted in this framework by setting

$$Q_{n+1}(x, dx') = G_n(x) M_{n+1}(x, dx').$$

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:

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

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 $\|\varepsilon_n G_n\| \leq 1$. 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 $\xi_n \rightsquigarrow \widehat{\xi}_n = (\widehat{\xi}_n^i)_{1 \leq i \leq N} \rightsquigarrow \xi_{n+1}$. During the selection stage, with probability $\varepsilon_n G_n(\xi_n^i)$, we set $\widehat{\xi}_n^i = \xi_n^i$; 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 $\widehat{\xi}_n^i \rightsquigarrow \xi_{n+1}^i$ evolves according to the transition M_{n+1}. If we set $\varepsilon_n = 0$, the above particle model reduces to the simple genetic-type model discussed in (1.7) in the Introduction.

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

$$K_{n,\eta}(x, dy) = \frac{1}{\sqrt{(2\pi)^d \det(Q_n)}} \times \exp\left\{-\frac{1}{2}\bigl(y - d_n(x,\eta)\bigr)' Q_n^{-1}\bigl(y - d_n(x,\eta)\bigr)\right\} dy, \tag{2.3}$$

with a nonsingular, 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 Appendix A.5, for $d = 1$ we will check that any linear drift function d_n of the form $d_n(x, \eta) = a_n(x) + \eta(b_n)c_n(x)$, with some measurable (and nonnecessarily bounded) function a_n, and some pair of functions b_n and $c_n \in \mathcal{B}(\mathbb{R})$, the first-order decomposition (1.9) is met with the first-order operator defined by

$$D_\mu \Phi_n(f)(x) := [K_{n,\mu}(f)(x) - \Phi_n(\mu)(f)] + b_n(x) \int \mu(dy) c_n(y) K_{n,\mu}(y, dz) f(z)\bigl(z - d_n(y, \mu)\bigr).$$

In this context, the N-mean field particle model is given by the following recursion:

$$\forall 1 \le i \le N \qquad \xi_n^i = d_n(\xi_{n-1}^i, \eta_{n-1}^N) + W_n^i,$$

where $(W_n^i)_{i \ge 0}$ is a collection of independent and identically distributed d-valued Gaussian random variables with covariance matrix Q_n. The connection between these discrete generation models and the more traditional continuous time McKean–Vlasov diffusion models is as follows. Consider the partial differential equation

$$\partial_t \mu_t = \frac{1}{2} \sum_{i,j=1}^d \partial^2_{x^i,x^j}(\mu_t) - \sum_{i=1}^d \partial_{x^i}(b_t^i(\cdot, \eta_t)\mu_t),$$

where μ_t is a probability measure on $\mathbb{R}^d$, and b_t some drift term associated with some interaction kernels b_t' and given by

$$b_t(x, \mu) = \int b_t'(x, x')\mu(dx').$$

Under appropriate regularity conditions, one can show that η_t is the marginal distribution at time t of the law of the solution of the nonlinear stochastic differential equation

$$d\overline{X}_t = b_t(\overline{X}_t, \mu_t)\, dt + dB_t, \tag{2.4}$$

where B_t is a d-dimensional Brownian motion. These models have been introduced in the late-1960s by McKean [23]. The convergence of the mean field particle model associated with the diffusion (2.4) has been deeply studied in the mid-1990s by Bossy and Talay [3, 4], Méléard [24] and Sznitman [32]. We also refer the reader to the more recent treatments on McKean–Vlasov diffusion models by Bolley, Guillin and Malrieu [1] and Bolley, Guillin and Villani [2]. Besides the fact that these continuous time probabilistic models are directly connected to a rather large class of physical equations, to get some computationally feasible solution, some kind of time discretization scheme is needed. Mimicking traditional time discretization techniques of deterministic dynamical systems, several natural strategies can be used. For instance, we can use a Euler-type discretization of the diffusion given by (2.4) as follows:

$$X_{t_n}^{\Delta} - X_{t_{n-1}}^{\Delta} = b_{t_{n-1}}(X_{t_{n-1}}^{\Delta}, \mu_{t_{n-1}})\Delta + (B_{t_n} - B_{t_{n-1}})$$

on the time mesh $(t_n)_{n\geq 0}$, with $(t_n - t_{n-1}) = \Delta$, with some initial random variable with distribution $\mu_0 = \mathrm{Law}(X_0^{\Delta})$. In this situation, the elementary transitions of the approximated random states $X_{t_n}^{\Delta}$ are of the form (2.3), with the identity covariance matrices $Q_n = \mathrm{Id}$, and the drift functions $d_n(x, \eta) = x + b_{t_{n-1}}(x, \eta)$. The refined convergence analysis of these discrete time approximation models for more general models, including granular media equations is developed Malrieu and Talay [21, 22].

We mention that the semigroup derivation approach for functional fluctuation theorems and concentration inequalities developed in this article do not apply directly to any nonlinear diffusion equations with general interaction kernels c_t. The semigroup derivation technique requires one to control recursively time the integrability properties of the semigroup associated with the first and second-order derivative terms. A rather crude sufficient condition is to assume that the drift terms d_n is of the form discussed in the discrete time model.

