# On the Exponential Ergodicity of the McKean-Vlasov SDE Depending on a Polynomial Interaction

Mohamed Alfaki Ag Aboubacrine Assadeck



October 18, 2024

# ON THE EXPONENTIAL ERGODICITY OF THE MCKEAN-VLASOV SDE DEPENDING ON A POLYNOMIAL INTERACTION *

**Mohamed, Alfaki ABOUBACRINE ASSADEK**
Univ Angers, CNRS, LAREMA,
SFR MATHSTIC,
F-49000 Angers, France
mohamedalfaki.agaboubacrineassadeck@univ-angers.fr

October 17, 2024

**ABSTRACT**

In this paper, we study the long time behaviour of the Fokker-Planck and the kinetic Fokker-Planck equations with *many body interaction*, more precisely with interaction defined by U-statistics, whose macroscopic limits are often called McKean-Vlasov and Vlasov-Fokker-Planck equations respectively. In the continuity of the recent papers [1, 2, 3] and [4, 5, 6], we establish *nonlinear functional inequalities* for the limiting McKean-Vlasov SDEs related to our particle systems. In the first order case, our results rely on *large deviations* for U-statistics and a *uniform logarithmic Sobolev inequality* in the number of particles for the invariant measure of the particle system. In the kinetic case, we first prove a uniform (in the number of particles) *exponential convergence to equilibrium* for the solutions in the weighted Sobolev space $H^1(\mu)$ with a rate of convergence which is explicitly computable and independent of the number of particles. In a second time, we quantitatively establish an exponential return to equilibrium in Wasserstein's $\mathcal{W}_2$-metric for the Vlasov-Fokker-Planck equation. Some concrete examples are also provided.

***K*eywords and phrases:** U-statistics · propagation of chaos · polynomial interaction · (kinetic) Fokker-Planck equation · McKean-Vlasov equation · functional inequalities · convergence to equilibrium · (hypo)coercivity

## Contents



* *Acknowledgements*: **The author thanks Fabien PANLOUP for extensive discussions and suggestions. Thanks also to everyone who has contributed in one way or another to the realization of this work. Last but not least, thankful for the benefits from HENRI LEBESGUE CENTER (program ANR-11-LABX-0020-0) such as MASTER SCHOLARSHIPS, LEBESGUE DOCTORAL MEETING, POST DOC POSITIONS... and thankful for the benefits for the ANR project RAWABRANCH number ANR-23-CE40-0008.**
*Mathematics Subject Classification*: **39B62; 82C31; 26D10; 47D07; 60G10; 60H10; 60J60**

## 1 Introduction

In the continuity of the recent papers [2] and [3], we establish exponential convergence towards equilibrium for a class of McKean-Vlasov and Vlasov-Fokker-Planck with *polynomial interaction* (macroscopic interaction associated with U-statistics and defined in Eq. (1.17) and Eq. (1.18)). Before going further into the details, we recall the general setting related to our problem.

**General homogeneous McKean-Vlasov diffusion.** The processes studied in this paper belong to the following class of stochastic differential equations:

$$dX_t = b(X_t, \mathbb{P}_{X_t})dt + \sigma(X_t, \mathbb{P}_{X_t})dB_t, \tag{1.1}$$

with respectively $b : \mathbb{R}^D \times \mathscr{P}(\mathbb{R}^D) \longrightarrow \mathbb{R}^D$ the drift coefficient, $\sigma : \mathbb{R}^D \times \mathscr{P}(\mathbb{R}^D) \longrightarrow \mathscr{M}_{D,p}(\mathbb{R})$ the diffusion coefficient and $(B_t)_{t\geq 0}$ a standard $p-$dimensional Brownian motion. More precisely, we are interested in the study of exponential ergodicity of the process defined by

$$dX_t = -(\mathscr{D}_m F(\mathbb{P}_{X_t}, X_t) + \frac{\sigma^2}{2}\nabla V(X_t))dt + \sigma dB_t, \tag{1.2}$$

where $F : \mathscr{P}(\mathbb{R}^D) \to \overline{\mathbb{R}}$, $\mathscr{D}_m F$ is the intrinsic derivative (**L**$-$derivation or derivation in the sense of Fréchet of F on the probability measure space, see Eq. (1.26) for precise definition) which is none other than the gradient of a flat derivative (see Eq. (1.26)) of F: $\mathscr{D}_m F(m,\cdot) := \nabla \frac{\delta F}{\delta m}(m,\cdot)$ (for example, if $F(m) = \int \varphi dm$, we have $\frac{\delta F}{\delta m}(m,x) = \varphi(x)$ then, $\mathscr{D}_m F(m,x) = \nabla\varphi(x)$), V is a *confinement potential* and $\sigma > 0$ (in this paper, without loss of generality and for the sake of standardization, we take $\sigma = \sqrt{2}$). Equation (1.2) also writes

$$dX_t = -\nabla \frac{\delta H}{\delta m}(\mathbb{P}_{X_t}, X_t)dt + \sigma dB_t \tag{1.3}$$

with the functional H given by

$$H(\mu) := F(\mu) + \frac{\sigma^2}{2}\int V d\mu. \tag{1.4}$$

With these notations, considering polynomial interactions means that F is a *polynomial on the probability space of degree at least two* (see Eq. (1.17) for details). The second term being a polynomial of degree 1, the function H is also a polynomial on the probability space (without constant term).

**The related mean-field particle system.** The $n-$particle associated with (1.2) is given by the following system of SDEs:

$$\forall i \in \{1,\dots,n\}, \quad dX_t^{i,n} = b(X_t^{i,n},\mu_{X_t^n})dt + \sigma(X_t^{i,n},\mu_{X_t^n})dB_t^i, \tag{1.5}$$

where $B^1,\dots,B^n$ are $n$ independent Brownian motions and $\mu_{X_t^n}$ denotes the empirical measure defined by

$$\mu_x := \frac{1}{n}\sum_{k=1}^{n}\delta_{x_k}, \quad x = (x_1,\dots,x_n) \in (\mathbb{R}^D)^n.$$

Under standard assumptions, $(X^{i,n})_{i=1}^n$ is a Markov process with infinitesimal generator defined on an appropriate subspace of $\mathscr{C}_b((\mathbb{R}^D)^n)$ by,

$$\mathscr{L}_n\varphi(x) := \sum_{i=1}^{n}\mathscr{L}_{\mu_x}\blacksquare_i\varphi(x) \tag{1.6}$$

where for a given $\mu \in \mathscr{P}(\mathbb{R}^D)$,

$$\mathscr{L}_\mu := b(\cdot,\mu)\cdot\nabla + \frac{1}{2}\mathbf{Tr}(\sigma\sigma^*(\cdot,\mu)\nabla^2), \quad \mu \in \mathscr{P}(\mathbb{R}^D) \tag{1.7}$$

and the notation $\mathscr{L}\blacksquare_i\varphi$ denotes the action of an operator $\mathscr{L}$ defined on (a subset of) $\mathscr{C}_b(\mathbb{R}^D)$ against the i-th variable of a function $\varphi \in \mathscr{C}_b((\mathbb{R}^D)^n)$; in other words, $\mathscr{L}\blacksquare_i\varphi$ is defined as the function:

$$x \in (\mathbb{R}^D)^n \longmapsto \mathscr{L}[y \longmapsto \varphi(x_1,\dots,x_{i-1},y,x_{i+1},\dots,x_n)](x_i) \in \mathbb{R}.$$

In the family of equations of type (1.5), *kinetic* particle systems correspond to the case where $Z_t^{i,n} := (X_t^{i,n},V_t^{i,n}) \in \mathbb{R}^d\times\mathbb{R}^d$ is a particle defined by two arguments, its position $X_t^{i,n}$ and its velocity $V_t^{i,n}$ defined as the time derivative of the position. The evolution of a system of kinetic particles is usually governed by Newton's laws of motion. In a random setting, the typical system of SDEs is thus the following:

$$\forall i \in \{1,\dots,n\}, \quad \begin{cases} dX_t^{i,n} = V_t^{i,n}dt \\ dV_t^{i,n} = \mathbf{F}(X_t^{i,n},V_t^{i,n},\mu_{X_t^n})dt + \sigma(X_t^{i,n},V_t^{i,n},\mu_{X_t^n})dB_t^i, \end{cases} \tag{1.8}$$

where $\mathbf{F}:\mathbb{R}^d\times\mathbb{R}^d\times\mathscr{P}(\mathbb{R}^d)\to\mathbb{R}^d$ and $\sigma:\mathbb{R}^d\times\mathbb{R}^d\times\mathscr{P}(\mathbb{R}^d)\to\mathscr{M}_d(\mathbb{R})$. Note that it is often assumed that the force field induced by the interactions between the particles depends only on their positions. Note that in the system Eq. (1.5) there are actually $n$D independent one-dimensional Brownian motions. In particular, for kinetic particles defined by their positions and velocities, the noise is often added on the velocity variable only (this case is nevertheless covered by Eq. (1.5) with a block-diagonal matrix $\sigma$ with a vanishing block on the position variable). This special case of the McKean-Vlasov diffusion in $\mathbb{R}^D = \mathbb{R}^d\times\mathbb{R}^d$ is also often called a *second order system* by opposition to the *first order systems* when $\mathbb{R}^D = \mathbb{R}^d$. In this paper, we will establish some uniform exponential convergence of the particle systems Eq. (3.5) and Eq. (3.16) (defined below) which in turn will allow us to derive the same properties for their mean-field limiting dynamics.

**McKean-Vlasov PDE.** It is classically assumed that the domain of the generator $\mathscr{L}_\mu$ does not depend on $\mu$. This domain will be denoted by $\mathscr{F}\subset\mathscr{C}_b(\mathbb{R}^D)$. In that case, it is easy to guess the form of the associated nonlinear system obtained when $n\to+\infty$. Taking a test function of the form $\varphi(x_1,\dots,x_n) := \psi(x_1)$, where $\psi\in\mathscr{F}$, one obtains the one-particle Kolmogorov equation:

$$\frac{d}{dt}\langle\mathbb{P}_{X_t^{1,n}},\psi\rangle = \int_{(\mathbb{R}^D)^n}\mathscr{L}_{\mu_x}\varphi(x)\mathbb{P}_{X_t^n}(dx) = \mathbb{E}[\mathscr{L}_{\mu_{X_t^n}}\varphi(X_t^n)]. \tag{1.9}$$

Note that the right-hand side depends on the $n-$particle distribution. If the limiting system exists (propagation of chaos) then, its law $\mu_t$ at time $t\geq 0$ is typically obtained as the limit of the empirical measure process:

$$\mu_{X_t^n} \overset{n\to+\infty}{\longrightarrow} \mu_t \tag{1.10}$$

This also implies $\mathbb{P}_{X_t^{1,n}} \overset{n\to+\infty}{\longrightarrow} \mu_t$. Reporting formally in the previous equation, it follows that $\mu_t$ should satisfy

$$\left(\forall\varphi\in\mathscr{F}, \quad \frac{d}{dt}\langle\mu_t,\varphi\rangle = \langle\mu_t,\mathscr{L}_{\mu_t}\varphi\rangle\right) \Longleftrightarrow \partial_t\mu_t = \mathscr{L}_{\mu_t}^\dagger\mu_t, \quad \text{where} \quad \mathscr{L}_{\mu_t}^\dagger \quad \text{is the weak adjoint of} \quad \mathscr{L}_{\mu_t}. \tag{1.11}$$

This is the *weak form* of the so-called the *(nonlinear) evolution equation* induced by (1.1). The evolution equation Eq. (1.11) can be written in a *strong form* (at least formally) and reads:

$$\partial_t\mu_t(x) = -\nabla_x\cdot(b(x,\mu_t)\mu_t) + \frac{1}{2}\sum_{i,j=1}^{D}\partial_{x_i}\partial_{x_j}\Big((\sigma\sigma^*)_{i,j}(x,\mu_t)\mu_t\Big). \tag{1.12}$$

This is a *nonlinear Fokker-Planck equation* which is used in many important modelling problems. This equation was obtained (formally) previously using only the generators when $n \to +\infty$. Here, there is an alternative way to derive the limiting system: looking at the SDE system Eq. (1.5), the empirical measure can be formally replaced by its expected limit $\mu_t$. Since all the particles are *exchangeable*, this can be done in any of the $n$ equations. The result is a process $(\overline{X}_t)_{t\geq 0}$ which solves the SDE: (McKean-Vlasov process)

$$d\overline{X}_t = b(\overline{X}_t, \mu_t)dt + \sigma(\overline{X}_t, \mu_t)dB_t, \tag{1.13}$$

where $(B_t)_{t\geq 0}$ is a Brownian motion and $\overline{X}_0 \sim \mu_0$. Moreover, since for all $i$, $X_t^{i,n}$ has law $\mathbb{P}_{X_t^{1,n}}$ and since it is expected that $\mathbb{P}_{X_t^{1,n}} \overset{n\to+\infty}{\longrightarrow} \mu_t$, the process $(\overline{X}_t)_{t\geq 0}$ and the distributions $(\mu_t)_{t\geq 0}$ should be linked by the relation: for all $t \geq 0$, $\overline{X}_t \sim \mu_t$. The dependency of the solution of a SDE on its law is a special case of what is called a nonlinear process in the sense of McKean (Eq. (1.1) is equivalent to Eq. (1.11) via mean-field system given by Eq. (1.6). Under appropriate conditions, the process Eq. (1.13) is well defined or (equivalently) the PDE Eq. (1.12) is well-posed (see [7, Proposition.1] or Theorem A.2 for details).

*Remark* 1.1. Note that when $\sigma = 0$, the limit equation Eq. (1.12) is the renowned *Vlasov equation* which is historically one of the first and most important models in plasma physics and celestial mechanics.

Equivalently, our main objective is the study of the long-time behavior of the solution flow of the nonlinear ($\mathscr{D}_m F$ must at least depend on the measure otherwise we find the standard Fokker-Planck PDE) Fokker-Planck equation:

$$\partial_t m = \nabla \cdot \Big( (\mathscr{D}_m F(m,\cdot) + \frac{\sigma^2}{2}\nabla V) m + \frac{\sigma^2}{2}\nabla m \Big). \tag{1.14}$$

**From two-body to many-body interactions.** Depending on the form of the drift and diffusion coefficients, the McKean-Vlasov diffusion can be used in a wide range of *modelling problems*. The first case is obtained when $b$ and $\sigma$ depend linearly on the measure argument. Namely, for $n, m \in \mathbb{N}$, let us consider two functions $K_1 : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}^n$, $K_2 : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}^m$, and let us take $b(x,\mu) = \overline{b}(x, K_1 \star \mu(x))$, $\sigma(x,\mu) = \overline{\sigma}(x, K_2 \star \mu(x))$, where $\overline{b} : \mathbb{R}^d \times \mathbb{R}^n \to \mathbb{R}^d$, $\overline{\sigma} : \mathbb{R}^d \times \mathbb{R}^m \to \mathscr{M}_d(\mathbb{R})$ and $K_i \star \mu(x) := \int K_i(x,y)\mu(dy)$. When $K_1, K_2$ and $\overline{b}, \overline{\sigma}$ are Lipschitz and bounded, the propagation of chaos result is the given by McKean's theorem.
In many applications, $\sigma$ is a constant diffusion matrix, $K_1(x,y) \equiv K(y-x)$ for a (usually symmetric) radial kernel $K : \mathbb{R}^d \to \mathbb{R}^d$ and $b(x,\mu) = K \star \mu(x)$. Note that the case where K has a singularity is much more delicate but contains many important cases (such as the Biot-Savart kernel or the 2D-incompressible Navier-Stokes model in fluid dynamics).
The case of *gradient systems* is an important sub-case when $\sigma(x,\mu) = \sigma \mathbf{Id}$ for a constant $\sigma > 0$ and

$$b(x,\mu) = -\nabla \mathbf{V}(x) - \int_{\mathbb{R}^d} \nabla W(x-y)\mu(dy) \tag{1.15}$$

where $\mathbf{V}, W$ are two potentials on $\mathbb{R}^d$ respectively called the *confinement potential* and the *interaction potential*. The limit Fokker-Planck equation

$$\partial_t \mu_t = \frac{\sigma^2}{2}\Delta \mu_t + \nabla \cdot \Big( \mu_t \nabla (\mathbf{V} + W \star \mu_t) \Big), \tag{1.16}$$

is called the *granular-media* equation. The above models are two-body interactions. This is characterized by the fact that $K_1$ or $K_2$ depend on only two variables or equivalently by the fact the functional $F : \mu \mapsto \int K_1(x,y)\mu(dx)\mu(dy)$ is a polynomial of degree two. Nevertheless, in some other models, one may find some interactions which involve more than two particles. This is for instance the case of the Skyrme model (see [8]). This is why in this paper, we choose to consider a polynomial dependence in the measure $\mu$ induced by *order statistics (many-body interaction)* in order to generalize the results obtained in the case of a linear interaction in the measure $\mu$ defined by the convolution via a potential two-body interaction ( [3], [2]). More exactly, under adequate assumptions (see $\mathbb{HMV}$3.1, $\mathbb{VFP}$3.2), we are interested in the exponential return to equilibrium of the solution of Eq. (1.14) in the case

$$F(\mu) = \sum_{k=2}^{N} \int W^{(k)} d\mu^{\otimes k}, \tag{1.17}$$

where $\forall k \in \{2, \dots, N\}$, $W^{(k)}$ is a *symmetric interaction potential* between $k$ particles and N represents the number of such potentials. The intrinsic derivative $\mathscr{D}_m F(\nu, y)$ associated with this functional is given by

$$\nabla \frac{\delta F}{\delta m}(\nu, y) = \sum_{k=2}^{N} \sum_{j=1}^{k} \int \nabla_{x_j} W^{(k)}(x_1, \dots, x_{j-1}, y, x_{j+1}, \dots, x_k) \nu^{\otimes k-1}(dx_1, \dots, dx_{j-1}, dx_{j+1}, \dots, dx_k) \tag{1.18}$$

The associated microscopic (particle-level) interaction is given by (U−statistic of order $k$ and kernel $\Phi \equiv W^{(k)}$)

$$U_n(W^{(k)}) := \frac{k!(n-k)!}{n!} \sum_{1\le i_1<\ldots<i_k\le n} W^{(k)}(X^{i_1,n},\ldots,X^{i_k,n}), \quad \text{where } X^n := (X^{1,n},\ldots,X^{n,n}) \in (\mathbb{R}^D)^n. \tag{1.19}$$

$U(X^n) := U_n(\Phi)$ is called U−statistic of order $k$ and kernel $\Phi$ associated with the sample $X^n$. This statistic corresponds to the arithmetic mean of the kernel $\Phi$ over all the parts at $k$ elements of the set of sample values. We will often write $U_n(W^{(n)})(X^n) =: U(X^n)$. We generalize this definition to the space of probabilities by the functional

$$\mu \in \mathscr{P}(\mathbb{R}^D) \longmapsto \int_{\mathbb{R}^{kD}} \Phi d\mu^{\otimes k}, \tag{1.20}$$

called monome of degree $k$ and coefficient $\Phi$ on the probability space $\mathscr{P}(\mathbb{R}^D)$. The link between these two microscopic and macroscopic interactions is given by

$$\sum_{k=2}^{N} U_n(W^{(k)}) = F(\mu_{X^n}). \tag{1.21}$$

*Remark* 1.2. Note that the *granular-media equation* (1.16) is a particular case of (1.14) with

$$F(\mu) = \int W^{(2)}(x,y)\mu(dx)\mu(dy) \quad W^{(2)}(x,y) = \frac{1}{2}W(x-y) \quad V(x) \equiv \frac{2}{\sigma^2}\mathbf{V}(x). \tag{1.22}$$

Indeed, in this case, we have

$$\frac{\delta F}{\delta m}(\mu,x) = \int W^{(2)}(x,y)\mu(dy) + \int W^{(2)}(y,x)\mu(dy) = \int W(x-y)\mu(dy) =: W \star \mu(x), \tag{1.23}$$

so that

$$\mathscr{D}_m F(\mu,x) = \nabla \frac{\delta F}{\delta m}(\mu,x) = \nabla W \star \mu(x).$$

**Energy and Large Deviations.** Consider $\mathscr{M}_1^p(\mathbb{R}^D)$ defined in Section 2, $G : \mathscr{M}_1^p(\mathbb{R}^D) \to \overline{\mathbb{R}}$ (which can be nonlinear) and the probability (Gibbs) measure $\alpha$ related to V, *i.e.* $\alpha(dx) = Z_V^{-1} e^{-V(x)}dx$ with $Z_V = \int e^{-V(x)}dx$ (where $Z_V$ is assumed to finite). For any $\sigma > 0$, we put

$$V^{\sigma,G}(m) := G(m) + \frac{\sigma^2}{2}\mathbf{H}[m|\alpha]. \tag{1.24}$$

$V^{\sigma,G}$ is an energy function regularised by the **KL**−divergence $\mathbf{H}[m|\alpha]$ which is given by Eq. (2.3) in Section 2. It is known (see e.g. [9, Proposition.2.5]) that $V^{\sigma,G}$ is minimized by a measure $m^{\sigma,\star}$ satisfying the following fixed point problem (it is noteworthy that the variational form of the invariant measure of the classic Langevin equation is a particular example of this first order condition)

$$m^{\sigma,\star}(dx) = \frac{1}{Z_\sigma} e^{-\frac{2}{\sigma^2}\left(\frac{\delta G}{\delta m}(m^{\sigma,\star},x) + \frac{\sigma^2}{2}V(x)\right)} dx, \tag{1.25}$$

where $Z_\sigma$ is the normalising constant, and for any $m \in \mathscr{M}_1^p(\mathbb{R}^D)$ and $x \in \mathbb{R}^D$, $\frac{\delta G}{\delta m}(m,x)$ denotes a flat derivative of G with respect to $m$, in the direction of $x$, evaluated at $m$. For any $\Theta_0, \Theta_1 \in \mathscr{M}_1^p(\mathbb{R}^D)$, the function $\frac{\delta G}{\delta m} : \mathscr{M}_1^p(\mathbb{R}^D) \times \mathbb{R}^D \to \mathbb{R}$ satisfies

$$G(\Theta_1) - G(\Theta_0) = \int_0^1 \int_{\mathbb{R}^D} \frac{\delta G}{\delta m}(\Theta_0 + \lambda(\Theta_1 - \Theta_0), x)(\Theta_1 - \Theta_0)(dx)\lambda_{\mathbb{R}}(d\lambda). \tag{1.26}$$

This notion of derivative appears in the literature under several different names, including the linear functional derivative (see e.g [10, Section.5.4.1]) or the first variation [11].

*Remark* 1.3. It is important to note that we have uniqueness modulo the choice of a function of the measure. The McKean-Vlasov SDE given by Eq. (1.2) (therefore the associated PDE given by Eq. (1.14) and the invariant measure given by Eq. (1.25)) does not depend on the choice of the function of the measure in the calculation of a flat derivative: we can therefore do the calculations with any flat derivative. By convexity of $\mathscr{M}_1^p(\mathbb{R}^d)$, for all $t \in [0,1]$, $\Theta_t := (1-t)\Theta_0 + t\Theta_1 \in [\Theta_0, \Theta_1] \subset \mathscr{M}_1^p(\mathbb{R}^d)$. Eq. (1.26) is equivalent to deriving the functional G along the end segment $\Theta_0$ and $\Theta_1$ parameterized by the path $t \in [0,1] \mapsto \Theta_t$:

$$\frac{d}{dt}G(\Theta_t) = \int_{\mathbb{R}^D} \frac{\delta G}{\delta m}(\Theta_t, x)\partial_t \Theta_t(dx). \tag{1.27}$$

In practice, this reformulation via derivation along paths lends itself better to calculations. For example, it is easy to check

$$\nabla \frac{\delta F}{\delta m}(\nu, y) = \sum_{k=2}^{N} k \int \nabla_{x_1} W^{(k)}(y, z) \nu^{\otimes k-1}(dz); \tag{1.28}$$

$$\mathscr{D}_m \mathbf{H}[\cdot|\alpha](\nu, y) = \nabla \log\left(\frac{d\nu}{d\alpha}\right)(y). \tag{1.29}$$

In [1], the authors showed a *Sanov theorem* by *large deviations* on U-statistics in which $V^{\sqrt{2},F}$ is a *good convergence rate function*, which implies that $V^{\sqrt{2},F}$ admits a *minimizer*: inf-compactness of the mean-field energy functional. As for *uniqueness*, it is not assured: in *statistical physics*, we speak of *phase transition*. We ensure uniqueness by a *contraction property* (see $\mathbb{HMV}$3.1) on the *invariant measure application* induced by Eq. (1.25). Large Deviation Principles imply propagation of chaos, but they do not always give a way to quantify it since the related results are often purely asymptotic (for instance, Sanov theorem is non-quantitative). Nevertheless, the results of large deviations turn out to be very useful for the technical passages in the macroscopic limits: when one makes tend the number of particles to infinity. In the seminal article [12], the authors improve results from [13] and [14] on Large Deviation Principles (LDP) for Gibbs measures and obtain as a byproduct a pathwise propagation of chaos result for the McKean-Vlasov diffusion. Firstly, [12, Theorem.A] (or Theorem A.3) states a large deviation principle for Gibbs measures with a polynomial potential. [12, Theorem.B] quantifies the fluctuations of $\mu_{X^n}$ in the non-degenerate case. Analogous results for the degenerate case are given in [12, Theorem.C]. For more details, see also [7, Theorem.4.7, Corollary.3]. We use the large deviations results obtained on the order statistics in [1]: In addition to the fact that the mean-field entropy functional (Eq. (A.62) or $V^{\sqrt{2},F}$ defined by Eq. (1.24)) is a rate function (Theorem A.3) for the random empirical measure $\mu_{X^n}$, the authors show that it is a good rate function that has good tensorization properties.

**Long time behavior.** In the present paper, we are concerned by the long-time convergence towards the solution to an optimization problem on the subspace $\mathscr{M}_1^p(\mathbb{R}^D)$ of probability measures $\mathscr{M}_1(\mathbb{R}^D)$: we consider a function $\mathbf{E} : \mathscr{M}_1^p(\mathbb{R}^D) \to \overline{\mathbb{R}}$ and we want to find a minimizing measure $m^\star := \mathbf{arginf}_{\mathscr{M}_1^p(\mathbb{R}^D)} \mathbf{E}$ such that for a *gradient flow* (see e.g. [11] and [15]) $(m_t)_{t\geq 0}$ associated with $\mathbf{E}$, we have an exponential estimate of the deviation $\mathbf{E}(m_t) - \mathbf{E}(m^\star)$ of the form (with $C \geq 1$ and $\rho > 0$)

$$\mathbf{E}(m_t) - \mathbf{E}(m^\star) \leq C(\mathbf{E}(m_0) - \mathbf{E}(m^\star))e^{-\rho t}. \tag{1.30}$$

Eq. (1.30)-type Inequalities are called *hypocoercive inequalities*. We call $\mathbf{E} - \mathbf{E}(m^\star)$ the *entropy functional* ([16],[17],[18]) of the system and $-\frac{d}{dt}(\mathbf{E}(m_t) - \mathbf{E}(m^\star))$ the *production of entropy* (usually called energy in mathematical literature). Clausius invents the concept of entropy, Boltzmann proposes to derive entropy along the flow. Generally speaking, an entropy is a *Lyapunov functional* of a specific form. It is however hard (and even somewhat artificial) to give a formal narrow definition of entropies that distinguishes them from, say, energies. An entropy is a quantity calculated from a solution, which decreases over time when the solution obeys an evolution equation, and which is stationary only for the stationary solutions of the equation. In conclusion, the concept of entropy is a tool that adapts to what we want to study. The notion of *hypocoercivity* was proposed by T. Gallay. The objective is typically to control the entropy at time $t$ by the initial entropy multiplied by a constant C (always greater than 1) and a exponential decay factor, with exponential decay rate as good as possible in big time. This theory is inspired by the *hypoelliptic theory of L. Hormander*, and the terminology hypocoercivity accounts for the relationship between entropy and its derivative with respect to $t$. There would be *coercivity* if $C = 1$, which is clearly not possible in most cases considered in kinetic theory. It is well known that, for the standard *Langevin equation* of Hamiltonian V (given by Eq. (1.2) in the case $F \equiv 0$), for $\rho > 0$, the following assertions are equivalent:

$$\forall \varphi \in \mathscr{C}_c^\infty(\mathbb{R}^d), \quad \rho \mathbf{Ent}_\alpha[\varphi^2] \leq 2 \int ||\nabla \varphi||^2 d\alpha. \tag{1.31}$$

$$\rho \mathbf{H}[\cdot|\alpha] \leq 2\mathbf{I}[\cdot|\alpha]. \tag{1.32}$$

$$\forall t \geq 0, \quad \mathbf{H}[\mu_t^V|\alpha] \leq \mathbf{H}[\mu_0^V|\alpha]e^{-\rho t}. \tag{1.33}$$

These three equivalent assertions imply the T2−*Talagrand inequality*

$$\rho \mathscr{W}_2^2(\cdot, \alpha) \leq 2\mathbf{H}[\cdot|\alpha], \tag{1.34}$$

inequality which, in turn, implies an *exponential contraction in wasserstein metric* $\mathscr{W}_2$, i.e. the exponential convergence of the flow $(\mu_t^V)_{t\geq 0}$ (solution of the Fokker-Planck equation associated with the standard Langevin process of

Hamiltonian V) to the maxwellian (invariant measure of the Langevin process that can also be seen from equivalently as the unique $\mathbf{argmin}_{\mathscr{P}(\mathbb{R}^d)}\mathbf{H}[\cdot|\alpha]$) $\alpha$ of the *Fokker-Planck PDE* given by Eq. (1.14) in the case $\mathrm{F}\equiv 0$:

$$\forall t\geq 0,\quad \mathscr{W}_2^2(\mu_t^{\mathrm{V}},\alpha)\leq \frac{2}{\rho}\mathbf{H}[\mu_0^{\mathrm{V}}|\alpha]e^{-\rho t}. \tag{1.35}$$

Eq. (1.31) and Eq. (1.32) respectively define the *logarithmic Sobolev inequality* ([17]) and its *dual version*. According to the *dimension curvature criterion of Bakry-Emery*, we have

$$\left(\exists\rho>0\quad\forall(x,h)\in\mathbb{R}^d\times\mathbb{R}^d,\quad\langle\nabla^2\mathrm{V}(x)h,h\rangle\geq\rho||h||_2^2\right)\Longrightarrow Eq.\ (1.31). \tag{1.36}$$

Note that in the case of the symmetric Langevin-Kolmogorov process, we have

$$m_t=\mu_t^{\mathrm{V}},\quad m^\star=\alpha,\quad \mathbf{E}=\mathbf{H}[\cdot|\alpha],\quad \mathbf{E}(m_t)-\mathbf{E}(m^\star)=\mathbf{H}[\mu_t^{\mathrm{V}}|\alpha], \tag{1.37}$$

$$-\frac{d}{dt}(\mathbf{E}(m_t)-\mathbf{E}(m^\star))=-\frac{d}{dt}\mathbf{H}[\mu_t^{\mathrm{V}}|\alpha]=\mathbf{I}[\mu_t^{\mathrm{V}}|\alpha]. \tag{1.38}$$

The objective of this work is to identify a flow of measures $(m_t^{\sigma,\mathrm{F}})_{t\geq 0}$ (flow solution of Eq. (1.14)) such that

$$\mathrm{V}^{\sigma,\mathrm{F}}(m_t^{\sigma,\mathrm{F}})-\mathrm{V}^{\sigma,\mathrm{F}}(m^{\sigma,\star})\overset{t\to+\infty}{\longrightarrow}0, \tag{1.39}$$

as well as conditions (HMV3.1, VFP3.2) that ensure that this convergence is exponential. To this end, we equip the space $\mathscr{M}_1^p(\mathbb{R}^{\mathrm{D}})$ with a suitable distance function $\mathbf{d}:\mathscr{M}_1^p(\mathbb{R}^{\mathrm{D}})\times\mathscr{M}_1^p(\mathbb{R}^{\mathrm{D}})\to\mathbb{R}_+$ and consider a corresponding gradient flow, where the form of the flow is dictated by the choice of $\mathbf{d}$. Such a problem has been dealt with in the case of the Fisher-Rao metric (see [19]): the authors established from a Polyak-Lojasiewicz inequality the exponential convergence of the gradient flow $(m_t^{\sigma,\mathrm{G}})_{t\geq 0}$ described by the birth-death equation along $\mathrm{V}^{\sigma,\mathrm{G}}$ towards $\mathrm{V}^{\sigma,\mathrm{G}}(m^{\sigma,\star})$. In our case, Eq. (1.30) implies the exponential decay in $\mathbf{d}$-metric (transport distance):

$$\mathbf{d}(m_t^{\sigma,\mathrm{F}},m^{\sigma,\star})\leq\gamma(\mathrm{V}^{\sigma,\mathrm{F}}(m_0^{\sigma,\mathrm{F}})-\mathrm{V}^{\sigma,\mathrm{F}}(m^{\sigma,\star}))e^{-\rho t}. \tag{1.40}$$

Eq. (1.40) is a consequence of transport inequalities (see [20]). Moreover, given a measure $m^{\sigma,\star}$ satisfying the first order condition Eq. (1.25), it is formally a stationary solution to Eq. (1.14) called the *Maxwellian of the McKean-Vlasov PDE*. Therefore, formally, we have already obtained the correspondence between the *minimiser of the free energy function* and the *invariant measure* of Eq. (1.2). In this paper, the connection is rigorously proved mainly with a probabilistic argument. The study of stationary solutions to nonlocal, diffusive Eq. (1.14) is classical topic with it roots in statistical physics literature and with strong links to Kac's program in Kinetic theory [21]. We also refer reader to the excellent monographs [11] and [22]. An important issue is the *long-time behaviour of gradient systems* which is often studied under *convexity assumptions* on the potentials. In particular, variational approach has been developed in [23] and [15] where authors studied *dissipation of entropy* for granular media equations Eq. (1.16) with the symmetric interaction potential of convolution type (interaction potential corresponds to term $\mathscr{D}_m\mathrm{F}$ in Eq. (1.14)). Following on from the work done in [15] and [23] (among others) on the long-time behavior of Eq. (1.16), in [2], the authors proved via a *uniform logarithmic Sobolev inequality* in the number of particles that

$$\forall t\geq 0,\quad \mathrm{H_W}[\nu_t]\leq\mathrm{H_W}[\nu_0]e^{-\rho_{\mathrm{LS}}\frac{t}{2}}\quad\text{and}\quad\mathscr{W}_2^2(\nu_t,\nu_\infty)\leq\frac{2}{\rho_{\mathrm{LS}}}\mathrm{H_W}[\nu_0]e^{-\rho_{\mathrm{LS}}\frac{t}{2}}. \tag{1.41}$$

Eq. (1.41) translates the exponential decrease of the *mean field entropy* $\mathrm{H_W}$ (given by Eq. (1.30) with $\mathbf{E}=\mathrm{V}^{\sigma,\mathrm{F}}$) and the *contraction* in Wasserstein metric ($\mathbf{d}=\mathscr{W}_2$) of the solution flow of Eq. (1.14) in the case

$$\sigma=\sqrt{2}\quad\text{and}\quad\mathrm{F}(\mu)=\frac{1}{2}\int\mathrm{W}(x,y)\mu(dx)\mu(dy). \tag{1.42}$$

The study of the long-time behaviour for the VFP equation is often more difficult than that of the McKean-Vlasov equation because of two reasons:

(i) it is a degenerate diffusion process where the Laplacian acts only on the volocity variable and;

(ii) it is not a gradient flows but simultaneously presents both Hamiltonian and gradient flows effects.

In [4], combining the results of [2] and [5], the trend to equilibrium in large time is studied for a large particle system (given by Eq. (3.16) in case of a two-body interaction) associated to a *Vlasov-Fokker-Planck equation* by the authors: they showed that under some conditions (that allow *non-convex confining potentials*), the convergence rate is proven to be independent from the number of particles. From this are derived *uniform in time propagation of chaos estimates* and an *exponentially fast convergence for the nonlinear equation itself.*

**Contributions.** In this paper, we are going to prove in a polynomial interaction setting

(i) propagation of chaos in Wasserstein's $\mathscr{W}_2$-metric for our particle systems given by Eq. (3.5) and Eq. (3.16).

(ii) entropic convergence to equilibrium for the nonlinear McKean-Vlasov SDE (mean field limit of the first order system given by Eq. (3.5)) generalizing results (given in Eq. (1.41) of [2]).

(iii) by *Villani's hypocoercivity theorem* (see e.g. [3, Theorem.3] or [24, Theorem.35]) the $H^1-$convergence for the kinetic Fokker-Planck equation with mean field interaction given by Eq. (3.16).

(iv) exponential convergence towards equilibrium in metric $\mathscr{W}_2-$Wasserstein for the flow solution of the Vlasov-Fokker-Planck equation: mean field limit of the second order system given by Eq. (3.16).

In the spirit of [2], we give *probabilistic proofs* (see Section 5, Fig. 1 and Section 6) based mainly on the *propagation of chaos* (see Theorem 5.1); the **H**-tensorization (see Proposition 5.10) and inf-compactness given by *large deviations principle*; the **I**-tensorization (see Proposition 5.11) given by *law of large numbers*; the *uniform log-Sobolev inequality* (see Theorem 5.3); and the *uniqueness argument* (see Proposition 5.9). In the kinetic case, we need additional results such as Villani's hypocoercivity ([3, Theorem.3] or [24, Theorem.18 and Theorem.35]) theorem (see Proposition 5.15) and *Hormander's form* (see e.g. respectively Theorem.7 and Theorem.10 in [5, [6]]). The fact that the interaction is polynomial is important in calculations, among other things, for passing to the limit in the number of particles: technical passage to the limit given, among others, by LDP.

**Plan of the paper.** Let us finish this introduction by the plan of the paper. In the next three sections, we will present our mean field systems (Eq. (3.5),Eq. (3.16)), our set of assumptions ($\mathbb{HMV}$3.1,$\mathbb{VFP}$3.2), the main results (and examples) (in Section 4) of the paper concerning logarithmic Sobolev inequality of mean field particles systems as well as exponential convergences to equilibrium for McKean-Vlasov (Theorem 4.1,Theorem 4.2), kinetic Fokker-Planck (Theorem 4.3) and Vlasov-Fokker-Planck (Theorem 4.4) SDEs. In Section 5, we sketch a proof of our results and we introduce the pre-proof tools. In Section 6, we prove our main results. And we end the paper with the appendix, the acknowledgments and the bibliographical references.

## 2 Notations and Definitions

We try to keep coherent definitions and notations throughout the article, but as the various objects and what they represent may become confusing, we list them here for reference :

**Notations.** For all $(u,v)\in\mathbb{R}^d\times\mathbb{R}^d$, we note $u\otimes v:=uv^{\mathrm{T}}=(u_iv_j)_{1\le i,j\le d}$ the tensor product matrix of two vectors and $u\cdot v:=u^{\mathrm{T}}v$ the standard Euclidean scalar product of two vectors. We note $||\cdot||_{\mathbf{op}}$ the matrix subordinate norm to the Euclidean norm which we will note indifferently $||\cdot||_2$ or $|\cdot|$. $\langle\cdot,\cdot\rangle$ represents indifferently the scalar product and the duality bracket. We note $|||\cdot|||_{H^1\to H^1}$ the operator norm associated with the weighted Sobolev $H^1(\mu_Z^n)$ space induced by the invariant measure $\mu_Z^n$ of our second-order system given by Eq. (3.16). We have

$$H^1(\mu_Z^n):=\left\{\varphi\in L^2(\mu_Z^n),\quad \nabla\varphi\in(L^2(\mu_Z^n))^n\right\},\quad ||\varphi||_{H^1}^2:=||\varphi||_{L^2(\mu_Z^n)}^2+\int\left(||\nabla_x\varphi||_2^2+||\nabla_v\varphi||_2^2\right)d\mu_Z^n. \tag{2.1}$$

$(B_t)_{t\ge0}$ represents the standard Brownian motion. We consider $((B^i)_{t\ge0})_{i\in\{1,\dots,n\}}$ $n$ independent copies of $(B_t)_{t\ge0}$. For all $n\ge1$, $\mathfrak{S}_n$ is the n-th symmetric group. For all $p\in[1+\infty)$, the Wasserstein $p$-distance between two probability measures $\mu$ and $\nu$ on $\mathbb{R}^D$ with finite $p$-moments is given by

$$\mathscr{W}_p(\mu,\nu):=\left(\inf_{\gamma\in\Gamma(\mu,\nu)}\int_{\mathbb{R}^D\times\mathbb{R}^D}|x-y|^p\gamma(dxdy)\right)^{\frac{1}{p}}, \tag{2.2}$$
$$\Gamma(\mu,\nu):=\left\{\gamma\in\mathscr{P}(\mathbb{R}^D\times\mathbb{R}^D),\quad \pi_1\gamma=\mu\quad\text{and}\quad\pi_2\gamma=\nu\right\}.$$

We note $\mathscr{M}_1^p(\mathbb{R}^D)$ the space of probability measures with finite $p-$moments.

**Definitions.**
*Relative entropy:* Let $\mu\in\mathscr{P}(\mathbb{R}^D)$. We define $\mathbf{H}[\cdot|\mu]:\mathscr{P}(\mathbb{R}^D)\longrightarrow[0,+\infty]$ such that

$$\mathbf{H}[\nu|\mu]=\begin{cases}\mathbb{E}_\nu[\log\frac{d\nu}{d\mu}]=:\ \ \mathbf{Ent}_\mu[\frac{d\nu}{d\mu}] & \text{if } \nu\ll\mu,\\ +\infty & \text{otherwise.}\end{cases} \tag{2.3}$$

And we recall that in the first case of absolute continuity, $\frac{d\nu}{d\mu}$ is the *Radon-Nikodym density* of ν with respect to μ.
*Relative Fisher information:* We also define the *Fisher-Donsker-Varadhan information* of ν with respect to μ by:

$$\mathbf{I}[\nu|\mu] = \int \left|\left|\nabla\sqrt{\frac{d\nu}{d\mu}}\right|\right|^2 d\mu = \frac{1}{4}\int \left|\left|\nabla \log\frac{d\nu}{d\mu}\right|\right|^2 d\nu = \frac{1}{4}\int \left|\left|\nabla \frac{\delta \mathbf{H}[\cdot|\mu]}{\delta m}(\nu,y)\right|\right|^2 \nu(dy) \tag{2.4}$$

if $\nu \ll \mu$ and $\sqrt{\frac{d\nu}{d\mu}} \in \mathbf{H}^1_\mu$, and $\mathbf{I}[\nu|\mu] = +\infty$ otherwise. $\mathbf{H}^1_\mu$ is the domain of the Dirichlet form

$$\mathscr{E}_\mu : g \longmapsto \int ||\nabla g||^2 d\mu. \tag{2.5}$$

**UPI.** We say that $\mu(dx) := \frac{1}{Z}e^{-\mathrm{H}(x)}dx$ (Gibbs probability measure of hamiltonian $\mathrm{H}:\mathbb{R}^{n\mathrm{D}} \to \mathbb{R}$) satisfies a uniform Poincaré inequality if

$$\exists \lambda > 0 \quad \forall n \geq 2 \quad \forall \varphi \in \mathscr{C}_c^\infty(\mathbb{R}^{n\mathrm{D}}), \quad \lambda \mathbb{V}_\mu[\varphi] \leq \mathbb{E}_\mu[||\nabla\varphi||^2]. \tag{2.6}$$

And we call Poincaré constant the best constant $\lambda_1(\mu)$ for which we have such an inequality.
**ULSI.** We say that μ satisfies a uniform logarithmic Sobolev inequality if

$$\exists \rho > 0 \quad \forall n \geq 2 \quad \forall \varphi \in \mathscr{C}_c^\infty(\mathbb{R}^{n\mathrm{D}}), \quad \rho\mathbf{Ent}_\mu[\varphi^2] \leq \mathbb{E}_\mu[||\nabla\varphi||^2]. \tag{2.7}$$

And the best constant $\rho_{\mathrm{LS}}(\mu)$ for which such an inequality holds is called the logarithmic Sobolev constant.
*Remark* 2.1. We recall that

$$\textbf{ULSI.} \Longrightarrow \textbf{UPI.} \tag{2.8}$$

The Poincaré and log-Sobolev inequalities for μ are equivalent to exponential decreases of the semigroup $(\mathrm{P}_t)_{t\geq 0}$ respectively in variance and in entropy, i.e.

- ▷ **Poincaré**

$$\forall \quad f \in \mathrm{L}^2(\mu) \quad t \geq 0, \quad ||\mathrm{P}_t f - \langle \mu, f\rangle||_{\mathrm{L}^2(\mu)} \leq e^{-\lambda_1(\mu)t}||f - \langle\mu, f\rangle||_{\mathrm{L}^2(\mu)}. \tag{2.9}$$

- ▷ **Log-Sobolev**

$$\forall \quad f \in \mathrm{L}^1(\mu)\log\mathrm{L}^1(\mu) \quad t \geq 0, \quad \mathbf{Ent}_\mu[\mathrm{P}_t f] \leq e^{-\rho_{\mathrm{LS}}(\mu)t}\mathbf{Ent}_\mu[f]. \tag{2.10}$$

  Here, the notation $\mathrm{L}^1(\mu)\log\mathrm{L}^1(\mu)$ denotes the entropy definition domain under μ.