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 $E_n = \bigcup_{s\in S_n} A_s$ 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

$$K_{n+1,\eta}(x, dy) = \int \nu_n(ds)\eta(du)a_n(s, u)M_{n+1}((s, x), dy). \tag{2.5}$$

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 of McKean's two-velocities model for Maxwellian gases corresponds 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:

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

We leave the reader to write out the mean field particle interpretation of this model. For more details on this model, we refer to [31]. In Appendix A.4, we will check that the first-order decomposition (1.9) is met with the first-order operator defined by

$$\begin{aligned} D_\mu\Phi_{n+1}(f)(x) &= [K_{n+1,\mu}(f)(x) - \Phi_{n+1}(\mu)(f)] \\ &\quad + \int \nu_n(ds)[a(s, x) - \mu(a(s, \cdot))]\mu(M_{n+1}(f)(s, \cdot)). \end{aligned}$$

3. 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. 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

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

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

$$|\mathcal{R}^\Phi(\mu, \eta)(f)| \le \int |(\mu - \eta)^{\otimes 2}(g)| R_\eta^\Phi(f, dg), \tag{3.3}$$

for some collection of integral operators R_η^Φ from $\mathcal{B}(F)$ into the set $\operatorname{Osc}_1(E)^2$ such that

$$\begin{aligned} &\sup_{\eta\in\mathcal{P}(E)} \int \operatorname{osc}(g_1)\operatorname{osc}(g_2) R_\eta^\Phi\bigl(f, d(g_1 \otimes g_2)\bigr) \le \operatorname{osc}(f)\delta(R^\Phi) \\ &\qquad\qquad\qquad\qquad\qquad\qquad \text{with } \delta(R^\Phi) < \infty. \end{aligned} \tag{3.4}$$

This rather weak first-order regularity property is satisfied for a large class of one-step transformations Φ_n associated with a nonlinear measure valued process (1.1). For instance, in Section 5.3 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) and for the McKean-type model of gases (2.5) presented in Section 2.3. The proof of this assertion is rather technical and it is postponed in Appendix A.4.

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 equation (1.1) are chosen so that $\Phi_n \in \Upsilon(E_{n-1}, E_n)$, for any $n \geq 1$. 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 measures $\mathcal{R}^{\Phi_{p,n}}(\eta, \mu)$. For further use, we let r_n be the second-order stochastic perturbation term related to the quadratic remainder measures $R^{\Phi_{p,n}}$ and defined by

$$r_n := \sum_{p=0}^{n} \delta(R^{\Phi_{p,n}}).$$

4. Some exponential concentration inequalities. Let us examine some more or less direct consequences of the concentration inequalities stated in Theorem 1.2.

When the Markov kernels $K_{n,\mu} = K_n$ do not depend on the measure μ, the N-particle model reduces 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

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

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:

$$(4.1)\qquad \begin{array}{cc} (M)_m & \exists m \geq 1, \exists \varepsilon_m > 0 \quad \text{s.t.} \\ \forall (x,y) \in E^2 & M^m(x,\cdot) \geq \varepsilon_m M^m(y,\cdot). \end{array}$$

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:

$$(4.2)\qquad \delta_m := \sup \prod_{0 \leq p < m} \big(G(x_p)/G(y_p)\big).$$

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

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

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 $\varpi_{k,l}(m) \leq m\delta_{m-1}\delta_m^k/\varepsilon_m^{k+2}$. The detailed proof of these estimates can be found in Appendix A.3.

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 1.2 resumes to the following Bennett and Hoeffding-type concentration events, respectively, given by

$$[\eta_n^N - \eta_n](f_n) \leq \overline{\sigma}_n^2 b_n^\star \varepsilon_1^{-1}\left(\frac{x}{N\overline{\sigma}_n^2}\right) \quad \text{and} \quad [\eta_n^N - \eta_n](f_n) \leq \sqrt{\frac{2x}{N}}\beta_n.$$

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

$$\varepsilon_1^{-1}(x) \leq \frac{\sqrt{2x} + (4x/3) - \log(1 + (x/3) + \sqrt{2x})}{\log(1 + (x/3) + \sqrt{2x})} \leq (x/3) + \sqrt{2x}.$$

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

$$\varepsilon_0^{-1}(x) \leq 2x + \log\big(1 + 2x + 2\sqrt{x}\big) + \frac{\log(1 + 2x + 2\sqrt{x}) - 2\sqrt{x}}{2x + 2\sqrt{x}} \leq 2x + 2\sqrt{x}.$$

A detailed proof of the upper bounds given above is detailed in Appendix A.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 4.1. *For any $N \geq 1$ and any $n \geq 0$, we have the following Bernstein-type concentration inequalities*:

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

and

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

In terms of the random fields V_n^N, the first concentration inequality stated in Corollary 4.1 takes the following form:

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

This proves the asymptotic estimate presented in (1.14).

Last, but not least, without further work, Theorem 1.2 leads to uniform concentration inequalities for mean field particle interpretations of Feynman–Kac semigroups.