We say that μ satisfies a $\mathrm{T}_p$−transport (Talagrand) inequality if there exists $\alpha > 0$ such that $\mathscr{W}_p(\cdot,\mu) \leq \sqrt{\alpha\mathbf{H}[\cdot|\mu]}$.
*Remark* 2.2. Moreover, as with the Poincaré and log-Sobolev inequalities, $\mathrm{T}_2$-inequality implies the T1-inequality : by definition and Cauchy-Schwarz inequality, we have

$$\mathscr{W}_1(\mu,\nu) := \inf_{\mathrm{X}\sim\mu,\mathrm{Y}\sim\nu} \mathbb{E}[||\mathrm{X}-\mathrm{Y}||] \leq \inf_{\mathrm{X}\sim\mu,\mathrm{Y}\sim\nu}\sqrt{\mathbb{E}[||\mathrm{X}-\mathrm{Y}||^2]} =: \mathscr{W}_2(\mu,\nu). \tag{2.11}$$

The class of probabilities verifying $\mathrm{T}_1$-inequality is identical to that having an exponential moment of finite order 2. The $\mathrm{T}_2$-inequality is significantly more structured than the $\mathrm{T}_1$-inequality since it involves a spectral gap inequality.

## 3 Mean-Field Systems and Assumptions

Throughout the paper, we consider a *confinement potential of a particle* $\mathrm{V}:\mathbb{R}^d \longrightarrow \mathbb{R} \in \mathscr{C}^2(\mathbb{R}^d)$ and N *interaction potentials* such that

$$\forall k \in \{2,\ldots,\mathrm{N}\}, \quad \mathrm{W}^{(k)} : (\mathbb{R}^d)^k \longrightarrow \mathbb{R} \in \mathscr{C}^2((\mathbb{R}^d)^k). \tag{3.1}$$

We recall that $\forall \sigma \in \mathfrak{S}_k$ and $\forall x = (x_1,\ldots,x_k)$,

$$\mathrm{W}^{(k)}(\sigma\cdot x) = \mathrm{W}^{(k)}(x), \quad \alpha(dx) := \frac{1}{\mathrm{C}}e^{-\mathrm{V}(x)}dx, \quad \mathrm{U}_n(\mathrm{W}^{(k)}) := \frac{1}{|\mathrm{I}_n^k|}\sum_{(i_1,\ldots,i_k)\in\mathrm{I}_n^k}\mathrm{W}^{(k)}(x_{i_1},\ldots,x_{i_k}), \tag{3.2}$$

where $\mathrm{I}_n^k := \{(i_1,\ldots,i_k)\in\mathbb{N}^k | i_p \neq i_q, \quad 1 \leq i_p \leq n\}$ is the set of possible arrangements of $k$ integers of the set of $n$ first nonzero integers, which gives $|\mathrm{I}_n^k| = \mathrm{A}_n^k := \frac{n!}{(n-k)!}$. We define $\mathrm{W}^{(k),-} := \max(-\mathrm{W}^{(k)},0)$ and $\mathrm{W}^{(k),+} := \max(\mathrm{W}^{(k)},0)$ the negative and positive parts of $\mathrm{W}^{(k)}$. ∀μ such that $\mathrm{W}^{(k),-} \in \mathrm{L}^1(\mu^{\otimes k})$,

$$\mathbf{W}^{(k)}[\mu] := \mathbb{E}_{\mu^{\otimes k}}[\mathrm{W}^{(k)}] = \mathbb{E}_{\mu^{\otimes k}}[\mathrm{W}^{(k),+}] - \mathbb{E}_{\mu^{\otimes k}}[\mathrm{W}^{(k),-}]. \tag{3.3}$$

### 3.1 Our Systems

**First order case.** We consider the *microscopic mean-field many-body interaction energy* given by

$$\mathrm{H}_n(x_1,\dots,x_n) := \sum_{j=1}^{n} \mathrm{V}(x_j) + n\sum_{k=2}^{\mathrm{N}} \mathrm{U}_n(\mathrm{W}^{(k)}). \tag{3.4}$$

The (non-kinetic) McKean-Vlasov process is defined as the mean field limit (under adequate assumptions given below) of the sequence $(\mathrm{X}^n)_{n\geq \mathrm{N}}$ of Langevin-Kolmogorov process of Hamiltonian $\mathrm{H}_n$, i.e.: (N fixed)

$$\forall n \geq \mathrm{N}, \quad d\mathrm{X}_t^n = \sqrt{2}d\mathrm{B}_t - \nabla \mathrm{H}_n(\mathrm{X}_t^n)dt. \tag{3.5}$$

Let

$$\mathscr{L}_n := \Delta - \nabla \mathrm{H}_n \cdot \nabla \tag{3.6}$$

be the *infinitesimal generator* and $(\mathrm{P}_t^n)_{t\geq 0}$ the associated *semigroup* of unique invariant measure (under $\mathbb{HMV}$3.1 below), the *Gibbs measure*

$$\mu_n(dx) := \frac{1}{\mathrm{Z}_n} e^{-\mathrm{H}_n(x)}dx \quad \text{with} \quad \mathrm{Z}_n := \int_{(\mathbb{R}^d)^n} e^{-\mathrm{H}_n(x)}dx < +\infty \tag{3.7}$$

is the normalization constant (called *partition function*). Note that

$$\mu_n(dx) = \frac{\mathrm{C}^n}{\mathrm{Z}_n} e^{-n\sum_{k=2}^{\mathrm{N}}\mathrm{U}_n(\mathrm{W}^{(k)})}\alpha^{\otimes n}(dx). \tag{3.8}$$

Without interaction (i.e. $\forall k$, $\mathrm{W}^{(k)} \equiv 0$ or constant), $\mu_n = \alpha^{\otimes n}$ (i.e. the particles are independent). We denote

$$\mathrm{L}_n(x;\cdot) := \frac{1}{n}\sum_{i=1}^{n} \delta_{x_i}(\cdot) \tag{3.9}$$

the empirical measurement application. We know that under general conditions, by propagation of chaos ([25]), $\mathrm{L}_n(\mathrm{X}^n;\cdot)$ converges weakly towards the solution of the nonlinear partial differential equation of McKean-Vlasov associated with the system of particles. We define

$$\mu_n^*(dx) := e^{-n\sum_{k=2}^{\mathrm{N}}\mathrm{U}_n(\mathrm{W}^{(k)})}\alpha^{\otimes n}(dx) = \frac{\mathrm{Z}_n}{\mathrm{C}^n}\mu_n(dx). \tag{3.10}$$

The *macroscopic mean-field energy* is given by

$$\mathbf{E}_\mathrm{W}[\mu] := \begin{cases} \mathbf{H}[\mu|\alpha] + \sum_{k=2}^{\mathrm{N}} \mathbf{W}^{(k)}[\mu] & \text{if } \mathbf{H}[\mu|\alpha] < +\infty \text{ and } \mathrm{W}^{(k),-} \in \mathrm{L}^1(\mu^{\otimes k}), \\ +\infty & \text{otherwise.} \end{cases} \tag{3.11}$$

Let

$$\mathbf{dom}(\mathbf{H}_\mathrm{W}) := \Big\{\mu, \quad \mathbf{H}[\mu|\alpha] < +\infty \quad \text{and} \quad \forall k, \quad \mathrm{W}^{(k),-} \in \mathrm{L}^1(\mu^{\otimes k})\Big\}. \tag{3.12}$$

*Remark* 3.1. $\mathbf{H}_\mathrm{W} := \mathbf{E}_\mathrm{W} - \inf \mathbf{E}_\mathrm{W}$ is called the *mean field entropy*. We can prove that $\mathbf{H}_\mathrm{W}$ is *inf-compact* (Theorem 5.2) and that there is at least one minimizer usually called *equilibrium point*. From the point of view of statistical physics, $\mathbf{H}_\mathrm{W}$ is an entropy or *free energy* associated to the *nonlinear McKean-Vlasov equation* given by Eq. (3.5). The uniqueness of the minimizer means that there is no phase transition for the mean-field. Concerning the work on uniqueness in the case of peer interaction, we can cite among others: [2], [26] and [23]. These authors ([26],[23]) showed that $\mathbf{H}_\mathrm{W}$ is *strictly displacement convex* (i.e. along the $\mathscr{W}_2$-geodesic) under various sufficient conditions on the convexity of the confinement potential V and the pair interaction potential $\mathrm{W}^{(2)}$. In case of a many-body interaction, under assumptions in $\mathbb{HMV}$3.1, we prove in Proposition 5.9 the uniqueness: then we denote $\mu_\infty$ this minimizer.

Analogously, we define the *mean-field Fisher information* by:

$$\mathbf{I}_\mathrm{W}[\mu] := \frac{1}{4}\int \left|\left|\nabla \frac{\delta \mathbf{E}_\mathrm{W}}{\delta m}(\mu, y)\right|\right|^2 \mu(dy). \tag{3.13}$$

*Remark* 3.2. Without interaction ($\forall k$, $\mathrm{W}^{(k)} \equiv w_k$), we find the Lyapunov functionals associated with the standard symmetric Langevin-Kolmogorov process whose Hamiltonian is given by the confinement potential V. More precisely, in this case:

$$\mathbf{E}_\mathrm{W} = \mathbf{H}[\cdot|\alpha] + \sum_{k=2}^{\mathrm{N}} w_k, \quad \mathbf{H}_\mathrm{W} = \mathbf{H}[\cdot|\alpha] \quad \text{and} \quad \mathbf{I}_\mathrm{W} = \mathbf{I}[\cdot|\alpha]. \tag{3.14}$$

**Kinetic case.** Set

$$z := (x_1,\dots,x_n,v_1,\dots,v_n) \in \mathbb{R}^{2nd}, \quad H_n^Z(z) = \frac{1}{2}\sum_{j=1}^{n} |v_j|^2 + 2V(x_j) + n\sum_{k=2}^{N} U_n(W^{(k)}) \tag{3.15}$$

and $Z^n := (X^{n,1},\dots,X^{n,n},V^{n,1},\dots,V^{n,n}) \in (\mathbb{R}^d\times\mathbb{R}^d)^n$ such that

$$\begin{cases} dX_t^{n,i} = \nabla_{v_i} H_n^Z(Z_t^n)dt \\ dV_t^{n,i} = -\left(\nabla_{x_i} H_n^Z(Z_t^n) + \nabla_{v_i} H_n^Z(Z_t^n)\right)dt + \sqrt{2}dB_t^i. \end{cases} \tag{3.16}$$

We are going to study the long-time behavior of the mean-field limit of the Langevin process $(Z_t^n)_{t\ge0}$ of Hamiltonian $H_n^Z(x,v) := S_{1,n}(x) + S_{2,n}(v)$ with $S_{1,n}$ is none other than the Hamiltonian $H_n$ of the McKean-Vlasov case and $S_{2,n}$ the velocity part ($S_{2,n} := H_n^Z - S_{1,n}$). Invariant measure of the Langevin process is given by

$$\mu_Z^n(dxdv) = \frac{1}{\tilde{C}} e^{-H_n^Z(z)} dxdv = \frac{1}{C_{1,n}} e^{-S_{1,n}(x)} dx \frac{1}{C_{2,n}} e^{-S_{2,n}(v)} dv = \mu_{1,n} \bigotimes \mu_{2,n}(dxdv). \tag{3.17}$$

And the parabolic PDE in the sense of the distributions associated with this Kolmogorov-Fokker-Planck SDE is:

$$\partial_t\mu = \Delta_v\mu + \nabla S_{2,n}\cdot\nabla_v\mu - \nabla S_{1,n}\cdot\nabla_v\mu + \nabla S_{2,n}\cdot\nabla_x\mu = \Delta_v\mu + v\cdot\nabla_v\mu - \nabla S_{1,n}\cdot\nabla_v\mu + v\cdot\nabla_x\mu = \mathscr{L}_{Z,n}^{\dagger}\mu \tag{3.18}$$

with

$$\mathscr{L}_{Z,n} := \Delta_v - v\cdot\nabla_v + \nabla S_{1,n}\cdot\nabla_v - v\cdot\nabla_x \tag{3.19}$$

the generator of the strongly continuous semigroup $(P_t^{Z,(n)})_{t\ge0}$ (if the *hessian* $\nabla^2 S_{1,n}$ is bounded, it is a Markovian semigroup defined by the Kolmogorov-Fokker-Planck SDE) and we note $\mathscr{L}_{Z,n}^{\dagger}$ adjoint in the sense of distributions. In other words, for any test function $\varphi \in \mathscr{C}_c^\infty((\mathbb{R}^d\times\mathbb{R}^d)^n)$, the function $(t,z)\longmapsto P_t^{Z,(n)}\varphi(z)$ is the unique solution of the Cauchy problem:

$$\begin{cases} \frac{\partial h}{\partial t} = \mathscr{L}_{Z,n}h, \\ h(0,\cdot) = \varphi. \end{cases} \iff \begin{cases} \frac{\partial \mu_t}{\partial t} = \mathscr{L}_{Z,n}^{\dagger}\mu_t, \\ \mu_0 = \delta_z. \end{cases} \tag{3.20}$$

Vlasov Fokker Planck free energy and associated mean field entropy are given by

$$\begin{aligned} \mathscr{E}[\mu] &:= \mathbf{H}[\mu|dxdv] + \frac{1}{2}\int_{\mathbb{R}^d\times\mathbb{R}^d} ||v||^2\mu(dxdv) + \sum_{k=2}^{N}\int_{(\mathbb{R}^d\times\mathbb{R}^d)^k} W^{(k)}d\mu^{\otimes k} + \int V(x)\mu(dxdv) \\ &= \mathbf{H}[\mu|\alpha\otimes\mathscr{N}(0,\mathbf{Id}_d)] + \sum_{k=2}^{N}\int_{(\mathbb{R}^d\times\mathbb{R}^d)^k} W^{(k)}d\mu^{\otimes k} \end{aligned} \tag{3.21}$$

and

$$\mathscr{S} := \mathscr{E} - \mathbf{Inf}\mathscr{E} = \mathscr{E} - \mathscr{E}[\mu_\infty^Z]. \tag{3.22}$$

They are *Lyapunov functionals for the Vlasov-Fokker-Planck partial differential equation* whose solutions are obtained as mean-field limits of our kinetic Fokker-Planck particle system given by Eq. (3.16). Mean Field Fisher Information for Vlasov-Fokker-Planck is given by $\left(A := \begin{pmatrix} 0 \\ \mathbf{Id}_d \end{pmatrix} \in \mathscr{M}_{2d,d}(\mathbb{R})\right)$

$$\mathscr{I}[\mu] := \int \left\langle \nabla_{x,v}\frac{\delta}{\delta m}\mathscr{E}(\mu,x,v), AA^*\nabla_{x,v}\frac{\delta}{\delta m}\mathscr{E}(\mu,x,v)\right\rangle\mu(dxdv) = \int \left|\left|\nabla_{x,v}\frac{\delta}{\delta m}\mathscr{E}(\mu,x,v)\right|\right|_{AA^*}^2\mu(dxdv). \tag{3.23}$$

The functional obtained by replacing A by $Z := \begin{pmatrix} z_1\mathbf{Id}_d \\ z_2\mathbf{Id}_d \end{pmatrix} \in \mathscr{M}_{2d,d}(\mathbb{R})$, we will talk about auxiliary Fisher information. We have

$$\frac{d}{dt}\mathscr{E}[\mu_t^{\mathbf{VFP}}] = \frac{d}{dt}\mathscr{S}[\mu_t^{\mathbf{VFP}}] = -\mathscr{I}[\mu_t^{\mathbf{VFP}}] \le 0. \tag{3.24}$$

### 3.2 Our Assumptions

**Assumption 3.1** ($\mathbb{HMV}$)**.** We put the following hypotheses on the potentials which will ensure properties of existence, uniqueness and contraction:

- ▷ (**H1**)**(Hessian)** The hessian of the confinement potential V is bounded from below and the hessians of the interaction potentials $W^{(k)}$, $k = 2,\dots,N$, are bounded.

▷ (**H2**)(**Lyapunov**) There are two positive constants $c_1$ and $c_2$ such that

$$\forall x \in \mathbb{R}^d, \quad x \cdot \nabla V(x) \geq c_1 |x|^2 - c_2. \tag{3.25}$$

This hypothesis is a Lyapunov condition.

*Remark* 3.3. Since the Hessian $\nabla^2 V$ of V is bounded from below and V satisfies a Lyapunov condition (**H2**), $\alpha \propto e^{-V}$ satisfies a *logarithmic Sobolev inequality* (see *e.g.* [27, 28]).

▷ (**H3**) For all $k \in \{2, \dots, N\}$,

$$\forall \lambda > 0, \quad \int e^{\lambda W^{(k),-}(x) - \sum_{j=1}^k V(x_j)} dx < +\infty. \tag{3.26}$$

*Remark* 3.4. This assumption is trivially satisfied if the $W^{(k)}$ are bounded from below. If (**H2**) holds, it is also always true if $W^{(k),-}(x_1, \dots, x_k) = o(\sum_{j=1}^k |x_j|^2)$ as $|x_1|^2 + \dots + |x_k|^2 \to +\infty$ (since (**H2**) involves that $\liminf_{|x| \to +\infty} V(x)/|x|^2 > 0$). Under the exponential integrability condition, if $\mathbf{H}[\mu|\alpha] < +\infty$, for all $k \in \{2, \dots, N\}$, by Donsker-Varadhan variational formula (in the bounded case) and Fatou's lemma (by approximating $W^{(k),-}$ with $(\min(W^{(k),-}, L))_{L \geq 0}$), we have $W^{(k),-} \in L^1(\mu^{\otimes k})$: then $\mathbf{dom}(\mathbf{H}_W)$ given by Eq. (3.12) satisfies

$$\mathbf{dom}(\mathbf{H}_W) = \Big\{\mu, \quad \mathbf{H}[\mu|\alpha] < +\infty\Big\}.$$

▷ (**H4**)(**Logsob**) The invariant measure $\mu_n$ of the system satisfies a logarithmic Sobolev inequality such that

$$\limsup_{n \to +\infty} \rho_{LS}(\mu_n) > 0. \tag{3.27}$$

▷ (**H5**)(**Contraction**) There exists a distance $\mathbf{d}_{Lip}$ on a subset $\mathscr{Z}$ of $\mathscr{P}(\mathbb{R}^d)$ such that $(\mathscr{P}_2(\mathbb{R}^d), \mathscr{W}_2)$ *continuously injects* into $(\mathscr{Z}, \mathbf{d}_{Lip})$ and $\Phi : \mu \in \mathscr{Z} \longmapsto \Phi(\mu)(dx) := \frac{1}{Z_\mu} e^{-\frac{\delta F}{\delta m}(\mu, x) - V(x)} dx \in \mathscr{Z}$ satisfies

$$\exists k \in (0, 1[, \quad \forall \mu, \nu \in \mathscr{Z}, \quad \mathbf{d}_{Lip}(\Phi(\mu), \Phi(\nu)) \leq k \mathbf{d}_{Lip}(\mu, \nu). \tag{3.28}$$

In others terms, $\Phi$ is $k$-Lipschitz (*contraction*) for $\mathbf{d}_{Lip}$.

*Remark* 3.5. The two above assumptions are not easy to check in practice. In Section 4.3, we thus provide several many-body interaction examples where these conditions apply. Nevertheless, let us give some first comments below

▷ **About** (**H4**): (**H4**) can be certainly satisfied under Bakry-Emery criterion (see Proposition 4.1). There also exist some specific conditions called Zegarlinski conditions (see [2],[29],[30]): we recall Zegarlinski conditions refer to specific conditions on the Hessian of the interaction potential, which are then used (together with other conditions) to prove a uniform logarithmic Sobolev inequality. Finally, let us note that as $\mu_n$ is a Gibbs measure with respect to $\alpha^{\otimes n}$ and its Hamiltonian is $H_{\alpha,n} := n \sum_{k=2}^N U_n(W^{(k)})$, if this Hamiltonian has bounded oscillations ($\mathbf{osc}(H_{\alpha,n}) := \sup H_{\alpha,n} - \inf H_{\alpha,n} < +\infty$) uniformly in $n$, then we can show that by property of tensorization and stability by bounded perturbation, we have (**H4**) seen that according to Royer's book [31, Proposition 3.1.18], $\rho_{LS}(\mu_n) \geq \rho_{LS}(\alpha) e^{-\mathbf{osc}(H_{\alpha,n})}$. In Proposition 4.1, our examples will be yet built with the help of the simpler Bakry-Emery condition. In Proposition 4.4, we provide another class of examples which do not require Bakry-Emery condition.

▷ **About** (**H5**): As concerns (**H5**), we also give some explicit conditions in Proposition 4.1 and Proposition 4.4 with $\mathscr{Z} = \mathscr{P}_2(\mathbb{R}^d)$ and $\mathbf{d}_{Lip} = \mathscr{W}_1$. (**H5**) ensures uniqueness of the fixed point (invariant measure of the McKean-Vlasov process): in *statistical physics*, we say that we have no *phase transition*. This is the crucial point for the proof: $\mathbf{H}_W = \mathbf{H}[\cdot|\Phi(\cdot)]$ (which justifies the name *mean field entropy*). The contractivity assumptions in Eq. (3.28) can follow from Eberle conditions (*lipschitzian spectral gap condition for one particle*): see [2]. To obtain uniqueness, some authors also require *displacement-convexity* (see e.g. [15],[23]): assuming that the functional G in $V^{\sigma,G} : \mu \mapsto \frac{\sigma^2}{2} \mathbf{H}[\mu|\alpha] + G(\mu)$ is displacement-convex. And as the relative entropy is strictly displacement-convex, $V^{\sigma,G}$ is also strictly displacement-convex, which implies the existence of an entropy minimizer ensuring its uniqueness.

**Assumption 3.2** ($\mathbb{VFP}$)**.** In this case, all the conditions stated in $\mathbb{HMV}$3.1 are assumed, together with the two following additional ones

▷ **VFP1.** Lipschitz interactions:

$$\forall k \in \{2, 3, \dots, N\} \quad \exists K > 0, \quad ||\nabla W^{(k)}|| \leq K. \tag{3.29}$$

▷ **VFP2.** Lyapunov condition on confinement:

$$||\nabla^2 V||_{\mathbf{op}} \leq K_1|\nabla V| + K_2. \tag{3.30}$$

*Remark* 3.6. Either of these conditions ensures that the kinetic Fokker-Planck semigroup converges exponentially ( as a family of operators of $\mathcal{H}^1(\mu_Z^n)$ indexed by time ) towards $\mu_Z^n$ and uniformly in the number of particles (see [3] or [24]).

# 4 Main Theorems

## 4.1 First-order case

Under $\mathbb{HMV}$3.1, we establish (see Section 6 for the proof) the following two main results (thus generalizing those of [2]). Let $(\mu_t)_{t\geq 0}$ (given by the arrow (**1**) in Fig. 1) be the flow of solution distributions of the McKean-Vlasov equation associated with the particle system defined by the U−statistic and the confinement potential. Then for any initial condition admitting a moment of order 2, the mean field entropy $\mathbf{H}_W$ decreases exponentially along the flow, i.e.:

**Theorem 4.1** (Exponential decreasing of mean-field entropy)**.** *Assume* $\mathbb{HMV}$*3.1 and let* $\mu_0 \in \mathcal{M}_1^2(\mathbb{R}^d)$ *be an initial condition. Then*

$$\forall t \geq 0, \quad \mathbf{H}_W[\mu_t] \leq \mathbf{H}_W[\mu_0]e^{-\rho_{LS}\frac{t}{2}}. \tag{4.1}$$

From the exponential decrease of the mean field entropy along the flow, we deduce the following exponential convergence in Wassertein metric:

**Theorem 4.2** (Exponential convergence in Wasserstein metric from flow to equilibrium)**.** *Assume* $\mathbb{HMV}$*3.1 give us an initial condition* $\mu_0 \in \mathcal{M}_1^2(\mathbb{R}^d)$*. Then*

$$\forall t \geq 0, \quad \mathcal{W}_2^2(\mu_t, \mu_\infty) \leq \frac{2}{\rho_{LS}}\mathbf{H}_W[\mu_0]e^{-\rho_{LS}\frac{t}{2}}. \tag{4.2}$$

## 4.2 Kinetic case

For kinetic type models, the extension of the above results relies on applications of hypocoercivity arguments (see e.g. [3] or [24] for background). In this setting, we first obtain an exponential decrease in $|||\cdot|||_{H^1\to H^1}$ norm (defined in Section 2).

**Theorem 4.3** (Uniform exponential convergence to equilibrium in the weighted Sobolev space)**.** *Assume* $\mathbb{VFP}$*3.2 and give us an initial condition* $\mu \in \mathcal{M}_1^2(\mathbb{R}^d \times \mathbb{R}^d)$*. Then*

$$\exists \alpha > 0 \quad \exists \beta > 0 \quad \forall n \geq 2, \quad \left|\left|\left|P_t^{Z,(n)} - \mu_Z^n\right|\right|\right|_{H^1\to H^1} \leq \alpha e^{-\beta t}. \tag{4.3}$$

*Remark* 4.1. We still have Theorem 4.3 if we replace the uniform logarithmic Sobolev inequality given in $\mathbb{HMV}$3.1 by a uniform Poincaré inequality. We keep the logarithmic Sobolev inequality to have the following Theorem 4.4. Note that the constants $\alpha > 0$ and $\beta > 0$ can be made explicit uniform. The originality of the proof relies on functional inequalities and hypocoercivity with Lyapunov type conditions, usually not suitable to provide adimensional results.

**Theorem 4.4** (Exponential decay in Wasserstein metric)**.** *Assume* $\mathbb{VFP}$*3.2 and* $\nabla^2 V$ *is bounded. Then there are constants* $C > 0$*,* $\xi > 0$ *and* $\kappa > 0$ *such that* $\forall \mu \in \mathcal{P}_2(\mathbb{R}^d \times \mathbb{R}^d)$*,* $\forall n \geq 2$ *and* $\forall t > 0$*,*

$$\mathbf{H}[\mu_Z^n(t)|\mu_Z^n] \leq C\mathbf{H}[\mu_Z^n(0)|\mu_Z^n]e^{-\xi t}, \tag{4.4}$$

$$\mathcal{W}_2^2(\mu_t^{\mathbf{VFP}}, \mu_\infty^Z) \leq \kappa C\mathcal{S}[\mu]e^{-\xi t}, \tag{4.5}$$

*where* $\mu$ *is the initial condition and* $\mathcal{S}$ *(defined in Eq.* (3.22)*) is the mean-field entropy associated with our second order system given by Eq.* (3.16).

## 4.3 Examples

Let us begin with results which provide some explicit conditions on V and the $W^{(k)}$ under which our results apply. We only focus on $\mathbb{HMV}$3.1 but the extension to $\mathbb{VFP}$3.2 only requires to add the constraints on $\nabla W^{(k)}$ and $\nabla^2 V$ introduced in **VFP1** and **VFP2**. We will use the notation $\underline{A}$ for the lowest eigenvalue of a symmetric matrix.

**Proposition 4.1** (Application in Example 4.1). *Assume that* (**H1**) *holds, that* $W^{(k),-}(x_1,\ldots,x_k) = o(\sum_{j=1}^k V(x_j))$ *as* $|x_1|^2+\ldots+|x_k|^2 \to +\infty$ *and that the following assumption is fulfilled:*

$$\underline{\lambda} > \sum_{k=2}^{N} k(k-1)\|\nabla_{12}^2 W^{(k)}\|_{\mathbf{op},\infty} \tag{4.6}$$

*where* $\|\nabla_{12}^2 W^{(k)}\|_{\mathbf{op},\infty} = \sup_{x\in(\mathbb{R}^d)^k}\|\nabla_{12}^2 W^{(k)}\|_{\mathbf{op}}$ *and*

$$\underline{\lambda} = \inf_{x\in\mathbb{R}^d}\left(\underline{\lambda}_{\nabla^2 V(x)} + \inf_{y\in(\mathbb{R}^d)^{k-1}}\sum_{k=2}^{N} k\underline{\lambda}_{\nabla_{11}^2 W^{(k)}(x,y)}\right).$$

*Then,* $\mathbb{HMV}$ *3.1 is true.*

*Proof.* The proof of this result is achieved in Section 6. □

**Proposition 4.2** (Application in Example 4.2). *Assume that*

*(i)* (**H1**), (**H2**) *and* (**H3**) *hold;*

*(ii)* $\forall k\in\{2,\ldots,N\}$, $x\mapsto \nabla_{x_1}W^{(k)}(x)$ *is bounded;*

*(iii)* $\mathscr{E}_W : \mu \mapsto \sum_{k=2}^N \int W^{(k)} d\mu^{\otimes k}$ *is convex in the flat interpolation sense:*

$$\forall t\in[0,1] \quad \forall(\mu,\nu)\in\mathscr{P}_2(\mathbb{R}^d)^2, \quad \mathscr{E}_W((1-t)\mu+t\nu) \le (1-t)\mathscr{E}_W(\mu)+t\mathscr{E}_W(\nu); \tag{4.7}$$

*(iv)* *Eq.* (4.6) *holds.*

*Then,* $\mathbb{HMV}$ *3.1 is true.*

*Proof.* The proof of this result is achieved in Section 6. □

*Remark* 4.2. For instance, Assumption $(iii)$ is true in Example 4.2.

(⋆) **Regularized Skyrme model**

*Example* 4.1 (Regularized Skyrme model). One of the main areas of research in nuclear physics is the study of nuclei under extreme conditions in spin and isospin. Microscopic methods of mean field type, including the Hartree-Fock method based on the independent particle approximation, are one of the most efficient tools for theoretical predictions in this field. Representing the interactions between nucleons in the nucleus, the effective forces nucleon-nucleon are the main ingredient of these self-consistent microscopic theories. The Skyrme interaction is a zero-range force allowing to construct the mean field in a relatively simple manner: effective phenomenological interaction of zero range which allows the interactions between nucleons in the nucleus to be modeled in a simple manner. Proposed by Skyrme ([8]), this force is limited to the sum of interactions between two and three nucleons. The interaction potential is given by

$$\overrightarrow{F} := \binom{n}{2} U_n(W^{(2)}) + \binom{n}{3} U_n(W^{(3)}) \tag{4.8}$$

with $W^{(2)}$ a potential causing two particles to interact and $W^{(3)}$ a potential causing three particles to interact. In this model, the potentials are functions of Dirac distributions: therefore singular. We will regularize the problem by replacing the Dirac distributions with a smooth approximation: setting $G_\sigma := \frac{1}{(2\pi\sigma^2)^{\frac{d}{2}}} e^{-\frac{|\cdot|^2}{2\sigma^2}}$, $\sigma>0$, we consider the particle system

$$dX_t^{(n)} = \sqrt{2}dB_t^{(n)} - \nabla H_n^{(3)}(X_t^{(n)})dt \tag{4.9}$$

where

$$H_n^{(3)}(x) := \sum_{j=1}^n V(x_j) + \frac{2}{n-1}\sum_{1\le i<j\le n} W^{(2)}(x_i,x_j) + \frac{6}{(n-1)(n-2)}\sum_{1\le i<j<k\le n} W^{(3)}(x_i,x_j,x_k), \tag{4.10}$$

with $W^{(2)}(x,y) = G_\sigma(x-y)$ and $W^{(3)}(x,y,z) = G_\sigma(x-y)G_\sigma(x-z)G_\sigma(y-z)$.

**Proposition 4.3.** *Assume that* V *is a* $\mathscr{C}^2$*-function. Then, if* $\inf_{x\in\mathbb{R}^d}\underline{\lambda}_{\nabla^2 V(x)} > c_1\sigma^{-2-d} + c_2\sigma^{-2-3d}$ *with* $c_1 = 4\times(2\pi)^{-\frac{d}{2}}$ *and* $c_2 = 12(1+4e^{-1})\times(2\pi)^{-\frac{3d}{2}}$*, the regularized Skyrme model satisfies* $\mathbb{HMV}$ *3.1. If furthermore,* ***VFP2*** *holds true for* V*, then,* $\mathbb{VFP}$ *3.2 is also satisfied.*

*Remark* 4.3. As expected, these conditions become more and more stringent when σ goes to 0. Thus, considering the long-time behavior of such models with singular kernels would probably require to develop specific techniques.

*Proof.* Since $W^{(2)}$ and $W^{(3)}$ are bounded, it is enough to check Assumption (4.6) of Proposition 4.1. First, $\nabla G_\sigma(x) = -\sigma^{-2}G_\sigma(x)x$ and $\nabla^2 G_\sigma(x) = \sigma^{-2}(\sigma^{-2}x\otimes x - \mathbb{I}_d)G_\sigma(x)$. Using that for all $x$, $x\otimes x$ is a nonnegative symmetric matrix with $x\otimes x \le |x|^2\mathbb{I}_d$, it follows that

$$-\frac{1}{(2\pi)^{\frac{d}{2}}\sigma^{2+d}}\mathbb{I}_d \le -\frac{1}{\sigma^2}G_\sigma(x)\mathbb{I}_d \le \nabla^2 G_\sigma(x) \le \frac{1}{\sigma^4}G_\sigma(x)\Big(|x|^2-\sigma^2\Big)\mathbb{I}_d. \tag{4.11}$$

We deduce that

$$\underline{\lambda}_{\nabla^2_{11}W^{(2)}(x,y)} = \underline{\lambda}_{\nabla^2 G_\sigma(x-y)} \ge -\frac{1}{(2\pi)^{\frac{d}{2}}\sigma^{2+d}}$$

and that

$$\|\nabla^2_{12}W^{(2)}(x,y)\|_{\mathbf{op},\infty} \le \frac{1}{(2\pi)^{\frac{d}{2}}\sigma^{2+d}}.$$

Using that

$$\nabla^2_{11}(G_\sigma(x-y)G_\sigma(x-z)) = \nabla^2 G_\sigma(x-y)G_\sigma(x-z) + \nabla^2 G_\sigma(x-z)G_\sigma(x-y) + u_{x,z}\otimes u_{x,y} + u_{x,y}\otimes u_{x,z},$$

with $u_{x,y} = \nabla G_\sigma(x-y)$, one also deduces that

$$\underline{\lambda}_{\nabla^2_{11}W^{(3)}(x,y,z)} \ge -\frac{2\|G_\sigma\|_\infty^2}{(2\pi)^{\frac{d}{2}}\sigma^{2+d}} - 2\sigma^{-4}\|G_\sigma\|_\infty|x-y|.|x-z|G_\sigma(x-y)G_\sigma(x-z) \ge -\frac{2(1+2e^{-1})}{(2\pi)^{\frac{3d}{2}}\sigma^{2+3d}},$$

where we used that $\sup_{u\in\mathbb{R}^d}\frac{|u|}{\sigma}G_\sigma(u) \le \sqrt{2e^{-1}}\|G_\sigma\|_\infty$. Finally, one similarly obtains that

$$\|\nabla^2_{12}W^{(3)}(x,y,z)\|_{\mathbf{op},\infty} \le \sigma^{-2}(1+6e^{-1})\|G_\sigma\|^3_\infty = \frac{1+6e^{-1}}{(2\pi)^{\frac{3d}{2}}\sigma^{2+3d}}.$$

Plugging these estimates into Assumption (4.6) yields the result. □

**(⋆⋆) Elementary Symmetric Polynomial Interaction Model (ESPIM):** $\mu\in\mathscr{P}_G(\mathbb{R}^d)\mapsto P(\langle\mu,G\rangle)$ **with** $P\in\mathbb{R}[X]$

*Example* 4.2 (ESPIM model). Let us finish this section with a class of examples with polynomial interaction inspired by elementary symmetric polynomials (ESP). Note that this class will not require Bakry-Emery criterion. The polynomial interaction is built though the $\mathscr{C}^2$-potential $G:\mathbb{R}^d\to\mathbb{R}$: for $N\ge 2$, $k\in\{2,3,\dots,N\}$ and $j\in\{1,\dots,k\}$, let $G^{(j)}:(\mathbb{R}^d)^k\to\mathbb{R}$ be the following symmetric function

$$G^{(j)}(x_1,\dots,x_k) := \sum_{I\subset\mathscr{P}_k(j)}\prod_{i\in I}G(x_i) \quad\text{with}\quad \mathscr{P}_k(j) := \Big\{I\subset\{1,\dots,k\},\quad \mathbf{card}(I)=j\Big\} \tag{4.12}$$

and assume that $W^{(k)}\in\mathbf{Vect}_{\mathbb{R}}\{G^{(1)},\dots,G^{(k)}\}$, i.e.

$$W^{(k)} = \sum_{j=1}^{k}\beta_j^{(k)}G^{(j)} \quad\text{with}\quad (\beta_1^{(k)},\dots,\beta_k^{(k)})\in\mathbb{R}^k. \tag{4.13}$$

Note that

$$G^{(j)}(x_1,\dots,x_k) = P_j(G(x_1),\dots,G(x_k)) \tag{4.14}$$

where $P_j$ denotes the $j^{th}$ ESP defined by

$$P_j(y_1,\dots,y_k) := \sum_{I\subset\mathscr{P}_k(j)}\prod_{i\in I}y_i. \tag{4.15}$$

The homogeneous polynomial $H:\mathscr{P}(\mathbb{R}^d)\to\mathbb{R}$ associated with the McKean-Vlasov equation is defined by

$$H(\mu) = \int V\,d\mu + \sum_{k=2}^{N}\int W^{(k)}\,d\mu^{\otimes k}. \tag{4.16}$$

By symmetry and the Fubini-Tonelli-Lebesgue Theorem,

$$\int W^{(k)} d\mu^{\otimes k} = \sum_{j=1}^{k} \beta_j^{(k)} \int G^{(j)} d\mu^{\otimes k} = \sum_{j=1}^{k} \beta_j^{(k)} \sum_{I\subset \mathscr{P}_k(j)} \int \prod_{i\in I} G(x_i) \prod_{i\in I} \mu(dx_i) = \sum_{j=1}^{k} \beta_j^{(k)} \binom{k}{j} \Big( \int G d\mu \Big)^j, \tag{4.17}$$

and hence,

$$H(\mu) = \int V d\mu + \sum_{k=2}^{N} \sum_{j=1}^{k} \beta_j^{(k)} \binom{k}{j} \Big( \int G d\mu \Big)^j = \int V d\mu + Q\Big( \int G d\mu \Big) \tag{4.18}$$

$$\text{with} \quad Q := \sum_{k=2}^{N} \sum_{j=1}^{k} \beta_j^{(k)} \binom{k}{j} X^j \in \mathbb{R}_N[X].$$

Since $H : \mu \mapsto \int V d\mu + Q \circ T(\mu)$ with $T : \mu \mapsto \int G d\mu$, we have

$$\frac{\delta H}{\delta m}(\mu, x) = V(x) + Q'\Big( \int G d\mu \Big) G(x) \quad \text{with} \quad Q' := \sum_{k=2}^{N} \sum_{j=1}^{k} j \beta_j^{(k)} \binom{k}{j} X^{j-1} \in \mathbb{R}_{N-1}[X]; \tag{4.19}$$

$$\Longrightarrow \nabla \frac{\delta H}{\delta m}(\mu, \cdot) = \nabla V + Q'\Big( \int G d\mu \Big) \nabla G \Longrightarrow \nabla^2 \frac{\delta H}{\delta m}(\mu, \cdot) = \nabla^2 V + Q'\Big( \int G d\mu \Big) \nabla^2 G. \tag{4.20}$$

Furthermore, we have

$$\sup_{\mu\in \mathscr{P}_G(\mathbb{R}^d)} \Big| Q'\Big( \int G d\mu \Big) \Big| \leq \sum_{k=2}^{N} \sum_{j=1}^{k} j |\beta_j^{(k)}| \binom{k}{j} ||G||_\infty^{j-1} =: \gamma_1; \tag{4.21}$$

$$\sup_{\mu\in \mathscr{P}_G(\mathbb{R}^d)} \Big| Q''\Big( \int G d\mu \Big) \Big| \leq \sum_{k=2}^{N} \sum_{j=1}^{k} j(j-1) |\beta_j^{(k)}| \binom{k}{j} ||G||_\infty^{j-2} =: \gamma_2. \tag{4.22}$$

**Proposition 4.4.** *Assume that* $G$, $\nabla G$ *and* $\nabla^2 G$ *are bounded. As* $||G||_\infty < +\infty$*, we have for all* $i = 1, 2$, $\gamma_i < +\infty$ *and*

$$\mathscr{P}_2(\mathbb{R}^d) \subset \mathscr{P}_G(\mathbb{R}^d) := \{\mu \in \mathscr{P}(\mathbb{R}^d), \quad G \in L^1(\mu)\} = \mathscr{P}(\mathbb{R}^d).$$

*Let us further assume that*

- ▹ $Q'' \geq 0$ *and the confinement potential* $V$ *satisfies* (**H1**) *and* (**H2**)*;*
- ▹ $$\underline{\lambda}^\star := \inf_{x\in\mathbb{R}^d} \Big( \underline{\lambda}_{\nabla^2 V(x)} + \gamma_3 \underline{\lambda}_{\nabla^2 G(x)} \Big) > 0 \quad \text{and} \quad \frac{\gamma_2}{\underline{\lambda}^\star} ||\nabla G||_\infty^2 < 1,$$
  *with* $\gamma_3 := \inf_{\mu\in\mathscr{P}_2(\mathbb{R}^d)} Q'(\int G d\mu) \quad (|\gamma_3| \leq \gamma_1)$*.*

*Then,* **ESPIM** *satisfies* $\mathbb{HMV}$*3.1. If furthermore,* ***VFP2*** *holds true for* $V$*, then,* $\mathbb{VFP}$*3.2 is also satisfied.*

*Proof.* The proof of this result is achieved in Section 6. □

*Remark* 4.4. If $\nabla^2 \frac{\delta H}{\delta m}(\mu, \cdot) \geq \rho \mathbf{Id}$ with $\rho > 0$, by the Bakry-Emery curvature criterion, $\Phi(\mu)(dx) = \frac{1}{Z_\mu} e^{-\frac{\delta H}{\delta m}(\mu, x)} dx$ satisfies a uniform logarithmic Sobolev inequality in the measure.

# 5 Sketch of proofs and preliminaries

## 5.1 Sketch of proofs

**First order case.** The diagram given in Fig. 1 summarizes the strategy of proof: we show (**4**) from (**1**) , (**2**) and (**3**). And in this diagram, the quantities involved are:

- ▹ $\mu^{\otimes n} P_t^{(n)} = \mu_n(t)$ the law at time $t$ of the particle system induced by the confinement potential and the U−statistics;
- ▹ $\mu_n^{(i)}(t)$ the $i$−th marginal of $\mu_n(t)$;
- ▹ $\mu_\infty^{(n)} = \mu_n$ the invariant measure of the particle system;

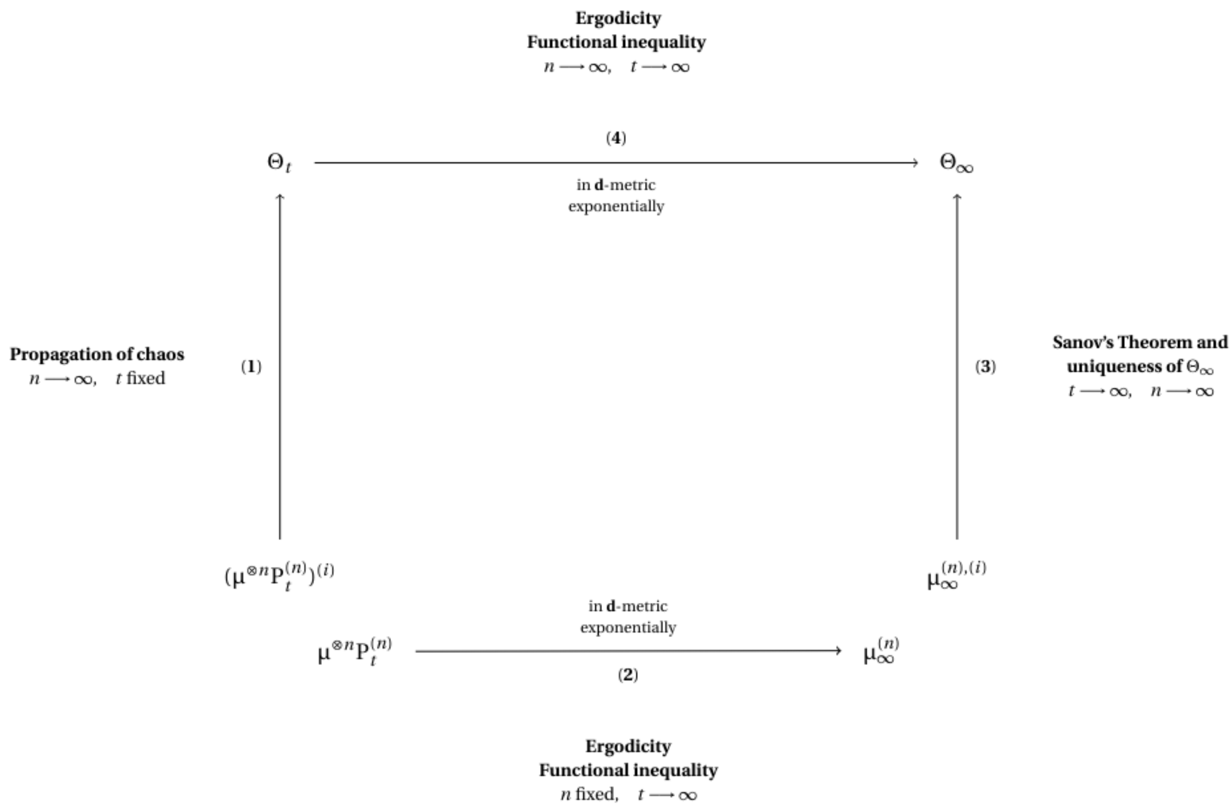


Figure 1: Diagram of convergences

- ▷ $\mu_n^{(i)}$ the $i$−th marginal of $\mu_n$;
- ▷ $\Theta_t = \mu_t$ the law at time $t$ of the McKean-Vlasov process obtained by propagation of chaos;
- ▷ $\Theta_\infty = \mu_\infty$ the invariant measure of the McKean-Vlasov process;
- ▷ $\mathbf{d} = \mathscr{W}_2$.