COROLLARY 4.2. *In the context of Feynman–Kac models, under the mixing type condition $(M)_m$ introduced in* (4.1), *for any $N \geq 1$, any $n \geq 0$ and any $x \geq 0$ the probability of each of the following pair of events*:

$$\begin{aligned} [\eta_n^N - \eta_n](f_n) \leq{}& \frac{4}{N} \varpi_{3,1}(m)\left(1 + \varepsilon_0^{-1}(x)\right) \\ &+ \frac{8\delta_m}{\varepsilon_m} \varpi_{2,2}(m) \sigma^2 \varepsilon_1^{-1}\left(\frac{x}{4\sigma^2 \varpi_{2,2}(m) N}\right) \end{aligned}$$

and

$$[\eta_n^N - \eta_n](f_n) \leq \frac{4}{N} \varpi_{3,1}(m)\left(1 + \varepsilon_0^{-1}(x)\right) + 2\sqrt{\frac{2\varpi_{2,2}(m) x}{N}}$$

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

5. A stochastic perturbation analysis.

5.1. *Proof of the functional central limit theorem.*

DEFINITION 5.1. 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 $f \in \mathrm{Osc}_1(F)$:

$$(K) \quad \|[K_\mu - K_\eta](f)\| \le \int |(\mu - \eta)(h)| T_\eta^K(f, dh). \tag{5.1}$$

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

$$\sup_{\eta \in \mathcal{P}(E)} \int \mathrm{osc}(h) T_\eta^K(f, dh) \le \mathrm{osc}(f) \delta(T^K), \tag{5.2}$$

for some finite constant $\delta(T^\Phi) < \infty$. 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 (5.1) 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 (5.1). In this situation, we already mention that the corresponding one-step mappings $\Phi_n(\eta) = \eta K_{n,\eta}$, and the corresponding semigroup $\Phi_{p,n}$ satisfies 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 (5.1). 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. As before, this condition is met for the Gaussian transitions introduced in (2.3) and for the McKean-type model of gases (2.5) presented in Section 2.3. For a more detailed discussion on these stability properties, we refer the reader to the Appendix, on page 23.

Notice that the centered random fields W_n^N introduced in (1.8) have conditional variance functions given by

$$\mathbb{E}(W_n^N(f_n)^2 | \mathcal{F}_{n-1}^N) = \eta_{n-1}^N [K_{n,\eta_{n-1}^N}((f_n - K_{n,\eta_{n-1}^N}(f_n))^2)]. \tag{5.3}$$

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

$$\mathbb{E}(|W_n^N(f_n)|^{2m} | \mathcal{F}_{n-1}^{(N)})^{1/(2m)} \le b(2m) \qquad \text{with } b(2m)^{2m} := 2^{-m}(2m)!/m!. \tag{5.4}$$

We can also prove the following theorem.

THEOREM 5.2. *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 1.1.*

The proof of this theorem follows the same line of arguments as those we used in [9] 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 Appendix A.2.

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

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

we find that

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

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

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

We are now in position to prove the fluctuation Theorem 1.1. Using the decomposition

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

we readily prove that

$$V_n^N = \sum_{p=0}^{n} W_p^N \mathcal{D}_{p,n} + \frac{1}{\sqrt{N}} \mathcal{R}_n^N, \tag{5.5}$$

with the remainder second-order measure

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

In the above display, $\mathcal{D}_{p,n} = \mathcal{D}_{p+1}, \dots, \mathcal{D}_{n-1}\mathcal{D}_n$ stands for the semigroup associated with the integral operators $\mathcal{D}_n := D_{\eta_{n-1}} \Phi_n$, with the usual convention $\mathcal{D}_{n,n} = \mathrm{Id}$, for $p = n$. Using a first-order derivation formula for the semigroup $\Phi_{p,n}$ (cf., e.g., Lemma A.1 on page 1039), 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

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

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

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

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 (5.5). This ends the proof of Theorem 1.1.

5.2. *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

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

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

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

and R_n^N is a random perturbation term such that

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

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

$$\overline{c}_n^2 := (b_n^\star)^{-2} \sum_{p=0}^{n} c_p^2 \quad \text{and} \quad \overline{\delta}_n^2 := \sum_{p=0}^{n} \delta_p^2$$

with the middle point

$$\delta_n := \frac{b_n - a_n}{2}.$$

LEMMA 5.3. *For any $N \geq 1$ and any $n \geq 0$, the probability of each of the following pair of events*

$$T_n^N \leq d_n\left(1 + \varepsilon_0^{-1}(x)\right) + N\overline{c}_n^2 b_n^\star \varepsilon_1^{-1}\left(\frac{x}{N\overline{c}_n^2}\right) \tag{5.6}$$

and

$$T_n^N \leq d_n\left(1 + \varepsilon_0^{-1}(x)\right) + \overline{\delta}_n\sqrt{2xN} \tag{5.7}$$

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

REMARK 5.4. Notice that (5.7) 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 (5.6) 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 of these inequalities based on Legendre–Fenchel transforms estimates developed in Appendix A.6. First, combining (A.16) with (A.15) 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

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

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 (A.8) imply that, with probability greater than $1 - e^{-x}$,

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

where

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

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

$$\begin{aligned} &-\frac{1}{N}\log\mathbb{P}\left(\frac{T_n^N}{N} \geq \frac{d_n}{N} + \lambda\right) \\ &\qquad \geq \frac{\lambda^2}{2}\left(\left(b_n^\star\overline{c}_n + \frac{\sqrt{2}d_n}{\sqrt{N}}\right)^2 + \lambda\left(2d_n + \frac{b_n^\star}{3}\right)\right)^{-1}. \end{aligned} \tag{5.8}$$