**Arrow** (**1**)**.** The McKean-Vlasov process classically appears as the mean-field limit of a particle system. This property is recalled and studied, among others, in [7]. In our case, see Theorem 5.1.

**Arrow** (**2**)**.** The process $X^n$ is a homogeneous diffusion process of the Langevin-Kolmogorov type which is a class of Markov processes. In the literature, the long-time behavior for this class is classically studied (see e.g. [17, [22]]). In order to ensure this property (see Section 5.2.Theorem 5.3), exponentially in time and uniformly in number of particle $n$, we rely on (**H4**) in $\mathbb{HMV}$3.1 and the equivalence between Sobolev's inequality, exponential decay of entropy and Talagrand's second inequality for Gibbs measures.

**Arrow** (**3**)**.** This arrow is ensured by (**H1**), (**H2**) and (**H3**) in $\mathbb{HMV}$3.1 which allow us to obtain large deviations principle and Sanov-type theorem (see Section 5.2.Theorem 5.2.Proposition 5.8).

**Arrow** (**4**)**.** To establish this last arrow, we will use the fact that the nonlinear Sobolev inequality ($\rho_{LS}\mathbf{H}_W \leq 2\mathbf{I}_W$) given in Section 5.2.Theorem 5.3 is also equivalent to the exponential decrease of the mean field entropy $\mathbf{H}_W$ along the flow $(\mu_t)_{t\geq 0}$ of the McKean-Vlasov distributions and to the second nonlinear Talagrand inequality ($\rho_{LS}\mathscr{W}_2^2(\cdot,\mu_\infty) \leq 2\mathbf{H}_W$). Note that Talagrand inequalities allow to recover usual Wasserstein convergence (and then convergence in law) from entropic convergence. Note that concentration inequalities could also stem from Talagrand inequalities, although the stronger Logarithmic Sobolev inequality is more often used in this context.

*Remark* 5.1. In the case of the two-body interaction (N = 2) associated with the equation of granular media given by Eq. (1.16), the exponential convergence in entropy (given in Theorem 4.1) should be equivalent to the mean field log-Sobolev inequality $\rho_{LS}\mathbf{H}_W \leq 2\mathbf{I}_W$ (in Theorem 5.3), basing on (gradient flow and Gronwall lemma)

$$-\frac{d}{dt}\mathbf{H}_W[\mu_t] = 4\mathbf{I}_W[\mu_t] \Longrightarrow \frac{d}{dt}\mathbf{H}_W[\mu_t] \leq -2\rho_{LS}\mathbf{H}_W[\mu_t] \Longrightarrow \mathbf{H}_W[\mu_t] \leq \mathbf{H}_W[\mu_0]e^{-2\rho_{LS}t} \tag{5.1}$$

noted by Carrillo-McCann-Villani in their convex framework. The proof of $-\frac{d}{dt}\mathbf{H}_W[\mu_t] = 4\mathbf{I}_W[\mu_t]$ demands the *regularity of* $t \longmapsto \mu_t$ (Fig. 1) which requires the *PDE theory of the McKean-Vlasov equation*. That is why we prefer to give a *rigorously probabilistic proof based directly on the log-Sobolev inequality of* $\mu_n$ (Fig. 1) in $\mathbb{HMV}$3.1.(**H4**). As for Theorem 4.2 on exponential decay in Wasserstein metric, it follows from the previous one (Theorem 4.1) via Talagrand's T2-inequality.

**Second order case.** The proof in this case, can also be described by the diagram given in Fig. 1 but with the following notations:

- ▷ $\mu^{\otimes n}P_t^{(n)} = \mu_Z^n(t)$ the law at time $t$ of the kinetic particle system induced by the confinement potential and the U−statistics;
- ▷ $(\mu^{\otimes n}P^{(n)})^{(i)} =: \mu_Z^{n,(i)}(t)$ the $i$−th marginal of $\mu_Z^n(t)$;
- ▷ $\mu_\infty^{(n)} = \mu_Z^n$ the invariant measure of the particle system;
- ▷ $\mu_Z^{n,(i)}$ the $i$−th marginal of $\mu_Z^n$;
- ▷ $\Theta_t = \mu_t^{\mathbf{VFP}}$ the law at time $t$ of the Vlasov-Fokker-Planck process obtained by propagation of chaos;
- ▷ $\Theta_\infty = \mu_\infty^Z$ the invariant measure of the Vlasov-Fokker-Planck process;
- ▷ $\mathbf{d} = |||\cdot - \cdot|||_{H^1 \to H^1}$ or $\mathbf{d} = \mathscr{W}_2$.

**Arrow (1).** We first recall the generator $\mathscr{L}_{Z,n}$ defined (in Hormander form) by Eq. (5.88) is a non-symmetric hypoelliptic operator (see Remark 5.4). The related $n$-particle system given by Eq. (3.16) converges to the Vlasov-Fokker-Planck equation (mean-field limit of Eq. (3.16)) when $n \to +\infty$ (see Theorem 5.1).

**Arrow (2).** The process $Z^n$ is a homogeneous diffusion process of the Langevin type usually called kinetic Fokker-Planck process. The study of the long-time behavior of the particle system requires the help of hypocoercivity tools (see e.g. [3] and [24]). We recall that

$$\forall x \in \mathbb{R}^{nd}, \quad S_{1,n}(x) := \sum_{i=1}^{n} V(x_i) + n\sum_{k=2}^{N} U_n(W^{(k)}). \tag{5.2}$$

In particular, 3.2 ensures the following Poincaré and log-Sobolev inequalities

- ▷ **UPI.** We say that $\mu_{1,n}$ satisfies a uniform Poincaré inequality if

$$\exists\lambda > 0 \quad \forall n \geq 2 \quad \forall\varphi \in \mathscr{C}_c^\infty(\mathbb{R}^{nd}), \quad \lambda\mathbb{V}_{\mu_{1,n}}[\varphi] \leq \mathbb{E}_{\mu_{1,n}}[||\nabla_x\varphi||^2]. \tag{5.3}$$

- ▷ **ULSI.** We say that $\mu_{1,n}$ satisfies a uniform logarithmic Sobolev inequality if

$$\exists\rho > 0 \quad \forall n \geq 2 \quad \forall\varphi \in \mathscr{C}_c^\infty(\mathbb{R}^{nd}), \quad \rho\mathbf{Ent}_{\mu_{1,n}}[\varphi^2] \leq \mathbb{E}_{\mu_{1,n}}[||\nabla_x\varphi||^2]. \tag{5.4}$$

Under **(UPI)**, we are able to obtain as an application of Villani's theorem the following exponential rate to equilibrium

$$\forall n \geq 2, \quad \left|\left|\left|P_t^{Z,(n)} - \mu_Z^n\right|\right|\right|_{H^1 \to H^1} \leq \alpha e^{-\beta t} \tag{5.5}$$

with constants $\alpha > 0$ and $\beta > 0$ make explicit uniform. The idea in Villani's proof of [3, Theorem.3] is as follows: if one could find a Hilbert space such that the operator $\mathscr{L}_{Z,n}$ is coercive with respect to its norm, then one has exponential convergence for the semigroup $(P_t^{Z,(n)})_{t\geq 0}$ under such a norm. If, in addition, this norm is equivalent to some usual norm (such as $\mathscr{H}^1(\mu_Z^n)$−norm), then one obtains exponential convergence under the usual norm as well. In his statement of [24, Theorem.35], the *boundedness condition* is verified by $||\nabla^2 S_{1,n}||_{\mathbf{op}} \leq C(1 + ||\nabla S_{1,n}||)$ with a constant M depending unfortunately on the dimension. The $L^2$ and $H^1$ norms are not suitable to obtain a result on the non-linear system (such as Eq. (4.4) and Eq. (4.5)). On the other hand, thanks to **(ULSI)** playing a fundamental role in the exponential return in Wasserstein metric (see e.g. [5, Theorem.7] or [6, Theorem.10]), we are able to prove Eq. (4.4) which in turn will allow us to deduce Eq. (4.5).

**Arrow (3).** The results of large deviations on the U−statistics in the non-kinetic case in Section 5.2 and the fact that $\mu_Z^n = \mu_{1,n} \otimes \mu_{2,n}$ allow to deduce that the random empire measurements of the kinetic particle system satisfy the principle of large deviations under $\mu_Z^n$ of good rate function defined by

$$\forall (\mu_x, \mu_v) \in \mathscr{P}_x(\mathbb{R}^d) \times \mathscr{P}_v(\mathbb{R}^d), \quad \mathbf{I}(\mu_x, \mu_v) := \mathbf{H}_W[\mu_x] + \mathbf{H}[\mu_v | \mathscr{N}(0, \mathbf{Id}_d)]. \tag{5.6}$$

Thus there exists by inf-compactness a Maxwellian to the nonlinear Vlasov-Fokker-Planck equation and this equilibrium (invariant measure of the nonlinear Vlasov-Fokker-Planck process) is unique. See Section 5.2.Theorem 5.2.Proposition 5.8.Proposition 5.9.Appendix A.4.

**Arrow (4).** This part is obtained by the first-order case by exploiting the uniform logarithmic Sobolev inequality and the Hormander form given by Eq. (5.88) (see Section 5.2).

*Remark* 5.2. By applying hypocoercivity tools to the system with $n$ particles given by Eq. (3.16), we obtain a (uniform in $n$) convergence rate to equilibrium which in turn extends to the limiting non linear system.

### 5.2 Preliminaries

**Propagation of chaos for polynomial interacting particle systems.** Below, we recall or extend some conditions on the interaction potentials which guarantee the propagation of chaos for some particle systems with polynomial interaction. Even if our proof only requires such properties in finite horizon, we also provide some properties which lead to propagation of chaos uniform in time.

To this end, we use the classical (synchrounous) coupling strategy: let $(X^{(n,p)})_{p=1}^n$ denote the particle system and $(X^{(p)})_{p=1}^n$ denote $n$ copies of the limiting Mc-Kean-Vlasov process built with the same Brownian motions than in the particle system. Assume that all the paths have the same initial condition $X_0 \sim \mu_0$. Then, the following proposition holds:

**Theorem 5.1** (Chaos propagation in Wasserstein $\mathscr{W}_2$ metric)**.** *Assume that* $V$ *and the* $W^{(k)}$ *are* $\mathscr{C}^2$ *and that* $\exists \beta \in \mathbb{R}$ $\forall k \in \{2, \dots, N\}$ $\exists \beta_k \in \mathbb{R}$ $\forall (x, y) \in \mathbb{R}^d \times \mathbb{R}^d$

$$\langle \nabla V(x) - \nabla V(y), x - y \rangle \geq -\beta ||x - y||^2; \tag{5.7}$$

$$\langle \nabla_{x_1} W^{(k)}(x, \cdot) - \nabla_{x_1} W^{(k)}(y, \cdot), x - y \rangle \geq -\beta_k ||x - y||^2. \tag{5.8}$$

*Then, for every* $T > 0$*, a constant* $K_T$ *exists such that for every* $\mu_0 \in \mathscr{P}_2(\mathbb{R}^d)$*,*

$$\sup_{0 \leq t \leq T} \mathbb{E}|X_t^{(n),1} - X_t^{(1)}|^2 \leq \frac{K_T}{n}. \tag{5.9}$$

*Furthermore, if* $\omega := \beta + \sum_{k=2}^N k\beta_k < 0$*, the upper bound is uniform in time, i.e.,*

$$\sup_{t \geq 0} \mathbb{E}|X_t^{(n),1} - X_t^{(1)}|^2 \leq \frac{K_\infty}{n}, K_\infty < +\infty. \tag{5.10}$$

*The constants are specified in the proof.*

*Proof.* See Appendix A.1 □

*Remark* 5.3. ▷ For the sake of simplicity, we only provided the result for the classical McKean-Vlasov process. The extension of the result in finite horizon easily extends to the kinetic setting as soon as (**H1**).

▷ As mentioned before, the statement also leads to uniform in time propagation of chaos but it certainly requires that the function V plays a confinement role which is characterized by the fact that $\omega$ is assumed to be negative.

▷ If for all $k \in \{2, \dots, N\}$, $\nabla_{x_1} W^{(k)}$ is just uniformly bounded and uniformly Lipschitzian in the first coordinate, then it does not necessarily satisfy the conditions of the Theorem 5.1 but if the confinement potentiel V satisfies them, we still have the conclusions at least in a short time: for all $(x,y)\in\mathbb{R}^d\times\mathbb{R}^d$ and $\mu\in\mathscr{P}(\mathbb{R}^d)$,

$$x\cdot\nabla V(x)\geq(-\beta-|\nabla V(0)|)|x|^2-|\nabla V(0)|; \tag{5.11}$$

$$\begin{aligned}-\langle\nabla_{x_1}W^{(k)}(x,\cdot)-\nabla_{x_1}W^{(k)}(y,\cdot),x-y\rangle &\leq|\nabla_{x_1}W^{(k)}(x,\cdot)-\nabla_{x_1}W^{(k)}(y,\cdot)|\cdot|x-y| \\ &\leq[x_1\mapsto\nabla_{x_1}W^{(k)}(x_1,\cdot)]_1|x-y|^2;\end{aligned} \tag{5.12}$$

$$\begin{aligned}x\cdot\nabla\frac{\delta H}{\delta m}(\mu,x)&=x\cdot\nabla V(x)+\sum_{k=2}^{N}kx\cdot\nabla_{x_1}W^{(k)}*\mu^{\otimes k-1}(x) \\ &\geq(-\beta-|\nabla V(0)|)|x|^2-|\nabla V(0)|-|x|\sum_{k=2}^{N}k[x_1\mapsto W^{(k)}(x_1,\cdot)]_1 \\ &\geq\Big(-\beta-|\nabla V(0)|-|\nabla V(0)|-\sum_{k=2}^{N}k[x_1\mapsto W^{(k)}(x_1,\cdot)]_1\Big)|x|^2 \\ &\quad-|\nabla V(0)|-\sum_{k=2}^{N}k[x_1\mapsto W^{(k)}(x_1,\cdot)]_1.\end{aligned} \tag{5.13}$$

And this last inequality comes from a disjunction of cases depending on whether the vector is on the unit ball or not.

**Back to U-statistics.** The results on the U−statistics (5.1,5.2,5.3,5.4) and the inf-compactness of the entropy functional $\mathbf{H}_W$ (5.5,5.2,5.8) are inspired by [1] in the case $S=\mathbb{R}^d$. We recall that the expectation of $W^{(k)}$ under $\mu^{\bigotimes k}$ exists if and only if

$$\mathbb{E}_{\mu^{\otimes k}}[W^{(k),+}]<+\infty\quad\text{or}\quad\mathbb{E}_{\mu^{\otimes k}}[W^{(k),-}]<+\infty. \tag{5.14}$$

First we present the law of large numbers of the U−statistic (see [[32], Corollary 3.1.1] or [[1], Lemma 3.1]). We recall that U-statistics are defined in Eq. (1.19).

**Proposition 5.1** (law of large numbers for U−statistics)**.** *Let $(X_n)_{n\geq1}$ be a sequence of independent and identically distributed random variables with values in a measurable space $(E,\mathscr{B}(E))$ equipped with its Borelian tribe and $\Phi:E^k\longrightarrow\mathbb{R}$ a symmetric measurable function such that*

$$\mathbb{E}[|\Phi(X_1,\ldots,X_k)|]<+\infty,\quad\text{then}\quad U_n(\Phi)\overset{n\to+\infty}{\longrightarrow}\mathbb{E}[\Phi(X_1,\ldots,X_k)]\quad\text{with probability}\quad 1. \tag{5.15}$$

*Proof.* See Appendix A.6 or [1]. □

In terms of integrals, this result means that for any function $\Phi\in\mathscr{M}_{sym}(E^k,\mathbb{R})$ with $E^k$ provided with the tensor tribe (or product) and any measure $\mu\in\mathscr{P}(E)$ such that $\Phi\in L^1(\mu^{\bigotimes k})$, we almost surely have

$$U_n(\Phi)\overset{n\to+\infty}{\longrightarrow}\mathbb{E}_{\mu^{\otimes k}}[\Phi]:=\int_{E^k}\Phi(x)\mu^{\bigotimes k}(dx). \tag{5.16}$$

This result can also be seen as a law of large numbers for U−statistics. From this result, we deduce that $\forall k\in\{2,\ldots,N\}$, if $W^{(k)}\in L^1(\mu^{\bigotimes k})$, then we almost surely have $U_n(W^{(k)})$ tends to $\mathbf{W}^{(k)}[\mu]$. We first recall the decoupling inequality of Victor H. De La Peña (see [[33], 1992]).

**Proposition 5.2** (Decoupling and Khintchine inequalities for U−statistics)**.** *Let $(X_n)_{n\geq1}$ be a sequence of random variables with values in a measurable space $(E,\mathscr{B}(E))$, independent and identically distributed. We assume that*

$$(X_1^j,\ldots,X_n^j)_{j=1,\ldots,k} \tag{5.17}$$

*are k independent copies of $(X_1,\ldots,X_n)$. Then for all increasing convex functions $\Psi:[0,+\infty)\longrightarrow\mathbb{R}$ and measurable symmetric $\Phi:E^k\longrightarrow\mathbb{R}$ such that $\mathbb{E}[|\Phi(X_1,\ldots,X_k)|]<+\infty$, we have*

$$\mathbb{E}\Big[\Psi\Big(\Big|\sum_{(i_1,\ldots,i_k)\in I_n^k}\Phi(X_{i_1},\ldots,X_{i_k})\Big|\Big)\Big]\leq\mathbb{E}\Big[\Psi\Big(C_k\Big|\sum_{(i_1,\ldots,i_k)\in I_n^k}\Phi(X_{i_1}^1,\ldots,X_{i_k}^k)\Big|\Big)\Big] \tag{5.18}$$

*with*

$$C_2:=8\quad\text{and}\quad\forall k\geq3,\quad C_k:=2^k\prod_{j=2}^{k}(j^j-1). \tag{5.19}$$

**Proposition 5.3.** *Let* $1 \le k \le n$, $(X_i^j)_{1\le i\le n, 1\le j\le k}$ *be independent random variables with values in* $(E, \mathscr{B}(E))$. *For all* $(i_1, \dots, i_k) \in I_n^k$, *defining* $\Phi_{i_1,\dots,i_k} : E^k \longrightarrow \mathbb{R}$ *a measurable function of k variables, we have*

$$\log \mathbb{E}\Big[\exp\Big(\frac{1}{|I_n^k|}\sum_{i\in I_n^k}\Phi_i(X_{i_1}^1,\dots,X_{i_k}^k)\Big)\Big] \le \frac{n-k+1}{|I_n^k|}\sum_{i\in I_n^k}\log\mathbb{E}\Big[\exp\Big(\frac{1}{n-k+1}\Phi_i(X_{i_1}^1,\dots,X_{i_k}^k)\Big)\Big]. \tag{5.20}$$

*Proof.* See Appendix A.6 or [1]. □

**Proposition 5.4** (Decoupling corollary)**.** *For* $(X_i)_{i\ge1}$ *a sequence of independent and identically distributed random variables according to* $\alpha$, *we denote* $\Lambda_n(\cdot, W^{(k)})$ *the log-Laplace transformation associated with the* U*−statistic of order k, i.e. to within a factor, the logarithm of the moment generating function, namely*

$$\forall n \ge k \ge 2, \quad \forall \lambda > 0, \quad \Lambda_n(\lambda, W^{(k)}) := \frac{1}{n}\log\mathbb{E}\Big[e^{n\lambda U_n(W^{(k)})}\Big]. \tag{5.21}$$

*If* $W^{(k)} \in L^1(\alpha^{\otimes k})$, *then*

$$\Lambda_n(\lambda, W^{(k)}) \le \frac{1}{k}\log\mathbb{E}\Big[\exp\Big(kC_k\lambda|W^{(k)}(X_1,\dots,X_k)|\Big)\Big]. \tag{5.22}$$

*Proof.* See Appendix A.6 or [1]. □

**Large deviations: inf-compactness of mean-field entropy and existence of an equilibrium point.** We will use a *large deviations result* ensuring the *infcompactness of the entropy functional* to show the *existence of an invariant measure* for the nonlinear process studied.

**Proposition 5.5** (Lower bound of large deviations for $L_n$ under $\mu_n^*$)**.** *Under the integrability assumptions on the interaction potentials* $(W^{(k)})_{2\le k\le N}$, *we have the lower bound of large deviations for* $\{\mu_n^*(L_n \in \cdot)\}_{n\ge N}$, *i.e.*

$$\begin{aligned}\forall \mathscr{O} \subset \mathscr{M}_1(\mathbb{R}^d) \quad \text{open}, \quad l^*(\mathscr{O}) &:= \liminf_{n\to+\infty}\frac{1}{n}\log(\mu_n^*(L_n \in \mathscr{O})) \\ &\ge -\inf\Big\{\mathbf{E}_W[\mu]\Big| \quad \mu \in \mathscr{O}, \quad \forall 2 \le k \le N, \quad W^{(k)} \in L^1(\mu^{\otimes k})\Big\}.\end{aligned} \tag{5.23}$$

*In particular, we have*

$$\liminf_{n\to+\infty}\Big\{\frac{1}{n}\log(Z_n) - \log(C)\Big\} \ge -\inf\Big\{\mathbf{E}_W[\mu]\Big| \quad \mu \in \mathscr{M}_1(\mathbb{R}^d), \quad \forall 2 \le k \le N, \quad W^{(k)} \in L^1(\mu^{\otimes k})\Big\}. \tag{5.24}$$

*Proof.* See Appendix A.6 or [1]. □

**Proposition 5.6** (Exponential approximation of the U−statistic)**.** *Assuming that for all* $\lambda > 0$,

$$\mathbb{E}[\exp(\lambda|W^{(k)}|(X_1,\dots,X_k))] < +\infty, \tag{5.25}$$

*then there exists a sequence* $(W_m^{(k)})_{m\ge1}$ *of bounded continuous functions such that*

$$\forall \delta > 0, \quad \lim_{m\to+\infty}\limsup_{n\to+\infty}\frac{1}{n}\log\mathbb{P}(|U_n(W_m^{(k)}) - U_n(W^{(k)})| > \delta) = -\infty. \tag{5.26}$$

*Proof.* See Appendix A.6 or [1]. □

**Theorem 5.2** (Large deviations principle for U−statistics)**.** *Let* $(X_i)_{i\ge1}$ *be a sequence of independent and identically distributed random variables with distribution* $\alpha$. *We assume that we have exponential integrability of the interaction potentials under the tensor products of* $\alpha$ *by itself, i.e.*

$$\forall k \in \{2,\dots,N\}, \quad \forall \lambda > 0, \quad \Big(\mathbb{E}\Big[e^{\lambda|W^{(k)}(X_1,\dots,X_k)|}\Big] < +\infty \Longleftrightarrow e^{\lambda|W^{(k)}|} \in L^1(\alpha^{\otimes k})\Big). \tag{5.27}$$

*Then*

$$\Big\{\mathbb{P}\Big((L_n, U_n(W^{(2)}), \dots, U_n(W^{(N)})) \in \cdot\Big)\Big\}_{n\ge N} \tag{5.28}$$

*satisfies a large deviations principle on the product space* $\mathscr{M}_1(\mathbb{R}^d) \times \mathbb{R}^{N-1}$ *and good rate function given by*

$$\mathbf{I}_U(\mu, x_2, \dots, x_N) := \begin{cases}\mathbf{H}[\mu|\alpha], & \text{if} \quad \forall k, \quad x_k = \mathbf{W}^{(k)}[\mu], \\ +\infty & \text{otherwise.}\end{cases} \tag{5.29}$$

*Proof.* Let $(W_m^{(k)})_{m\geq 1}$ be the sequence of bounded continuous functions of the proof of Proposition 5.6 (see Appendix A.6) such that for all $\lambda > 0$,

$$\varepsilon(\lambda, m, k) := \log \int_{(\mathbb{R}^d)^k} e^{\lambda |W^{(k)} - W_m^{(k)}|} d\alpha^{\otimes k} \overset{m\to+\infty}{\longrightarrow} 0. \tag{5.30}$$

For all $m \geq 1$, we set

$$f_m(\mu) := \Big(\mu, \mathbf{W}_m^{(2)}[\mu], \dots, \mathbf{W}_m^{(N)}[\mu]\Big), \quad f(\mu) := \Big(\mu, \mathbf{W}^{(2)}[\mu], \dots, \mathbf{W}^{(N)}[\mu]\Big). \tag{5.31}$$

We consider the following metric on the product space

$$\mathbf{d}\Big((\mu, x_2, \dots, x_N), (\nu, y_2, \dots, y_N)\Big) := d_{LP}(\mu, \nu) + \sum_{k=2}^{N} |x_k - y_k| = d_{LP}(\mu, \nu) + |x - y|_1, \tag{5.32}$$

and note that

$$\mathbf{d}(f_m(\mu), f(\mu)) = \sum_{k=2}^{N} \Big| \int_{(\mathbb{R}^d)^k} (W_m^{(k)} - W^{(k)}) d\mu^{\otimes k} \Big|. \tag{5.33}$$

The sequel of the proof is divided in three steps.

**Step 1: Continuity of $f_m$.** For this step, it suffices to show that for all $k \in \{2, \dots, N\}$, $\mu \in \mathscr{M}_1(\mathbb{R}^d) \longmapsto \mathbf{W}_m^{(k)}[\mu]$ is continuous for the convergence topology weak. Let $\mu_n \overset{n\to+\infty}{\longrightarrow} \mu$ in $(\mathscr{M}_1(\mathbb{R}^d), d_{LP})$ . By the Skorokhod representation theorem, there exists a sequence $(Y_n)_n$ of random variables with values in $\mathbb{R}^d$ such that $Y_n \sim \mu_n$ and almost surely, $Y_n \overset{n\to+\infty}{\longrightarrow} Y \sim \mu$. Let $(Y_n^{(i)}, n \geq 0, Y^{(i)})_{1\leq i\leq k}$ be independent copies of $(Y_n, n \geq 0, Y)$. We have for all $i$, almost surely, $Y_n^{(i)} \overset{n\to+\infty}{\longrightarrow} Y^{(i)}$, which implies that almost surely, $(Y_n^{(1)}, \dots, Y_n^{(k)}) \overset{n\to+\infty}{\longrightarrow} (Y^{(1)}, \dots, Y^{(k)})$. In particular, $\mu_n^{\otimes k}$ tends weakly to $\mu^{\otimes k}$, which proves the continuity of the above functional.

**Step 2: Good exponential approximation of $(L_n, U_n(W^{(2)}), \dots, U_n(W^{(N)}))$ by $f_m(L_n)$.** By exponential approximation of the U−statistic, we have for all $\delta > 0$,

$$\lim_{m\to+\infty} \limsup_{n\to+\infty} \frac{1}{n} \log \mathbb{P}\Big(\mathbf{d}\Big((L_n, U_n(W^{(2)}), \dots, U_n(W^{(N)})), (L_n, U_n(W_m^{(2)}), \dots, U_n(W_m^{(N)}))\Big) > \delta\Big) = -\infty,$$

i.e. $(L_n, U_n(W_m^{(2)}), \dots, U_n(W_m^{(N)}))$ is a good exponential approximation of $(L_n, U_n(W^{(2)}), \dots, U_n(W^{(N)}))$ .
Moreover, $(L_n, U_n(W^{(2)}), \dots, U_n(W^{(N)}))$ and $f_m(L_n)$ are exponentially equivalent because we have the following uniform estimate

$$\begin{aligned}\Big|U_n(W_m^{(k)}) - \int W_m^{(k)} dL_n^{\otimes k}\Big| &\leq \Big(1 - \frac{|I_n^k|}{n^k}\Big)\Big(|U_n(W_m^{(k)})| + ||W_m^{(k)}||_\infty\Big) \\ &\leq 2\Big(1 - \frac{|I_n^k|}{n^k}\Big)||W_m^{(k)}||_\infty \overset{n\to+\infty}{\longrightarrow} 0.\end{aligned} \tag{5.34}$$

We get that when $m \to +\infty$, $f_m(L_n)$ is a good exponential approximation of $(L_n, U_n(W^{(2)}), \dots, U_n(W^{(N)}))$.

**Step 3: LDP.** By Sanov theorem and the LDP approximation theorems, to get the desired LDP, it suffices to show that for all $L > 0$,

$$\sup_{\mu,\ \mathbf{H}[\mu|\alpha]\leq L} \mathbf{d}(f_m(\mu), f(\mu)) \overset{m\to+\infty}{\longrightarrow} 0. \tag{5.35}$$

Indeed, for all $\lambda > 0$, $L > 0$ and $\mu$ such that $\mathbf{H}[\mu|\alpha] \leq L$, by the variational formula of Donsker-Varadhan and Fatou's lemma, we have for all $k \in \{2 \dots, N\}$,

$$\int |W_m^{(k)} - W^{(k)}| d\mu^{\otimes k} \leq \frac{1}{\lambda}\Big(\mathbf{H}[\mu^{\otimes k}|\alpha^{\otimes k}] + \log \int e^{\lambda |W_m^{(k)} - W^{(k)}|} d\alpha^{\otimes k}\Big) \leq \frac{1}{\lambda}\Big(kL + \varepsilon(\lambda, m, k)\Big). \tag{5.36}$$

This completes the proof of the theorem because $\lambda$ is arbitrary and for all $\lambda > 0$, $\varepsilon(\lambda, m, k) \overset{m\to+\infty}{\longrightarrow} 0$. □

We are now able to prove the inf-compactness of the mean-field entropy functional.

**Proposition 5.7** (Inf-compactness of the mean-field entropy functional)**.** *The mean-field entropy functional is inf-compact.*

*Proof of Proposition 5.7.* We will do the proof in three steps. We recall that if we have a good rate function, then its infimum on any closed nonempty is reached, that is to say that this infimum is a minimum.

First, Assumption (A1) of [1] is clearly satisfied: taking $\mu = h\alpha$ with $h$ continuous with compact support, we deduce from the local boundedness of $W^{(k)}$ that

$$\mathbf{H}[\mu|\alpha] + \mathbb{E}_{\mu^{\otimes k}}[W^{(k),+}] < +\infty.$$

By (**H3**), we have the strong exponential integrability condition of (A1).

Moreover, under (**H2**), we necessarily have that $\lim_{|x|\to+\infty} \frac{V(x)}{|x|^2} > 0$. Actually, by (**H2**), there exists a positive M such that for every $x$ with $|x| > M$, $x.\nabla V(x) \geq \frac{c_1}{2}|x|^2$. Then, for any $x$ with $|x| \geq M$, set $\rho = \frac{M}{|x|}$. We have

$$\begin{aligned} V(x) - V(\rho x) &= (1-\rho)\int_0^1 \nabla V((1+u(1-\rho))x).x\,du = \int_0^1 \nabla V((1+u(1-\rho))x).(1+u(1-\rho))x \times \frac{1-\rho}{1+u(1-\rho)}du \\ &\geq \frac{c_1|x|^2}{2}\int_0^1 (1+u(1-\rho))(1-\rho)\,du \geq \frac{(1-\rho)c_1|x|^2}{2} \geq \frac{c_1|x|^2}{4} \end{aligned}$$

as soon as $|x| \geq 2M$ (since in this case, $\rho = \frac{M}{|x|} \leq \frac{1}{2}$). We deduce that for any $p \in [1,2)$ and any $\lambda > 0$,

$$\int_{\mathbb{R}^d} e^{\lambda|x|^p}\alpha(dx) < +\infty$$

which corresponds to the last assumption of [1]...

**Step 1: $W^{(k)}$ bounded from above.** In this case, we have for all $\lambda > 0$,

$$\mathbb{E}[e^{\lambda|W^{(k)}|(X_1,\ldots,X_k)}] < +\infty. \tag{5.37}$$

In principle, large deviations for the U−statistic, under $\mathbb{P} := \alpha^{\otimes N}$, $(L_n, U_n(W^{(2)}),\ldots,U_n(W^{(N)}))$ satisfies a large deviations principle on $\mathcal{M}_1(\mathbb{R}^d)\times\mathbb{R}^{N-1}$ of good rate function $\mathbf{I}_U$. As

$$\sum_{k=2}^N U_n(W^{(k)}) \quad \text{is continuous in} \quad (L_n, U_n(W^{(2)}),\ldots,U_n(W^{(N)})), \tag{5.38}$$

$$\forall p > 1, \quad \limsup_{n\to+\infty}\frac{1}{n}\log\mathbb{E}\Big[e^{-np\sum_{k=2}^N U_n(W^{(k)})}\Big] < +\infty,$$

by what precedes and the theorem of R.Ellis, we deduce that $\mu_n((L_n, U_n(W^{(2)}),\ldots,U_n(W^{(N)})) \in \cdot)$ satisfies a large deviations principle with rate function defined by

$$\tilde{I}(\mu, x_2,\ldots,x_N) = \mathbf{I}_U(\mu, x_2,\ldots,x_N) + \sum_{k=2}^N x_k - \inf_{\mu,x_2,\ldots,x_N}\Big\{\mathbf{I}_U(\mu,x_2,\ldots,x_N) + \sum_{k=2}^N x_k\Big\}. \tag{5.39}$$

So

$$\tilde{I}(\mu, x_2,\ldots,x_N) = \begin{cases} \mathbf{E}_W[\mu] - \inf_\eta \mathbf{E}_W[\eta] & \text{if} \quad \mathbf{H}[\mu|\alpha] < +\infty, \quad \forall k, \quad x_k = \mathbf{W}^{(k)}[\mu], \\ +\infty & \text{otherwise.} \end{cases} \tag{5.40}$$

We conclude by the principle of contraction that $\mu_n(L_n \in \cdot)$ satisfies a PGD of rate function $\mathbf{H}_W$. Note in this case that $\mathbf{E}_W$ is inf-compact, so $\mathbf{H}_W$ too.

**Step 2: General case.** In this case, for all $L > 0$, we set $W_L^{(k)} := \min(W^{(k)}, L)$. So

$$\mathbf{E}_{W_L}[\mu] = \begin{cases} \mathbf{H}[\mu|\alpha] + \sum_{k=2}^N \mathbf{W}_L^{(k)}[\mu] & \text{if} \quad \mathbf{H}[\mu|\alpha] < +\infty, \\ +\infty & \text{otherwise.} \end{cases} \tag{5.41}$$

is inf-compact on $\mathcal{M}_1(\mathbb{R}^d)$ by **step 1**. This proves that $\mathbf{H}_W$ is also inf -compact by passing to the monotonous limit. For all closed $\mathcal{F} \subset \mathcal{M}_1(\mathbb{R}^d)$ and $L > 0$, we have

$$\begin{aligned} \mu_n^*(L_n \in \mathcal{F}) = \int \mathbb{1}_{L_n\in\mathcal{F}} \exp\Big(-n\sum_{k=2}^N U_n(W^{(k)})\Big)d\alpha^{\otimes n} &\leq \int \mathbb{1}_{L_n\in\mathcal{F}} \exp\Big(-n\sum_{k=2}^N U_n(W_L^{(k)})\Big)d\alpha^{\otimes n} \\ &\leq \exp\Big(-n\inf_{\mu\in\mathcal{F}}\mathbf{E}_{W_L}[\mu] + \mathbf{o}(n)\Big) \end{aligned} \tag{5.42}$$

and this last inequality is given by the LDP for the U−statistic and the Varadhan-Laplace lemma. It follows

$$\limsup_{n\to+\infty}\frac{1}{n}\log\mu_n^*(L_n\in\mathscr{F})\le-\inf_{\mu\in\mathscr{F}}\mathbf{E}_{W_L}[\mu]\Longrightarrow\limsup_{n\to+\infty}\frac{1}{n}\log\mu_n^*(L_n\in\mathscr{F})\le-\inf_{\mu\in\mathscr{F}}\mathbf{E}_W[\mu] \tag{5.43}$$

by monotone limit and inf-compactness. In particular, for $\mathscr{F}=\mathscr{M}_1(\mathbb{R}^d)$, we deduce that

$$\limsup_{n\to+\infty}\left\{\frac{1}{n}\log Z_n-\log C\right\}\le-\inf_{\mu\in\mathscr{M}_1(\mathbb{R}^d)}\mathbf{E}_W[\mu]. \tag{5.44}$$

By the lower bound of the large deviations for $L_n$ under $\mu_n^*$ obtained, this upper bound and given that $\mathbf{E}_W[\mu]=+\infty$ if for a $k\in\{2,\dots,N\}$, $W^{(k)}\notin L^1(\mu^{\otimes k})$, we derive that

$$\lim_{n\to+\infty}\left\{\frac{1}{n}\log Z_n-\log C\right\}=-\inf_{\mu\in\mathscr{M}_1(\mathbb{R}^d)}\mathbf{E}_W[\mu] \tag{5.45}$$

which is a finite quantity by assumptions and inf-compactness. With this equality, we thus obtain upper and lower bounds of large deviations for $\{\mu_n(L_n\in\cdot)\}_{n\ge N}$.

□

**Proposition 5.8** (Sanov's theorem for the Wasserstein metric by Wang et.al)**.** *Let $(X_n)_{n\ge1}$ be a sequence of independent random variables, identically distributed, with values in $\mathbb{R}^d$ endowed with one of its norms that we will denote $||\cdot||$ and law $\alpha$. We have equivalence between the following two assertions*

*(i)* $(\mathbb{P}(L_n\in\cdot))_{n\ge1}$ *satisfies a principle of large deviations on the Wasserstein space $(\mathscr{M}_1^p(\mathbb{R}^d),\mathscr{W}_p)$ with speed $n$ and good rate function $\mathbf{H}[\cdot|\alpha]$ .*

*(ii)*

$$\forall\lambda>0\quad x_0\in\mathbb{R}^d,\quad\int_{\mathbb{R}^d}e^{\lambda||x-x_0||^p}\alpha(dx)<+\infty. \tag{5.46}$$

*Proof.* Since we have established a LDP for the random empirical measure $L_n$ under $\mu_n$ on $\mathscr{M}_1(\mathbb{R}^d)$ equipped with the topology of weak convergence, it suffices to prove the exponential tension of $(\mu_n(L_n\in\cdot))_{n\ge N}$ on $(\mathscr{M}_1^p(\mathbb{R}^d),\mathscr{W}_p)$. Let $K\subset\mathscr{M}_1^p(\mathbb{R}^d)$ be compact and $(a,b)\in[1,+\infty]^2$ a pair of conjugate exponents ($\frac{1}{a}+\frac{1}{b}=1$). By Holder's inequality, we have

$$\begin{aligned}\mu_n(L_n\notin K)&=\frac{C^n}{Z_n}\int\mathbb{1}_{L_n\notin K}\exp\Big(-n\sum_{k=2}^{N}U_n(W^{(k)})\Big)d\alpha^{\otimes n}\\&\le\frac{C^n}{Z_n}\Big(\alpha^{\otimes n}(L_n\notin K)\Big)^{\frac{1}{a}}\Big(\int\exp\Big(-nb\sum_{k=2}^{N}U_n(W^{(k)})\Big)d\alpha^{\otimes n}\Big)^{\frac{1}{b}}.\end{aligned} \tag{5.47}$$

It is deduced that

$$\begin{aligned}\limsup_{n\to+\infty}\frac{1}{n}\log\mu_n(L_n\notin K)&\le\frac{1}{a}\limsup_{n\to+\infty}\frac{1}{n}\log\alpha^{\otimes n}(L_n\notin K)-\limsup_{n\to+\infty}\frac{1}{n}\log\frac{Z_n}{C^n}\\&\quad+\frac{1}{b}\limsup_{n\to+\infty}\frac{1}{n}\int\exp\Big(-n\sum_{k=2}^{N}U_n(bW^{(k)})\Big)d\alpha^{\otimes n}.\end{aligned} \tag{5.48}$$

Now the right-hand side of this inequality is upper bounded by

$$\frac{1}{a}\limsup_{n\to+\infty}\frac{1}{n}\log\alpha^{\otimes n}(L_n\notin K)+\inf_{\mu\in\mathscr{M}_1(\mathbb{R}^d)}\mathbf{E}_W[\mu]-\frac{1}{b}\inf_{\mu\in\mathscr{M}_1(\mathbb{R}^d)}\mathbf{E}_{bW}[\mu], \tag{5.49}$$

and from the above, $\inf_{\mu\in\mathscr{M}_1(\mathbb{R}^d)}\mathbf{E}_W[\mu]$ and

$$\inf_{\mu\in\mathscr{M}_1(\mathbb{R}^d)}\mathbf{E}_{bW}[\mu]:=\inf_{\mu\in\mathscr{M}_1(\mathbb{R}^d)}\Big\{\mathbf{H}[\mu|\alpha]+\sum_{k=2}^{N}\int bW^{(k)}d\mu^{\otimes k}\Big\},\quad\text{are finite quantities.} \tag{5.50}$$

Under (**H1**), (**H2**) and (**H3**) in $\mathbb{HMV}$3.1, the LDP holds for $L_n$ under $\alpha^{\otimes n}$ on the Wasserstein space. So, for all $L>0$, there is a compact $K_L\subset\mathscr{M}_1^p(\mathbb{R}^d)$ such that

$$\limsup_{n\to+\infty}\frac{1}{n}\log\alpha^{\otimes n}(L_n\notin K_L)\le-aL-a\inf_{\mu\in\mathscr{M}_1(\mathbb{R}^d)}\mathbf{E}_W[\mu]+\frac{a}{b}\inf_{\mu\in\mathscr{M}_1(\mathbb{R}^d)}\mathbf{E}_{bW}[\mu]. \tag{5.51}$$

It follows that

$$\limsup_{n\to+\infty}\frac{1}{n}\log\mu_n(\mathrm{L}_n\notin\mathrm{K}_\mathrm{L})\leq -\mathrm{L}. \tag{5.52}$$

□

**Uniqueness of invariant measure.** The assumptions on the interaction potentials and the confinement potential ensure the existence (via the inf-compactness of the entropy functional proven in Section 5.2 and [1]) of an invariant measure (global minimum point for the entropy functional) for the McKean-Vlasov process obtained by propagation of chaos. It remains to prove the uniqueness.

**Proposition 5.9** (Fixed point uniqueness). *Under* (**H1**), (**H2**) *and* (**H3**) *in* $\mathbb{HMV}$*3.1, there exists* $\mu_\infty$ *a minimizer of* $\mathbf{H}_\mathrm{W}$. *We have equivalently*

- ▹ *Critical points of energy:* $\mu_\infty\in\{\nu,\quad \frac{\delta\boldsymbol{E}_\mathrm{W}}{\delta m}(\nu,\cdot)=0\}\neq\emptyset$.
- ▹ *Fixed points:* $\mu_\infty\in\{\nu,\quad \Phi(\nu)=\nu\}\neq\emptyset$.
- ▹ *Maxwellians:* $\mu_\infty$ *is also an invariant probability of the McKean-Vlasov process:* $\mu_\infty\mathscr{L}_{\mu_\infty}=0$ *or equivalently* $f_\infty:=\frac{d\mu_\infty}{dx}$ *(existence of density) satisfies*

$$\mathbf{div}\Big(f_\infty\nabla\frac{\delta\mathrm{H}}{\delta m}(\mu_\infty,\cdot)+\nabla f_\infty\Big)=0.$$

*The contraction assumption* (**H5**) *in* $\mathbb{HMV}$*3.1 ensures the uniqueness of* $\mu_\infty$.

*Proof.* To do this, we will use the characterization of the local extrema of a differentiable functional in the sense of Fréchet (flat derivation) on an open set. Let

$$\begin{aligned}\mathscr{O}&:=\Big\{\mu\in\mathscr{P}(\mathbb{R}^d),\quad \mathbf{H}[\mu|\alpha]<+\infty,\quad \forall k,\quad \int \mathrm{W}^{(k),-}d\mu^{\otimes k}<+\infty\Big\}\\&=\mathbf{H}[\cdot|\alpha]^{-1}(]-\infty,+\infty[)\bigcap\Psi^{-1}(]-\infty,+\infty[^{\mathrm{N}-1}),\end{aligned} \tag{5.53}$$

with

$$\Psi:\mu\longmapsto\Big(\int\mathrm{W}^{(2),-}d\mu^{\otimes 2},\ldots,\int\mathrm{W}^{(\mathrm{N}),-}d\mu^{\otimes\mathrm{N}}\Big). \tag{5.54}$$

We know that $\mathbf{E}_\mathrm{W}\equiv+\infty$ over $\mathscr{O}^\mathrm{C}$. By Fréchet differentiability of the relative entropy $\mathbf{H}[\cdot|\alpha]$ and of $\Psi$ on $\mathscr{M}_1(\mathbb{R}^d)$ endowed with its structure of differential Fréchet manifold, $\mathscr{O}$ is open as an intersection of open sets. We deduce that the local extrema (here minimum) of $\mathbf{E_W}$ are critical points on $\mathscr{O}$, i.e. $\mu\in\mathscr{O}$ such that

$$\mathrm{Z}_\mu:=\int e^{-\frac{\delta\mathrm{F}}{\delta m}(\mu,x)-\mathrm{V}(x)}dx<+\infty,\qquad \frac{\delta\mathbf{E}_\mathrm{W}}{\delta m}(\mu,\cdot)\equiv 0\Longleftrightarrow\mu(dx)=\frac{1}{\mathrm{Z}_\mu}e^{-\frac{\delta\mathrm{F}}{\delta m}(\mu,x)-\mathrm{V}(x)}dx. \tag{5.55}$$

According to the hypothesis (**H5**) of $\mathbb{HMV}$3.1, we have

$$\mathbf{d}_{Lip}(\Phi(\mu),\Phi(\nu))\leq k\mathbf{d}_{Lip}(\mu,\nu), \tag{5.56}$$

and since there is a fixed point, suppose by absurd that there is more than one, i.e. there is $\mu_1,\mu_2\in\mathscr{O}$ such that $\mu_1\neq\mu_2$ and for all $i$, $\Phi(\mu_i)=\mu_i$. It follows that $k\geq 1$ which is absurd because $k<1$. □

**Cesãro tensorial: About entropies and Fisher Informations.** We will establish *convergences in entropy and Fisher information* which are useful for the proof of the *exponential decrease of the mean field entropy* and the establishment of the *nonlinear Talagrand inequality*.