In much the same way, starting from (5.6), we have, with probability greater than $1 - e^{-x}$,

$$T_n^N \leq d_n\big(1 + 2\big(x + \sqrt{x}\big)\big) + \overline{\delta}_n \sqrt{2xN} = d_n + A_n x + \sqrt{2x B_n^N}, \tag{5.9}$$

with the pair of constants

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

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

$$-\frac{1}{N} \log \mathbb{P}\left(\frac{T_n^N}{N} \geq \frac{d_n}{N} + \lambda\right) \geq \frac{\lambda^2}{2}\left(\left(\overline{\delta}_n + \frac{\sqrt{2} d_n}{\sqrt{N}}\right)^2 + 2d_n\lambda\right)^{-1}. \tag{5.10}$$

PROOF OF LEMMA 5.3. First, we observe that

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

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

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

From this observation, we readily find that

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

Using (A.7), we obtain the following almost sure inequality:

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

It implies that

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

with the increasing and convex function $\alpha_{1,n}^N(t) = N\overline{c}_n^2 \alpha_1(b_n^\star t)$.

Using (A.8), we now obtain the following Cramér–Chernoff estimate:

$$\forall x \geq 0 \qquad \mathbb{P}\big(S_n^N + R_n^N \geq r_n + (L_{0,n}^{N\star})^{-1}(x) + (L_{1,n}^{N\star})^{-1}(x)\big) \leq e^{-x}. \tag{5.11}$$

In other words, the probability that

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

is greater than $1 - e^{-x}$, which, together with the homogeneity properties of the inverses of Legendre–Fenchel transforms recalled in Appendix A.6, gives (5.6).

The proof of (5.7) is based on Hoeffding's inequality,

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

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

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

We end the proof of the second assertion using (5.11). This ends the proof of the lemma. □

5.3. *Concentration properties of mean field models.* This section is concerned with the proof of Theorem 1.2. 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

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

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

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

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

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

In what follows f_n stands for some test function $f_n \in \mathrm{Osc}_1(E_n)$. Combining (5.4) with the generalized Minkowski integral inequality we find that

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

from which we readily conclude that

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

Notice that

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

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

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

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 $[0, n]$,

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

and

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

At the final time horizon, we have

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

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

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

In addition, we readily check the following almost sure estimates:

$$|X_p^{(N,i)}| \le \beta(\mathcal{D}\Phi_{p,n}) \quad \text{and} \quad \mathbb{E}((X_p^{(N,i)})^2 | \mathcal{F}_{p-1}^N)^{1/2} \le \sigma_p \beta(\mathcal{D}\Phi_{p,n})$$

for any $0 \le p \le n$. The proof of the theorem is now a direct consequence of Lemma 5.3.

APPENDIX

A.1. A first-order composition lemma.

LEMMA A.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*

$$\mathcal{D}_\eta(\Phi_2 \circ \Phi_1) = \mathcal{D}_\eta \Phi_1 \mathcal{D}_{\Phi_1(\eta)} \Phi_2. \tag{A.1}$$

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

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

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

$$\sup_{\eta\in\mathcal{P}(E)} \int \mathrm{osc}(h) T_\eta^\Phi(f,dh) \leq \mathrm{osc}(f)\delta(T^\Phi) \tag{A.2}$$

for some finite constant $\delta(T^\Phi)<\infty$. Using this property, we easily check that (A.1) is met with

$$\beta\bigl(\mathcal{D}(\Phi_2\circ\Phi_1)\bigr) \leq \beta(\mathcal{D}\Phi_2)\beta(\mathcal{D}\Phi_1)$$

and

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

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

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

Suppose we are given a mapping Φ defined in terms of a nonlinear 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

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

we readily check that

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

A.2. Proof of Theorem 5.2. 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

$$M_n^N(f) = \sum_{p=0}^n [\eta_p^N(f_p) - \Phi_p(\eta_{p-1}^N)(f_p)], \tag{A.3}$$

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

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

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 of using the CLT for triangular arrays of $\mathbb{R}^d$-valued random variables ([18], Theorem 3.33, page 437). We first rewrite the martingale $\sqrt{N} M_n^N(f)$ in the following form:

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

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

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

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 $pN + j \le k$. It can be checked that, for any $1 \le u < v \le d$ and for any $1 \le k \le (n+1)N$ with $k = pN + i$ for some $i = 1, \dots, N$ and $p = 0, \dots, n$, we have $\mathbb{E}(U_k^N(f^u)|\mathcal{G}_{k-1}^N) = 0$ and

$$\begin{aligned} &\mathbb{E}(U_k^N(f^u) U_k^N(f^v)|\mathcal{G}_{k-1}^N) \\ &\quad = \frac{1}{N} K_{p,\eta_{p-1}^N} [(f_p^u - K_{p,\eta_{p-1}^N} f_p^u)(f_p^v - K_{p,\eta_{p-1}^N} f_p^v)]\bigl(X_{p-1}^{(N,i)}\bigr). \end{aligned}$$