**Proposition 5.10** (**H**-Tensorization). *For any probability measure* $\nu$ *on* $\mathbb{R}^d$ *such that* $\mathbf{H}[\nu|\alpha]<+\infty$, *we have:*

$$\frac{1}{n}\mathbf{H}[\nu^{\otimes n}|\mu_n]\overset{n\to+\infty}{\longrightarrow}\mathbf{H}_\mathrm{W}[\nu],\quad\text{where}\quad \mu_n\quad\text{is defined in}\quad \mathrm{E}q.\ (3.8). \tag{5.57}$$

*Proof.* For $\mu$ such that $\mu\ll\alpha$ and for all $k\in\{2,\ldots,\mathrm{N}\}$, $\mathrm{W}^{(k),-}\in\mathrm{L}^1(\mu^{\otimes k})$, we have

$$\begin{aligned}\frac{1}{n}\mathbf{H}[\mu^{\otimes n}|\mu_n]&=\frac{1}{n}\mathbb{E}_{\mu^{\otimes n}}\Big[\frac{d\mu^{\otimes n}}{d\alpha^{\otimes n}}+n\sum_{k=2}^{\mathrm{N}}\mathrm{U}_n(\mathrm{W}^{(k)})+\log\frac{\mathrm{Z}_n}{\mathrm{C}^n}\Big]\\&=\mathbf{H}[\mu|\alpha]+\sum_{k=2}^{\mathrm{N}}\int\mathrm{W}^{(k)}d\mu^{\otimes n}+\frac{1}{n}\log\mathrm{Z}_n-\log\mathrm{C}.\end{aligned} \tag{5.58}$$

We recall that $\alpha(dx) := \frac{e^{-V(x)}}{C}dx$. Under the assumption (**H2**) in $\mathbb{HMV}$3.1, we know that $\exists\, \lambda_0 > 0$ such that:

$$\int_{\mathbb{R}^d} e^{\lambda_0 |x|^2}\alpha(dx) < +\infty. \tag{5.59}$$

By asking:

$$\widetilde{Z}_n := \int_{(\mathbb{R}^d)^n} e^{-n\sum_{k=2}^N U_n(W^{(k)})}\alpha^{\otimes n}(dx_1,\ldots,dx_n), \tag{5.60}$$

we get: (by Fubini-Tonelli)

$$\mu_n(dx) = \frac{C^n}{Z_n}e^{-n\sum_{k=2}^N U_n(W^{(k)})}\alpha^{\bigotimes n}(dx). \tag{5.61}$$

Let $\nu \in \mathscr{M}_1(\mathbb{R}^d)$ be such that $\mathbf{H}[\nu|\alpha] < +\infty$. Since

$$\mathbf{H}[\nu^{\otimes k}|\alpha^{\otimes k}] = k\mathbf{H}[\nu|\alpha], \tag{5.62}$$

$x \longmapsto e^{\lambda_0|x|^2} \in L^1(\alpha)$ and

$$\forall k \in \{2,\ldots,N\} \quad \forall x \in \mathbb{R}^{kd}, \quad |W^{(k)}(x)| \le \beta(1 + \sum_{j=1}^k ||x_j||^2) \tag{5.63}$$

by boundedness of its hessian $\nabla^2 W^{(k)}$ (hypothesis (**H1**) in $\mathbb{HMV}$3.1), according to Donsker-Varadhan variational formula of entropy, we have $W^{(k)} \in L^1(\nu^{\otimes k})$. We have successively: (by a direct calculation and application of the Fubini-Tonelli theorem)

$$\frac{1}{n}\mathbf{H}[\nu^{\otimes n}|\mu_n] = \frac{1}{n}\mathbf{Ent}_{\mu_n}\Big[\frac{d\nu^{\otimes n}}{d\mu_n}\Big] = \frac{1}{n}\int_{(\mathbb{R}^d)^n}\log\Big(\frac{d\nu^{\otimes n}}{d\mu_n}\Big)d\nu^{\otimes n} \tag{5.64}$$

We deduce that:

$$\frac{1}{n}\mathbf{H}[\nu^{\otimes n}|\mu_n] = \frac{1}{n}\int\sum_{i=1}^n \log\Big(\frac{d\nu}{d\alpha}(x_i)\Big)d\nu^{\otimes n} + \sum_{k=2}^N\int U_n(W^{(k)}d\nu^{\otimes n} + \frac{1}{n}\log(\widetilde{Z}_n) \tag{5.65}$$

$\lim_{n\to+\infty}\frac{1}{n}\log(\widetilde{Z}_n) = -\inf_{\eta\in\mathscr{M}_1(\mathbb{R}^d)} \mathbf{E}_W[\eta]$ (see Theorem 5.2),

$$\frac{1}{n}\int\sum_{i=1}^n \log\Big(\frac{d\nu}{d\alpha}(x_i)\Big)d\nu^{\otimes n} = \mathbf{H}[\nu|\alpha] \tag{5.66}$$

and finally, we also have: (see Proposition 5.1)

$$\sum_{k=2}^N\int U_n(W^{(k)})d\nu^{\otimes n} = \sum_{k=2}^N\int W^{(k)}(x)\nu^{\otimes k}(dx). \tag{5.67}$$

Thereby:

$$\frac{1}{n}\mathbf{H}[\nu^{\otimes n}|\mu_n] \overset{n\to+\infty}{\longrightarrow} \mathbf{H}[\nu|\alpha] + \sum_{k=2}^N\int W^{(k)}(x)\nu^{\otimes k}(dx) - \inf_{\eta\in\mathscr{M}_1(\mathbb{R}^d)}\mathbf{E}_W[\eta] = \mathbf{H}_W[\nu]. \tag{5.68}$$

What needed to be proven. □

**Proposition 5.11** (**I**-Tensorization)**.** *If* $\mathbf{I}[\nu|\alpha] < +\infty$*, we have:*

$$\frac{1}{n}\mathbf{I}[\nu^{\otimes n}|\mu_n] \overset{n\to+\infty}{\longrightarrow} \mathbf{I}_W[\nu]. \tag{5.69}$$

*Proof.* For any probability measure $\nu$ on $\mathbb{R}^d$ such that $\mathbf{I}[\nu|\alpha] < +\infty$, by the Lyapunov condition (**H2**) in $\mathbb{HMV}$3.1 on the potential V, we have:

$$c_1\int |x|^2 d\nu \le c_2 + \mathbf{I}[\nu|\alpha] < +\infty. \tag{5.70}$$

As the second order derivatives of $W^{(k)}$ are bounded by the condition (**H1**) in $\mathbb{HMV}$3.1 on its Hessian, $\nabla_{x_j}W^{(k)}$ has a linear increase. So $\nabla_{x_j}W^{(k)} \in L^2(\nu^{\otimes k})$. By the law of large numbers for independent and identically distributed sequences, we have successively:

$$\frac{1}{n}\mathbf{I}[\nu^{\otimes n}|\mu_n] = \frac{1}{4n}\int \left\|\nabla \log\left(\frac{d\nu^{\otimes n}}{d\mu_n}\right)\right\|^2 d\nu^{\otimes n} \tag{5.71}$$
$$= \frac{1}{4n}\int \sum_{i=1}^{n}\left\|\nabla_{x_i}\log\left(\frac{d\nu^{\otimes n}}{d\alpha^{\otimes n}}\right) + n\sum_{k=2}^{N}\nabla_{x_i}U_n(W^{(k)})\right\|^2 d\nu^{\otimes n}$$
$$= \frac{1}{4}\int \left\|\nabla \log\left(\frac{d\nu}{d\alpha}\right)(x_1) + n\sum_{k=2}^{N}\nabla_{x_1}U_n(W^{(k)})\right\|^2 d\nu^{\otimes n}$$
$$\overset{n\to+\infty}{\longrightarrow} \frac{1}{4}\int \left\|\nabla \log\left(\frac{d\nu}{d\alpha}\right)(y) + \sum_{k=2}^{N}\sum_{j=1}^{k}\int \nabla_{x_j}W^{(k)}(x_1,\dots,x_{j-1},y,x_{j+1},\dots,x_k)\nu^{\otimes k-1}\Big(\prod_{i=1,i\neq j}^{k}dx_i\Big)\right\|^2 \nu(dy)$$
$$= \mathbf{I}_W[\nu].$$

□

We recall the tensorisation property of relative entropy: The Proposition 5.12 on entropy and tensor product allows us, in what follows, to show the exponential decreasing of mean-field entropy along the flow of solution distributions of the McKean-Vlasov equation associated with the particle system.

**Proposition 5.12** (Relative entropy and tensor product)**.** *Let* $\prod_{i=1}^{N}\alpha_i$ *and* $Q$ *respectively be a product probability measure and a probability measure defined on* $E_1\times\cdots\times E_N$ *a product of Polish spaces. Denoting* $Q_i$ *the marginal distribution of* $x_i$ *under* $Q$*, we have:*

$$\mathbf{H}[Q|\prod_{i=1}^{N}\alpha_i] \geq \sum_{i=1}^{N}\mathbf{H}[Q_i|\alpha_i]. \tag{5.72}$$

*Proof.* See Appendix A.6 or [2]. □

**Proposition 5.13** (Relative entropy and Boltzmann measure)**.** *Let* $\mu$ *be a probability measure on a Polish space* $E$ *and* $U: E\longrightarrow(-\infty,+\infty]$ *be a measurable potential such that:*

$$\int e^{-pU}d\mu < +\infty \tag{5.73}$$

*for some* $p>1$*. Considering the Boltzmann probability measure* $\mu_U := \frac{e^{-U}}{C}d\mu$*, if for some measure* $\nu$*,* $\mathbf{H}[\nu|\mu_U]<+\infty$*, we have successively:*

*(i)* $\mathbf{H}[\nu|\mu]<+\infty$ *and* $U\in L^1(\nu)$.

*(ii)*
$$\mathbf{H}[\nu|\mu_U] = \mathbf{H}[\nu|\mu] + \int U d\nu + \log\int e^{-U}d\mu. \tag{5.74}$$

*Proof.* See Appendix A.6 or [2]. □

**Functional and transportation inequalities.** Functional inequalities are powerful tools to quantify the trend to equilibrium of Markov semigroups and have a wide range of important applications to the concentration of measure phenomenon and hypercontractivity. $\forall n$, we recall that $\mu_n(t) := \mathbb{P}\circ(X_t^n)^{-1}$ and $\beta_n := \rho_{LS}(\mu_n)$.

**Theorem 5.3** (Transportation inequalities)**.** *Under the assumptions in* $\mathbb{HMV}$ *3.1, we have*

*(i)*
$$\mathbf{H}[\mu_n(t)|\mu_n] \leq \mathbf{H}[\mu_n(0)|\mu_n]e^{-\beta_n\frac{t}{2}} = \mathbf{H}[\mu_0^{\bigotimes n}|\mu_n]e^{-\beta_n\frac{t}{2}}; \tag{5.75}$$
$$\rho_{LS}(\mu_n)\mathbf{H}[\cdot|\mu_n] \leq 2\mathbf{I}[\cdot|\mu_n]; \tag{5.76}$$
$$\rho_{LS}(\mu_n)\mathscr{W}_2^2(\cdot,\mu_n) \leq 2\mathbf{H}[\cdot|\mu_n]. \tag{5.77}$$

*(ii)* $\exists!\ \mu_\infty\in\mathscr{M}_1^2(\mathbb{R}^d)$ *such that: (Section 5.2.Proposition 5.9)*

$$\mu_\infty = \mathbf{argmin}\Big\{\mathbf{H}_W[\nu],\nu\in\mathscr{M}_1(\mathbb{R}^d)\Big\}, \tag{5.78}$$

*with* $\mathbf{H}_W$ *the mean field entropy.*

*(iii)* $\rho_{LS} := \limsup_{n\to+\infty} \rho_{LS}(\mu_n) > 0$ *checks:*

$$\forall \nu \in \mathcal{M}_1^2(\mathbb{R}^d), \quad \rho_{LS}\mathbf{H}_W[\nu] \le 2\mathbf{I}_W[\nu] \quad \text{and} \quad \rho_{LS}\mathcal{W}_2^2(\nu,\mu_\infty) \le 2\mathbf{H}_W[\nu]. \tag{5.79}$$

*We say that we have a nonlinear log-Sobolev inequality for the first inequality and a Talagrand transport inequality for the second.*

*Proof of Theorem 5.3.* The logarithmic Sobolev inequality of constant $\beta_n := \rho_{LS}(\mu_n)$ for $\mu_n$ given by (**H4**) in $\mathbb{HMV}$3.1, the large deviations principle ( Sanov's theorem) in Section 5.2 and the uniqueness of the minimum argument $(\mu_\infty)$ in Proposition 5.9 of the mean field entropy ensure that we have successively:

- ▷ $\forall\mu$ such as $\mathbf{H}[\mu|\alpha] < +\infty$, (Section 5.2.Proposition 5.10.Proposition 5.11)

$$\frac{1}{n}\mathbf{H}[\mu^{\otimes n}|\mu_n] \overset{n\to+\infty}{\longrightarrow} \mathbf{H}_W[\mu] \quad \text{and} \quad \frac{1}{n}\mathbf{I}[\mu^{\otimes n}|\mu_n] \overset{n\to+\infty}{\longrightarrow} \mathbf{I}_W[\mu]. \tag{5.80}$$

- ▷ Equivalence between Sobolev's inequality, exponential decay of entropy and Talagrand's second inequality for Gibbs measures (Otto-Villani,[34],[20])

$$\beta_n\mathbf{H}[\cdot|\mu_n] \le 2\mathbf{I}[\cdot|\mu_n] \quad \text{and} \quad \beta_n\mathcal{W}_2^2(\cdot,\mu_n) \le 2\mathbf{H}[\cdot|\mu_n]. \tag{5.81}$$

- ▷ *Chaos propagation.* (Theorem 5.1) Denoting $(\mu_t)_{t\ge0}$ the flow of solution distributions of the McKean-Vlasov equation associated with the particle system defined by the U− statistic and the confinement potential, if $\mu_0 \in \mathcal{M}_1^2(\mathbb{R}^d)$, then for any non-empty set $I \subset \mathbb{N}^*$ of finite cardinality, $\mathbb{P}_{(X_t^n(i))_{i\in I}}$ converges in metric $L^2-$Wasserstein to $\mu_t^{\otimes \mathbf{Card}(I)}$ (arrow (**1**) in Fig. 1).

- ▷ Denoting $\mu_n^{(i)}$ the i-th marginal distribution of $\mu_n$, we have by uniqueness and LDP (arrow (**3**) in Fig. 1)

$$\mu_n^{(i)} \overset{\mathcal{L}}{\longrightarrow} \mu_\infty. \tag{5.82}$$

- ▷ By symmetry of $\mu_n$, all its marginal distributions are identical and as

$$\mathcal{W}_2^2(\mu^{\otimes n},\mu_n) \ge \sum_{i=1}^n \mathcal{W}_2^2(\mu_n^{(i)},\mu) = n\mathcal{W}_2^2(\mu_n^{(1)},\mu), \tag{5.83}$$

  we deduce:

$$n\beta_n\mathcal{W}_2^2(\mu_n^{(1)},\mu) \le 2\mathbf{H}[\mu^{\otimes n}|\mu_n]. \tag{5.84}$$

By equivalence of the logarithmic Sobolev inequality to the exponential decrease of entropy along the semigroup, we have (arrow (**2**) in Fig. 1)

$$\mathbf{H}[\mu_n(t)|\mu_n] \le \mathbf{H}[\mu_n(0)|\mu_n]e^{-\beta_n\frac{t}{2}} = \mathbf{H}[\mu_0^{\otimes n}|\mu_n]e^{-\beta_n\frac{t}{2}}, \quad \mu_n(t) := \mathbb{P}\circ(X_t^n)^{-1}. \tag{5.85}$$

And by lower semi-continuity of the Wasserstein metric, we deduce the *nonlinear* $T_2-$*Talagrand inequality* given by (arrow (**4**) in Fig. 1)

$$\rho_{LS}\mathcal{W}_2^2(\mu,\mu_\infty) \le \rho_{LS}\liminf_{n\to+\infty}\mathcal{W}_2^2(\mu,\mu_n^{(1)}) \le 2\mathbf{H}_W[\mu], \quad \rho_{LS} := \limsup_{n\to+\infty}\beta_n > 0. \tag{5.86}$$

We also have the *nonlinear logarithmic Sobolev inequality* given by (arrow (**4**) in Fig. 1)

$$\rho_{LS}\mathbf{H}_W[\cdot] \le 2\mathbf{I}_W[\cdot]. \tag{5.87}$$

□

**In Kinetic case.** We consider $\mathcal{H}_n := \Delta_x - \nabla_x S_{1,n}\cdot\nabla = \mathcal{L}_n$ the elliptical generator associated with $\mu_{1,n} = \mu_n$ .

*Remark* 5.4. $\mathcal{L}_{Z,n}$ admits the following Hormander form

$$\mathcal{L}_{Z,n} = X_0 + Y + \sum_{i=1}^n\sum_{j=1}^d X_{i,j}^2, \quad X_{i,j} = \frac{\partial}{\partial v_{i,j}} \quad X_0 = -v\cdot\nabla_v \quad Y = \nabla S_{1,n}\cdot\nabla_v - v\cdot\nabla_x \tag{5.88}$$

The family

$$\left\{X_{1,1},\ldots,X_{i,j},\ldots,X_{i,j},\cdots,X_{n,d},[Y,X_{1,1}],\ldots,[Y,X_{i,j}],\ldots,[Y,X_{n,d}]\right\} \tag{5.89}$$

form a basis of $\mathbb{R}^{2nd}$ at any point. Which implies by Hormander's theorem that $\mathcal{L}_{Z,n}$ is hypoelliptic. Moreover, $\mathcal{L}_{Z,n}$ is non-symmetric, i.e. in $L^2(\mu_Z^n)$, we have :

$$\mathcal{L}_{Z,n}^* = \mathcal{L}_{Z,n} - 2Y \Longrightarrow (\mathcal{L}_{Z,n}^*,\mathcal{D}(\mathcal{L}_{Z,n}^*)) \quad \text{is not a closed extension of} \quad (\mathcal{L}_{Z,n},\mathcal{D}(\mathcal{L}_{Z,n})). \tag{5.90}$$

The following known lemma is a key to the Lyapunov type conditions. We include its simple proof for completeness.

**Proposition 5.14** (Lemma.8 in [3]). *For any function $\varphi \in \mathscr{C}^2(\mathbb{R}^{nd})$ strictly positive ($\varphi > 0$), we have*

$$\forall \psi \in \mathscr{H}^1(\mu_{1,n}), \quad \int -\frac{\mathscr{H}_n \varphi}{\varphi} \psi^2 d\mu_{1,n} \leq \int |\nabla \psi|^2 d\mu_{1,n}. \tag{5.91}$$

*Proof of Proposition 5.14.* Indeed, by integrating by parts, we successively obtain

$$\begin{aligned}\int -\frac{\mathscr{H}_n \varphi}{\varphi} \psi^2 d\mu_{1,n} &\leq \int \left\langle \nabla \varphi, \nabla \frac{\psi^2}{\varphi} \right\rangle d\mu_{1,n} \\ &\leq \int \left\langle \nabla \varphi, \frac{2\psi \nabla \psi}{\varphi} - \frac{\psi^2 \nabla \varphi}{\varphi^2} \right\rangle d\mu_{1,n} \\ &\leq \int |\nabla \psi|^2 d\mu_{1,n}.\end{aligned} \tag{5.92}$$

And this last inequality follows from the inequality

$$\left\langle 2\psi \nabla \psi, \frac{\nabla \varphi}{\varphi} \right\rangle \leq \frac{\psi^2 |\nabla \varphi|^2}{\varphi^2} + |\nabla \psi|^2. \tag{5.93}$$

□

This second Proposition 5.15 is the heart of the proof of Theorem 4.3: this proposition is inspired by [3, Lemma.10] for the two-body interaction. It uses Lyapunov conditions, yet well know for being highly dimensional, but at the marginal level, thus providing results independent of the number of particles.

**Proposition 5.15.** *Under the conditions in* $\mathbb{VFP}$*3.2 giving* **UPI***, there are two constants* $C_1$ *and* $C_2$ *depending on* $N, K, K_1, K_2$ *and* $d$ *(dimension of* $\mathbb{R}^d$*) and such that*

$$\forall \psi \in \mathscr{H}^1(\mu_{1,n}), \quad \int ||\nabla^2 V(x_i)||^2_{\mathbf{op}} \psi^2 d\mu_{1,n} \leq C_1 \int |\nabla_x \psi|^2 d\mu_{1,n} + C_2 \int \psi^2 d\mu_{1,n}. \tag{5.94}$$

*Proof of Proposition 5.15.* This lemma follows from the Lyapunov property, from the particular form of the invariant measure generator[2] $\mu_{1,n}$ and from the previous Proposition 5.14. Indeed, we have:

(i)

$$||\nabla^2 V||^2_{\mathbf{op}} \leq \eta_1 \Big( (1-\gamma) ||\nabla V||^2 - \Delta V \Big) + \eta_2, \tag{5.95}$$

$$\eta_1 := 5K_1^2 \quad \eta_2 := 4K_2^2 + \frac{25 K_1^4 d^2}{4} \quad \gamma := \frac{1}{5}. \tag{5.96}$$

(ii) Since the interactions are Lipschitz, we know that $\forall k \in \{2, \ldots, N\} \exists K^{(k)}$ such that $||\nabla W^{(k)}|| \leq K^{(k)}$. Let $K := \max\{K^{(k)}, \quad k = 2, \ldots, N\}$. It follows that

$$-n \sum_{k=2}^{N} \nabla_{x_i} U_n(W^{(k)}) \cdot \nabla V(x_i) \leq (N-1) K |\nabla V|(x_i) \leq (N-1) \Big( \frac{K^2}{2\gamma} + \frac{\gamma}{2} |\nabla V|^2(x_i) \Big). \tag{5.97}$$

But for $\varphi(x) := e^{\frac{\gamma}{2} V(x_i)}$, we have

$$\frac{\mathscr{H}_n \varphi}{\varphi} = \frac{\mathscr{T}_i \varphi}{\varphi} = \frac{\gamma}{2} \Big( \Delta V(x_i) + (\frac{\gamma}{2} - 1) |\nabla V|^2(x_i) - n \sum_{k=2}^{N} \nabla_{x_i} U_n(W^{(k)}) \cdot \nabla V(x_i) \Big). \tag{5.98}$$

Thereby

$$2 \frac{\mathscr{H}_n \varphi}{\gamma \varphi} \leq \Delta V(x_i) + (\frac{N\gamma}{2} - 1) ||\nabla V||^2(x_i) + \frac{(N-1) K^2}{2\gamma}. \tag{5.99}$$

[2] We have

$$\mathscr{H}_n = \sum_{i=1}^{n} \mathscr{T}_i, \quad \mathscr{T}_i := \Delta_{x_i} - \nabla V(x_i) \cdot \nabla_{x_i} - n \sum_{k=2}^{N} \nabla_{x_i} U_n(W^{(k)}) \cdot \nabla_{x_i}.$$

Moreover, we have

$$(1-\gamma)||\nabla V||^2(x_i) - \Delta V(x_i) \leq -2\frac{\mathcal{H}_n\varphi}{\gamma\varphi} + \frac{(N-1)K^2}{2\gamma}. \tag{5.100}$$

Therefore, by the inequality obtained in **(i)**,

$$||\nabla^2 V(x_i)||^2_{\mathbf{op}} \leq \eta_1\Big(-2\frac{\mathcal{H}_n\varphi}{\gamma\varphi} + \frac{(N-1)K^2}{2\gamma}\Big) + \eta_2 \tag{5.101}$$

Integrating with respect to $\psi^2 d\mu_{1,n}$, we obtain

$$\int ||\nabla^2 V(x_i)||^2_{\mathbf{op}}\psi^2 d\mu_{1,n} \leq \frac{2\eta_1}{\gamma}\int -\frac{\mathcal{H}_n\varphi}{\varphi}\psi^2 d\mu_{1,n} + \Big(\eta_2 + \frac{(N-1)K^2}{2\gamma}\eta_1\Big)\int \psi^2 d\mu_{1,n}. \tag{5.102}$$

And we conclude by the previous Proposition 5.14 that

$$\int ||\nabla^2 V(x_i)||^2_{\mathbf{op}}\psi^2 d\mu_{1,n} \leq C_1\int |\nabla_x\psi|^2 d\mu_{1,n} + C_2\int \psi^2 d\mu_{1,n}, \tag{5.103}$$

where $C_1 = \frac{2\eta_1}{\gamma}$ and $C_2 = \eta_2 + \frac{(N-1)K^2}{2\gamma}\eta_1$.