This also yields that

$$\begin{aligned} &\sum_{k=pN+1}^{pN+N} \mathbb{E}(U_k^N(f^u) U_k^N(f^v)|\mathcal{F}_{k-1}^N) \\ &\quad = \eta_{p-1}^N \bigl[K_{p,\eta_{p-1}^N} [(f_p^u - K_{p,\eta_{p-1}^N} f_p^u)(f_p^v - K_{p,\eta_{p-1}^N} f_p^v)]\bigr]. \end{aligned}$$

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

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

where, for any $n \ge 0$ and $1 \le u, v \le d$,

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

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

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

with

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

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

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

Since $\|U_k^N(f)\| \le \frac{2}{\sqrt{N}}(\bigvee_{p\le n} \|f_p\|)$, for any $1 \le k \le [Nt] + N$, 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 $1 \le u, v \le d$ 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 $M_n(f) = (M_n(f^u))_{u=1,\dots,d}$ such that for any $n \ge 0$ and $1 \le u, v \le d$

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

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

To take the final step, we let $(\varphi_n)_{n\ge 0}$ 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 $f = (f_p)_{0\le p\le n}$ defined for any $0 \le p \le n$ by the following formula:

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

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

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

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

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

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 $0 \leq u, v \leq n$,

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

This ends the proof of the theorem.

A.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:

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

For $p = n$, we use the convention $Q_{n,n} = \mathrm{Id}$, 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

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

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

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

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

$$\forall f \in \mathrm{Osc}_1(E_n) \qquad \|D_\eta \Phi_{p,n}(f)\| \le q_{p,n}\beta(P_{p,n})$$

with

$$q_{p,n} = \sup_{x,y} \frac{Q_{p,n}(1)(x)}{Q_{p,n}(1)(y)}.$$

This implies that

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

Finally, we observe that

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

from which one concludes that

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

We end this section with the analysis of these quantities for the time homogeneous models discussed in (4.1) and (4.2). Under the condition $(M)_m$ we have for any $n \ge m \ge 1$ and $p \ge 1$,

$$\text{(A.4)} \qquad q_{p,p+n} \le \delta_m/\varepsilon_m \quad \text{and} \quad \beta(P_{p,p+n}) \le (1 - \varepsilon_m^2/\delta_{m-1})^{\lfloor n/m \rfloor}.$$

The proof of these estimates relies on semigroup techniques (see [9], Chapter 4, for details). Several contraction inequalities can be deduced from these results, given below.

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

$$\text{(A.5)} \qquad \sum_{p=0}^{n} q_{p,n}^k \beta(P_{p,n})^l \le \varpi_{k,l}(m) := \frac{m(\delta_m/\varepsilon_m)^k}{1 - ((1 - \varepsilon_m^2/\delta_{m-1}))^l}.$$

Notice that

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

and that

$$r_n \le 4\varpi_{3,1}(m) \quad \text{and} \quad b_n^\star \le 2\delta_m/\varepsilon_m$$

as well as

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

A.4. McKean mean field model of gases. We consider McKean-type models of gases (2.5) 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

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

Observe that

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

Using the decomposition

$$\Phi(\eta) - \Phi(\mu) = (\eta - \mu)K_\mu + \mu(K_\eta - K_\mu) + [\eta - \mu](K_\eta - K_\mu) \tag{A.6}$$

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

$$\begin{aligned} D_\mu\Phi(f)(x) &= [K_\mu(f)(x) - \Phi(\mu)(f)] \\ &\quad + \int \nu(ds)[a(s,x) - \mu(a(s,\cdot))]\mu(M(f)(s,\cdot)) \end{aligned}$$

and the second-order remainder measure

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

In this situation, we notice that

$$\beta(\mathcal{D}\Phi) \le \beta(M)\left[1 + \int \nu(ds)\operatorname{osc}(a(s,\cdot))\right]$$

and

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

A.5. Gaussian semigroups. To simplify the presentation, we only discuss time homogenous and one-dimensional models. We consider the one-dimensional gaussian transitions on $E = \mathbb{R}$ defined below:

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

with some linear drift function d_n of the form $d(x,\eta) = a(x) + \eta(b)c(x)$, with some measurable (and nonnecessarily bounded) function a, and some pair of functions b and $c \in \mathcal{B}(\mathbb{R})$. We use the decomposition

$$\begin{aligned} [K_\eta - K_\mu](f)(x) &= \int K_\mu(x,dy)\Theta(\Delta_{\mu,\eta}(x,y))f(y) \\ &\quad + \int K_\mu(x,dy)\Delta_{\mu,\eta}(x,y)f(y) \end{aligned}$$

with $\Theta(u) = e^u - 1 - u$ and the function $\Delta_{\mu,\eta}(x, y)$ defined by

$$\Delta_{\mu,\eta}(x, y) = \log \frac{dK_\eta(x, \cdot)}{dK_\mu(x, \cdot)}(y) = \Delta^{(1)}_{\mu,\eta}(x, y) + \Delta^{(2)}_{\mu,\eta}(x, y)$$

with

$$\Delta^{(1)}_{\mu,\eta}(x, y) := [d(x, \eta) - d(x, \mu)][y - d(x, \mu)] = c(x)(\eta - \mu)(b)[y - d(x, \mu)],$$

$$\Delta^{(2)}_{\mu,\eta}(x, y) := -\tfrac{1}{2}[d(x, \eta) - d(x, \mu)]^2 = -\tfrac{1}{2}c(x)^2[(\eta - \mu)(b)]^2.$$

Under our assumptions on the drift function d, we have

$$\left|\Delta^{(1)}_{\mu,\eta}(x, y)\right| \le \|c\| \operatorname{osc}(b)|y - d(x, \mu)| \quad \text{and} \quad \left|\Delta^{(2)}_{\mu,\eta}(x, y)\right| \le \|c\|^2 \operatorname{osc}(b)^2/2.$$

Using the fact that $|\Theta(u)| \le e^{|u|}u^2/2$, after some elementary manipulations we prove that

$$\sup_{x\in\mathbb{R}} \left| [K_\eta - K_\mu](f)(x) - \int K_\mu(x, dy)\Delta^{(1)}_{\mu,\eta}(x, y) f(y) \right| \le C[(\eta - \mu)(b)]^2 \|f\|$$

with some finite constant $C < \infty$ whose values only depend on $\|c\|$ and $\operatorname{osc}(b)$. On the other hand, we have

$$\int (\eta - \mu)(dx) \int K_\mu(x, dy)\Delta^{(1)}_{\mu,\eta}(x, y) f(y) = (\eta - \mu)^{\otimes 2}\big(b \otimes (K'_\mu(f))\big)$$

and

$$\int \mu(dx) \int K_\mu(x, dy)\Delta^{(1)}_{\mu,\eta}(x, y) f(y) = (\eta - \mu)(b)\mu(K'_\mu(f))$$

with the bounded integral operator K'_μ defined by

$$K'_\mu(f)(x) = c(x) \int K_\mu(x, dy)[y - d(x, \mu)] f(y).$$

Using the decomposition (A.6) we prove that

$$\Phi(\eta)\big(f - \Phi(\mu)(f)\big) = (\eta - \mu)D_\mu\Phi(f) + \mathcal{R}^\Phi(\eta, \mu)(f),$$

with the first-order operator

$$D_\mu\Phi(f) = K_\mu\big(f - \Phi(\mu)(f)\big) + b\mu(K'_\mu(f))$$

and a second-order remainder term such that

$$|\mathcal{R}^\Phi(\eta, \mu)(f)| \le C'\big[\big|(\eta - \mu)^{\otimes 2}\big(b \otimes (K'_\mu(f))\big)\big| + [(\eta - \mu)(b)]^2 \operatorname{osc}(f)\big]$$

with some finite constant $C' < \infty$ whose values only depend on $\|c\|$ and $\operatorname{osc}(b)$. Using the fact that

$$K'_\mu(1) = 0 \quad \text{and} \quad \|K'_\mu(f)\| = \big\|K'_\mu\big(f - \Phi(\mu)(f)\big)\big\| \le \|c\| \operatorname{osc}(f),$$

we conclude that (3.3) and (3.4) are met with $\delta(R^\Phi) \le C' \operatorname{osc}(b)(2\|c\| + \operatorname{osc}(b))$, and condition (3.2) is satisfied with $\beta(\mathcal{D}\Phi) \le 1 + \|c\| \operatorname{osc}(b)$.

A.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 $\mathrm{Dom}(L) \subset \mathbb{R}_+$, with $L(0) = 0$, the Legendre–Fenchel transform $L^\star$ defined by the variational formula

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

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 nonnegative functions (α_0, α_1) given below:

$$\forall t \in [0, 1/2[\qquad \alpha_0(t) := -t - \tfrac{1}{2} \log(1 - 2t)$$

and

$$\forall t \geq 0 \qquad \alpha_1(t) := e^t - 1 - t$$

are simply given by

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

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

$$\mathbb{E}(e^{tY}) \leq \frac{ve^t + e^{-vt}}{1 + v} \leq 1 + v\alpha_1(t) \leq \exp(v\alpha_1(t)). \tag{A.7}$$

We refer to [6] for a proof of (A.7) and for more precise results. For any pair of such functions (L_1, L_2), it is readily checked that

$$\forall t \in \mathrm{Dom}(L_2) \qquad L_1(t) \leq L_2(t) \quad \text{and} \quad \mathrm{Dom}(L_2) \subset \mathrm{Dom}(L_1)$$
$$\Downarrow$$
$$L_2^\star \leq L_1^\star \quad \text{and} \quad (L_1^\star)^{-1} \leq (L_2^\star)^{-1}.$$

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

$$\forall t \in v^{-1}\,\mathrm{Dom}(L_2) \qquad L_1(t) = uL_2(vt)$$
$$\Downarrow$$
$$\forall \lambda \geq 0 \qquad L_1^\star(\lambda) = uL_2^\star\left(\frac{\lambda}{uv}\right) \quad \text{and} \quad \forall x \geq 0 \qquad (L_1^\star)^{-1}(x) = uv(L_2^\star)^{-1}\left(\frac{x}{u}\right).$$

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 5.3:

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

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 $A+B$,

$$(\alpha^\star_{A+B})^{-1}(t) \le (\alpha^\star_A)^{-1}(t) + (\alpha^\star_B)^{-1}(t) \tag{A.8}$$

for any positive t (see [25], Lemma 2.1).

In order to obtain analytic approximations of these inverse functions, 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

$$F(z) = \left(\frac{\alpha((\alpha')^{-1}(z)) + x}{(\alpha')^{-1}(z)}\right). \tag{A.9}$$

Now recall the variational formulation of the inverse of the Legendre–Fenchel transform,

$$(\alpha^\star)^{-1}(x) = \inf_{t>0} t^{-1}\bigl(\alpha(t) + x\bigr), \tag{A.10}$$

valid for any $x \ge 0$ (see [26], page 159 for a proof of this formula). From this formula, assuming that $\alpha''(0) > 0$ and setting $z = \alpha'(t)$, we get that

$$(\alpha^\star)^{-1}(x) = \inf_{z \in \alpha'(\mathrm{Dom}(\alpha))} F(z). \tag{A.11}$$

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 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 a 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. Assume that $z_0 > z(x)$ and let A be a positive real such that $F''(z) \le 2A$ for any z in $[z(x), z_0]$. Then, by the Taylor formula at order 2,

$$0 \le z_n - z(x) \le A^{2^n-1}\bigl(z_0 - z(x)\bigr)^{(2^n)}, \tag{A.12}$$

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

$$A := \frac{1}{2} \sup_{z \geq z(x)} F''(z) \leq \frac{(\alpha^\star)^{-1}(x)}{2x\alpha''(0)}. \tag{A.13}$$

To prove (A.13), 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

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

Since $f(z) \geq x$ and $g'(z(x)) \leq g'(0) = 1/\alpha''(0)$, we get (A.13), 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), page 153 in [26], we get that

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

Also, by the second part of Theorem B.2 in [26], the function $\overline{\alpha}_1^\star$, which is the inverse function of the above function, satisfies

$$\overline{\alpha}_1^\star(t) \geq \frac{t^2}{2(1 + (t/3))}, \tag{A.14}$$

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

$$
\begin{aligned}
(\alpha_1^\star)^{-1}(x) \le z_1 &:= \frac{\sqrt{2x} + (4x/3) - \log(1 + (x/3) + \sqrt{2x})}{\log(1 + (x/3) + \sqrt{2x})} \\
&\le (x/3) + \sqrt{2x}.
\end{aligned}
\tag{A.15}
$$

Furthermore, from (A.12) and (A.13) and the fact that $z_0 - z(x) \le x/3$,

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

which ensures that

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

In the same way, noticing that

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

we get

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

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

$$
\begin{aligned}
(\alpha_0^\star)^{-1}(x) \le z_1 &:= 2x + \log\bigl(1 + 2x + 2\sqrt{x}\bigr) \\
&\qquad + \frac{\log(1 + 2x + 2\sqrt{x}) - 2\sqrt{x}}{2x + 2\sqrt{x}} \\
&\le 2x + 2\sqrt{x},
\end{aligned}
\tag{A.16}
$$

which improves on the previous upper bound. Furthermore, from (A.12) and (A.13)

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

which ensures that

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

REFERENCES

[1] BOLLEY, F., GUILLIN, A. and MALRIEU, F. (2010). Trend to equilibrium and particle approximation for a weakly selfconsistent Vlasov–Fokker–Planck equation. *M2AN* **44** 867–884.

[2] BOLLEY, F., GUILLIN, A. and VILLANI, C. (2007). Quantitative concentration inequalities for empirical measures on non-compact spaces. *Probab. Theory Related Fields* **137** 541–593. MR2280433

[3] BOSSY, M. and TALAY, D. (1997). A stochastic particle method for the McKean–Vlasov and the Burgers equation. *Math. Comp.* **66** 157–192. MR1370849

[4] BOSSY, M. and TALAY, D. (1996). Convergence rate for the approximation of the limit law of weakly interacting particles: Application to the Burgers equation. *Ann. Appl. Probab.* **6** 818–861. MR1410117

[5] CEROU, F., DEL MORAL, P. and GUYADER, A. (2010). A non asymptotic variance theorem for unnormalized Feynman-Kac particle models. *Ann. Inst. H. Poincaré*. To appear.

[6] BENTKUS, V. (2004). On Hoeffding's inequalities. *Ann. Probab.* **32** 1650–1673. MR2060313

[7] CHOPIN, N. (2004). Central limit theorem for sequential Monte Carlo methods and its application to Bayesian inference. *Ann. Statist.* **32** 2385–2411. MR2153989

[8] CROOKS, G. E. (1998). Nonequilibrium measurements of free energy differences for microscopically reversible Markovian systems. *J. Stat. Phys.* **90** 1481–1487. MR1628273

[9] DEL MORAL, P. (2004). *Feynman–Kac Formulae: Genealogical and Interacting Particle Systems with Applications*. Springer, New York. MR2044973

[10] DEL MORAL, P., DOUCET, A. and JASRA, A. (2006). Sequential Monte Carlo samplers. *J. R. Stat. Soc. Ser. B Stat. Methodol.* **68** 411–436. MR2278333

[11] DEL MORAL, P. and MICLO, L. (2003). Particle approximations of Lyapunov exponents connected to Schrödinger operators and Feynman–Kac semigroups. *ESAIM Probab. Stat.* **7** 171–208. MR1956078

[12] DEL MORAL, P. and MICLO, L. (2000). Branching and interacting particle systems approximations of Feynman–Kac formulae with applications to non-linear filtering. In *Séminaire de Probabilités, XXXIV. Lecture Notes in Math.* **1729** 1–145. Springer, Berlin. MR1768060

[13] DEL MORAL, P. and DOUCET, A. (2004). Particle motions in absorbing medium with hard and soft obstacles. *Stoch. Anal. Appl.* **22** 1175–1207. MR2089064

[14] DEL MORAL, P. and GUIONNET, A. (1999). Central limit theorem for nonlinear filtering and interacting particle systems. *Ann. Appl. Probab.* **9** 275–297. MR1687359

[15] DOUCET, A., DE FREITAS, N. and GORDON, N. (2001). *Sequential Monte Carlo Methods in Practice*. Springer, New York. MR1847783

[16] STEINSALTZ, D., EVANS, S. N. and WACHTER, K. W. (2005). A generalized model of mutation-selection balance with applications to aging. *Adv. in Appl. Math.* **35** 16–33. MR2141503

[17] HARRIS, T. E. and KAHN, H. (1951). Estimation of particle transmission by random sampling. *Natl. Bur. Stand. Appl. Math. Ser.* **12** 27–30.

[18] JACOD, J. and SHIRYAEV, A. N. (1987). *Limit Theorems for Stochastic Processes. Grundlehren der Mathematischen Wissenschaften* [*Fundamental Principles of Mathematical Sciences*] **288**. Springer, Berlin. MR0959133

[19] JARZYNSKI, C. (1997). Nonequilibrium equality for free energy differences. *Phys. Rev. Lett.* **78** 2690–2693.

[20] JARZYNSKI, C. (1997). Equilibrium free-energy differences from nonequilibrium measurements: A master-equation approach. *Phys. Rev. E* **56** 5018.

[21] MALRIEU, F. (2003). Convergence to equilibrium for granular media equations and their Euler schemes. *Ann. Appl. Probab.* **13** 540–560. MR1970276

[22] MALRIEU, F. and TALAY, D. (2005). Concentration inequalities for Euler schemes. In *Monte Carlo and Quasi-Monte Carlo Methods* 2004 355–371. Springer, Berlin. MR2208718

[23] MCKEAN, H. P. JR. (1967). Propagation of chaos for a class of non-linear parabolic equations. In *Stochastic Differential Equations* (*Lecture Series in Differential Equations*, *Session* 7, *Catholic Univ*., 1967) 41–57. Air Force Office Sci. Res., Arlington, VA. MR0233437

[24] MÉLÉARD, S. (1996). Asymptotic behaviour of some interacting particle systems; McKean–Vlasov and Boltzmann models. In *Probabilistic Models for Nonlinear Partial Differential Equations* (*Montecatini Terme*, 1995). *Lecture Notes in Math*. **1627** 42–95. Springer, Berlin. MR1431299

[25] RIO, E. (1994). Local invariance principles and their application to density estimation. *Probab. Theory Related Fields* **98** 21–45. MR1254823

[26] RIO, E. (2000). *Théorie Asymptotique des Processus Aléatoires Faiblement Dépendants*. *Mathématiques and Applications* (*Berlin*) [*Mathematics and Applications*] **31**. Springer, Berlin. MR2117923

[27] ROSENBLUTH, M. N. and ROSENBLUTH, A. W. (1955). Monte-Carlo calculations of the average extension of macromolecular chains. *J. Chem*. *Phys*. **23** 356–359.

[28] ROUSSET, M. (2006). On the control of an interacting particle estimation of Schrödinger ground states. *SIAM J. Math*. *Anal*. **38** 824–844 (electronic). MR2262944

[29] RUBINSTEIN, R. (2009). The Gibbs cloner for combinatorial optimization, counting and sampling. *Methodol*. *Comput*. *Appl*. *Probab.* **11** 491–549. MR2551567

[30] RUBINSTEIN, R. Y. (2010). Randomized algorithms with splitting: Why the classic randomized algorithms do not work and how to make them work. *Methodol*. *Comput*. *Appl*. *Probab.* **12** 1–50.

[31] SHIGA, T. and TANAKA, H. (1985). Central limit theorem for a system of Markovian particles with mean field interactions. *Z. Wahrsch*. *Verw*. *Gebiete* **69** 439–459. MR0787607

[32] SZNITMAN, A.-S. (1991). Topics in propagation of chaos. In *École D'Été de Probabilités de Saint–Flour XIX*—1989. *Lecture Notes in Math*. **1464** 165–251. Springer, Berlin. MR1108185

CENTRE INRIA BORDEAUX-SUD-OUEST
INSTITUT DE MATHÉMATIQUES DE BORDEAUX
UNIVERSITÉ BORDEAUX 1
351 COURS DE LA LIBÉRATION
33405 TALENCE CEDEX
FRANCE
E-MAIL: Pierre.Del-Moral@inria.fr

CENTRE INRIA BORDEAUX-SUD-OUEST
AND
LABORATOIRE DE MATHÉMATIQUES DE VERSAILLES
UNIVERSITÉ DE VERSAILLES, BÂTIMENT FERMAT
45 AV. DES ETATS-UNIS
78035 VERSAILLES CEDEX
FRANCE
E-MAIL: rio@math.uvsq.fr