□

# 6 Proofs of Main Theorems

*Proof of Theorem 4.1.* Indeed, we have the inequality (Proposition 5.12)

$$\frac{1}{n}\mathbf{H}[\mu_n(t)|\alpha^{\otimes n}] \geq \mathbf{H}[\mu_n^{(1)}(t)|\alpha] \tag{6.1}$$

and by lower semi-continuity of relative entropy and propagation of chaos,

$$\liminf_{n\to+\infty} \quad \mathbf{H}[\mu_n^{(1)}(t)|\alpha] \geq \mathbf{H}[\mu_t|\alpha] \quad . \tag{6.2}$$

On the other hand, we have (Theorem 5.2.Theorem 5.3.Eq. (5.58))

$$\frac{1}{n}\mathbf{H}[\mu_n(t)|\mu_n] \leq \frac{1}{n}\mathbf{H}[\mu_0^{\otimes n}|\mu_n]e^{-\beta_n\frac{t}{2}} \quad \text{and} \quad \liminf_{n\to+\infty}\frac{1}{n}\mathbf{H}[\mu_0^{\otimes n}|\mu_n]e^{-\beta_n\frac{t}{2}} = \mathbf{H}_W[\mu_0]e^{-\rho_{LS}\frac{t}{2}}. \tag{6.3}$$

Also, as

$$\mu_n(dx) = \frac{C^n}{Z_n}e^{-n\sum_{k=2}^N U_n(W^{(k)})}\alpha^{\otimes n}(dx), \tag{6.4}$$

we also have

$$\frac{1}{n}\mathbf{H}[\mu_n(t)|\mu_n] = \frac{1}{n}\mathbf{H}[\mu_n(t)|\alpha^{\otimes n}] + \sum_{k=2}^N\int U_n(W^{(k)})d\mu_n(t) + \Big(\frac{1}{n}\log(Z_n) - \log(C)\Big), \tag{6.5}$$

and (Section 5.2.Eq. (5.58))

$$\sum_{k=2}^N\int U_n(W^{(k)})d\mu_n(t) \overset{n\to+\infty}{\longrightarrow} \sum_{k=2}^N\int W^{(k)}d\mu_t^{\otimes k} = \sum_{k=2}^N \mathbf{W}^{(k)}[\mu_t], \quad \frac{1}{n}\log(Z_n) - \log(C) \overset{n\to+\infty}{\longrightarrow} -\inf_{\mu\in\mathcal{M}_1(\mathbb{R}^d)}\mathbf{E}_W[\mu]. \tag{6.6}$$

It is deduced that

$$\forall t \geq 0, \quad \mathbf{H}_W[\mu_0]e^{-\rho_{LS}\frac{t}{2}} \geq \liminf_{n\to+\infty}\frac{1}{n}\mathbf{H}[\mu_n(t)|\mu_n] \geq \mathbf{H}[\mu_t|\alpha] + \sum_{k=2}^N \mathbf{W}^{(k)}[\mu_t] - \inf_{\mu\in\mathcal{M}_1(\mathbb{R}^d)}\mathbf{E}_W[\mu] = \mathbf{H}_W[\mu_t]. \tag{6.7}$$

This completes the proof of the exponential decrease of entropy along the flow. □

*Proof of Theorem 4.2.* Just use the nonlinear $T_2$−Talagrand inequality, i.e.: (Theorem 5.3)

$$\forall t \geq 0, \quad \rho_{LS}\mathcal{W}_2^2(\mu_t,\mu_\infty) \leq 2\mathbf{H}_W[\mu_t]. \tag{6.8}$$

This completes the proof of the desired inequality: We conclude with the Theorem 4.1. □

*Proof of Theorem 4.3.* By the Lyapunov condition in the assumptions 3.2, we can apply Proposition 5.15 and obtain that for any $\psi \in \mathcal{H}^1(\mu_{1,n})$, it holds

$$\int ||\nabla^2 V(x_i)||^2_{\mathbf{op}} \psi^2 d\mu_{1,n} \le C_1 \int |\nabla_x \psi|^2 d\mu_{1,n} + C_2 \int \psi^2 d\mu_{1,n}, \tag{6.9}$$

with $C_1 = \frac{2\eta_1}{\gamma}$ and $C_2 = \eta_2 + \frac{(N-1)K^2}{2\gamma}\eta_1$ for instance which are independent of the number $n$ of particles. It follows that the boundedness condition in Villani's theorem holds. Since the uniform Sobolev inequality implies the uniform Poincaré inequality, we can apply Villani's hypocoercivity theorem ([3, Theorem.3] or [24, Theorem.18 and Theorem.35]), which completes the proof. □

*Proof of Theorem 4.4.* Note that $(\mu^n_Z(t))_{t\ge 0}$ is a solution of a (large dimensional) linear Fokker-Planck equation, for which the exponential decay of the entropy is already known under assumptions including 3.2 (see e.g. [24]). Consider the generator $\mathcal{L}_{Z,n}$ given by Eq. (3.19). Then $\Psi_n := \frac{d\mu^n_Z(t)}{d\mu^n_Z}$, the density of the law of the particle system given by Eq. (3.16) with respect to its equilibrium distribution, solves

$$\partial_t \Psi_n = \mathcal{L}^*_{Z,n} \Psi_n. \tag{6.10}$$

This is a linear kinetic Fokker-Planck equation, for which convergence to equilibrium has been proven by many ways. All we need to check is that the explicit estimates we obtain do not depend on $n$ (see e.g. respectively Theorem.7 and Theorem.10 in [5, [6]]). The key point in Eq. (4.4) is that C and $\xi$ do not depend on $n$: Indeed, as $\mu^n_Z = \mu_{1,n} \otimes \mu_{2,n}$ and these measures satisfy logarithmic Sobolev inequalities of constants $\rho_{LS}(\mu_{1,n}) = \rho$ and $\rho_{LS}(\mu_{2,n}) = 1$, $\mu^n_Z$ satisfies an inequality of logarithmic Sobolev of constant $\rho_{LS}(\mu^n_Z) := \max(\rho, 1)$; moreover, the Hessian of its Hamiltonian is uniformly bounded. This enables us to prove the following: (T2−inequality)

$$\exists \kappa > 0 \quad \forall n \quad \forall t, \quad \mathcal{W}_2^2(\mu^n_Z(t), \mu^n_Z) \le \kappa e^{-\xi t} \mathbf{H}[\mu^n_Z(0)|\mu^n_Z], \quad \mu^n_Z(0) = \mu^{\otimes n}, \quad \mu \in \mathcal{P}_2(\mathbb{R}^d \times \mathbb{R}^d). \tag{6.11}$$

By symmetry, propagation of chaos and Sanov's theorem (LDP), we have respectively

$$\forall i \in \{1,\dots,n\}, \quad n\mathcal{W}_2^2(\mu^{n,(i)}_Z(t), \mu^{n,(i)}_Z) \le \mathcal{W}_2^2(\mu^n_Z(t), \mu^n_Z), \quad \mu^{n,(i)}_Z(t) \overset{n\to+\infty}{\longrightarrow} \mu^{\mathbf{VFP}}_t, \quad \mu^{n,(i)}_Z \overset{n\to+\infty}{\longrightarrow} \mu^Z_\infty. \tag{6.12}$$

We have by lower semi-continuity

$$\forall \mu \in \mathcal{P}_2(\mathbb{R}^d \times \mathbb{R}^d), \quad \mathcal{W}_2^2(\mu, \mu^Z_\infty) \le \liminf_{n\to+\infty} \mathcal{W}_2^2(\mu, \mu^{n,(i)}_Z) \le \kappa \liminf_{n\to+\infty} \frac{1}{n} \mathbf{H}[\mu^n_Z(0)|\mu^n_Z] = \kappa \mathcal{S}[\mu], \tag{6.13}$$

$$\frac{1}{n}\mathbf{H}[\mu^{\otimes n}|\mu^n_Z] = \mathbf{H}[\mu|\alpha \otimes \mathcal{N}(0, \mathbf{Id}_d)] + \sum_{k=2}^N \int U_n(W^{(k)}) d\mu^{\otimes n} + \frac{1}{n}\log(Z_n) - \log(C) \overset{n\to+\infty}{\longrightarrow} \mathcal{S}[\mu] := \mathcal{E}[\mu] - \mathcal{E}[\mu^Z_\infty]. \tag{6.14}$$

According to Eq. (4.4), we have

$$\mathcal{S}[\mu^{\mathbf{VFP}}_t] \le C\mathcal{S}[\mu] e^{-\xi t}. \tag{6.15}$$

It follows that

$$\mathcal{W}_2^2(\mu^{\mathbf{VFP}}_t, \mu^Z_\infty) \le \kappa \mathcal{S}[\mu^{\mathbf{VFP}}_t] \le \kappa C \mathcal{S}[\mu] e^{-\xi t}. \tag{6.16}$$

□

*Proof of Proposition 4.1.* First, note that (4.6) ensures that $\nabla^2 V \ge \rho I_d$ with $\rho > 0$, which in turn implies (**H2**). For (**H3**), this follows from the assumption: $W^{(k),-}(x_1,\dots,x_k) = o(\sum_{j=1}^k V(x_j))$ as $|x_1|^2 + \ldots + |x_k|^2 \to +\infty$.

To prove (**H4**), we use the classical Bakry-Emery criterion. To this end, let us denote by $A_{ij} = \nabla^2_{ij} H_n(x_1,\dots,x_n)$ where we abusively omit the dependence in $(x_1,\dots,x_n)$. Note that $A_{ij}$ is a $d\times d$-matrix. Using the symmetry of the $W^{(k)}$, we have

$$A_{ii} = \nabla^2 V(x_i) + \sum_{k=2}^N k \sum_{(i_1,\dots,i_{k-1})\in I_n^{-i}} \binom{n-1}{k-1}^{-1} \nabla^2_{11} W^{(k)}(x_i, x_{i_1}, \dots, x_{i_{k-1}})$$

and if $i \ne j$,

$$A_{ij} = \sum_{k=2}^N k \sum_{(i_1,\dots,i_{k-2})\in I_n^{-ij}} \binom{n-1}{k-1}^{-1} \nabla^2_{12} W^{(k)}(x_i, x_j, x_{i_1}, \dots, x_{i_{k-2}})$$

where $I_n^{-i}$ (resp. $I_n^{-ij}$) denotes the set of increasing sequences $i_1 < \ldots < i_{k-1}$ of $[\![1,n]\!]\backslash\{i\}$ (resp. $i_1 < \ldots < i_{k-2}$ of $[\![1,n]\!]\backslash\{i,j\}$). With the notations of the proposition, one easily checks one can find a small enough $\varepsilon$ and a large enough $n_\varepsilon$ such that for any $\lambda \leq \varepsilon$, $n \geq n_\varepsilon$ and $(x_1,\ldots,x_n) \in (\mathbb{R}^d)^n$, we have for all $i \in [\![1,n]\!]$,

$$\|A_{ii}^{-1} - \lambda\|_{\mathbf{op}}^{-1} \geq \underline{\lambda} - 2\varepsilon > \sum_{k=2}^{N} k(k-1)\|\nabla_{12}^2 W^{(k)}\|_\infty + \varepsilon \geq \sum_{j\neq i} \|\nabla_{i,j}^2 A_{ij}\|_{\mathbf{op}}.$$

This implies that for any $\lambda \in (-\infty, \varepsilon)$, the matrix $\nabla^2 H_n(x_1,\ldots,x_n) - \lambda I_{nd}$ is block-diagonally dominant and thus invertible. Hence, $\nabla^2 H_n \geq \varepsilon I_{nd}$ which in turn implies (**H4**) by the Bakry-Emery criterion.

Let us now prove that (**H5**) holds. Let $\nu_0, \nu_1 \in \mathscr{P}_2(\mathbb{R}^d)$ and set $\nu_t = (1-t)\nu_0 + t\nu_1$, $t \in [0,1]$. Let $f : \mathbb{R}^d \to \mathbb{R}$ be a 1-Lipschitz smooth function. From the very definition of $\Phi$,

$$\langle \Phi(\nu_t), f\rangle = \frac{1}{Z_{\nu_t}} \int_{\mathbb{R}^d} f(x) e^{-V(x) - \sum_{k=2}^N k \int W^{(k)}(x,y)\nu_t^{\otimes k-1}(dy)} dx, \tag{6.17}$$

so that setting $g_t(x) = -\partial_t(\sum_{k=2}^N k \int W^{(k)}(x,y)\nu_t^{\otimes k-1}(dy))$, we get

$$\frac{d}{dt}\langle \Phi(\nu_t), f\rangle = -\frac{\partial_t Z_{\nu_t}}{Z_{\nu_t}}\langle \Phi(\nu_t), f\rangle + \langle \Phi(\nu_t), f g_t\rangle = \langle \Phi(\nu_t), f g_t\rangle - \langle \Phi(\nu_t), f\rangle\langle \Phi(\nu_t), g_t\rangle = \mathrm{Cov}_{\Phi(\nu_t)}(f, g_t).$$

For a probability $\mu$, let $\mathscr{L}_\mu$ be the operator defined on $\mathscr{C}^2$-functions by

$$\mathscr{L}_\mu f = -\nabla f.\nabla(\frac{\delta F}{\delta m}(\mu,\cdot) + V) + \Delta f, \tag{6.18}$$

with $\frac{\delta F}{\delta m}$ defined by (1.18). Denoting by $\varphi_t$ be the solution of the Poisson equation $f - \Phi(\nu_t)(f) = \mathscr{L}_{\nu_t}\varphi_t$ and using that $\mathscr{L}_{\nu_t}$ is self-adjoint in $L^2(\Phi(\nu_t))$, we get

$$\mathrm{Cov}_{\Phi(\nu_t)}(f, g_t) = \langle \varphi_t, \mathscr{L}_{\nu_t} g_t\rangle_{L^2(\Phi(\nu_t))} = -\langle \nabla\varphi_t, \nabla g_t\rangle_{L^2(\Phi(\nu_t))}.$$

Note that for the second equality, we used the fact that for some $\mathscr{C}^2$-functions $f$ and $g$, $\mathscr{L}_{\nu_t}(f.g) = f\mathscr{L}_{\nu_t} g + g\mathscr{L}_{\nu_t} f + 2\nabla f.\nabla g$. With the help of Cauchy-Schwarz inequality and Lemma 6.1, this leads to

$$|\mathrm{Cov}_{\Phi(\nu_t)}(f, g_t)| \leq \|\nabla\varphi_t\|_{L^2(\Phi(\nu_t))}\|\nabla g_t\|_{L^2(\Phi(\nu_t))} \leq \frac{1}{\underline{\lambda}}\|\nabla g_t\|_{L^2(\Phi(\nu_t))}. \tag{6.19}$$

Let us finally focus on $\nabla g_t$. First, one checks that

$$\begin{aligned}\partial_t \int W^{(k)}(x,y)\nu_t^{\otimes k-1}(dy) &= \int W^{(k)}(x,y)\sum_{j=1}^{k-1}(\nu_1 - \nu_0)(dy_j)\prod_{i\neq j}\nu_t(dy_i) \\ &= (k-1)\int\int W^{(k)}(x,x_2,y)(\nu_1-\nu_0)(dx_2)\nu_t^{\otimes k-2}(dy).\end{aligned} \tag{6.20}$$

Note that we used the symmetry of $W^{(k)}$ and the fact that $|W^{(k)}|$ is subquadratic (due to (**H1**)), which ensures sufficient integrability properties for the above equalities. Now, denoting by $(X_{\nu_0}, X_{\nu_1})$ an optimal coupling of $\nu_0$ and $\nu_1$ for the 1-Wasserstein distance, one obtains

$$\begin{aligned}\left|\nabla_x\left(\int W^{(k)}(x,x_2,y)(\nu_1-\nu_0)(dx_2)\right)\right| &\leq |\mathbb{E}[\nabla_1 W^{(k)}(x, X_{\nu_1}, y) - \nabla_1 W^{(k)}(x, X_{\nu_0}, y)]| \\ &\leq \|\nabla_{12}^2 W^{(k)}\|_{\mathbf{op},\infty}\mathscr{W}_1(\nu_0,\nu_1).\end{aligned} \tag{6.21}$$

Hence,

$$\|\nabla g_t\|_{L^2(\Phi(\nu_t))} \leq \|\nabla g_t\|_\infty \leq \left(\sum_{k=2}^N k(k-1)\|\nabla_{12}^2 W^{(k)}\|_{\mathbf{op},\infty}\right)\mathscr{W}_1(\nu_0,\nu_1),$$

and by (6.19), we get for any smooth 1-Lipschitz function

$$|\langle \Phi(\nu_1), f\rangle - \langle \Phi(\nu_0), f\rangle| \leq \int_0^1 \left|\frac{d}{dt}\langle \Phi(\nu_t), f\rangle\right| dt \leq \frac{\sum_{k=2}^N k(k-1)\|\nabla_{12}^2 W^{(k)}\|_{\mathbf{op},\infty}}{\underline{\lambda}}\mathscr{W}_1(\nu_0,\nu_1).$$

By (4.6), a density argument and the Kantorovich-Rubinstein duality relation, it follows that $\Phi$ is a contraction on $(\mathscr{P}_2(\mathbb{R}^d), \mathscr{W}_1)$. □

*Proof of Proposition 4.2.* We need to show that (**H4**) and (**H5**) are true. Concerning (**H5**), this is a consequence of assumption (4.6), exactly as in the proof of Proposition 4.1. For (**H4**), we apply [35, Theorem 1]. The inspection of the related assumptions is divided in three steps.

**Step 1: Regularity properties.** For any $\mu \in \mathscr{P}(\mathbb{R}^d)$, we consider the standard Langevin process $(X_t^\mu)_{t\geq 0}$ of Hamiltonian $x_1 \mapsto \frac{\delta H}{\delta m}(\mu, x_1) = V(x_1) + \sum_{k=2}^N k \int W^{(k)}(x_1, y)\mu^{\otimes k-1}(dy)$, i.e.

$$dX_t^\mu = \sqrt{2}dB_t^\mu - \nabla\frac{\delta H}{\delta m}(\mu, X_t^\mu)dt. \tag{6.22}$$

This Langevin process admits for invariant distribution $\Phi(\mu)(dx) = \frac{1}{Z_\mu}e^{-\frac{\delta H}{\delta m}(\mu,x)}dx$. The functional H admits first and second-order flat derivatives that are jointly continuous and are $\mathscr{C}^2$ in the spatial variables. $H_n$ the Hamiltonian of the particle system and its derivatives are related to the macroscopic Hamiltonian H and its derivatives by

$$\begin{aligned} H_n(\mathbf{x}) &= nH(\mu_{\mathbf{x}}) \qquad (6.23)\\ &= nH\Big(\frac{1}{n}\sum_{p=1}^n \delta_{x_p}\Big)\\ &= \sum_{p=1}^n V(x_p) + n\sum_{k=2}^N U_n(W^{(k)});\\ \nabla_{x_i} H_n(\mathbf{x}) &= \mathscr{D}_m H(\mu_{\mathbf{x}}, x_i); \qquad (6.24)\\ \nabla^2_{x_j x_i} H_n(\mathbf{x}) &= \nabla\mathscr{D}_m H(\mu_{\mathbf{x}}, x_i)\delta_{ij} + \frac{1}{n}\mathscr{D}_m^2 H(\mu_{\mathbf{x}}, x_i, x_j). \qquad (6.25) \end{aligned}$$

As

$$\frac{\delta^2 H}{\delta m^2}(\mu, x_1, x_2) = \sum_{k=2}^N k(k-1)\int W^{(k)}(x_1, x_2, y)\mu^{\otimes k-2}(dy), \tag{6.26}$$

and that for all $k \in \{2, \ldots, N\}$, $\nabla_{x_1} W^{(k)}$ and $\nabla^2_{x_1 x_2} W^{(k)}$ are bounded, the functions $(\mu, x_1, x_2) \mapsto \nabla_{x_1}\frac{\delta^2 H}{\delta m^2}(\mu, x_1, x_2)$ and $(\mu, x_1, x_2) \mapsto \nabla^2_{x_1 x_2}\frac{\delta^2 H}{\delta m^2}(\mu, x_1, x_2)$ are bounded. These properties imply that Assumptions 1 and 2 of [35, Theorem 1] are true. On the other hand, Assumption 5 exactly corresponds to Assumption $(iii)$ of Proposition 4.2. It thus remains to establish Assumption 3 concerning uniform logarithmic Sobolev inequality for $\Phi(\mu)$ and Assumption 4 concerning uniform Poincaré inequality for the i-th conditional marginal of the invariant distribution of the particle system given by $\mu_n(d\mathbf{x}) = \frac{1}{Z_n}e^{-H_n(\mathbf{x})}d\mathbf{x}$.

**Step 2: Uniform log-Sobolev inequality.** For this step, one uses [27, 28] which gives some sufficient conditions on a potential U to ensure a $\rho_U$-log-Sobolev inequality for $\alpha_U(dx) := \frac{e^{-U(x)}}{Z_U}dx$. More precisely, such a result holds if U is $\mathscr{C}^2$ with lower bounded Hessian and satisfies the following Lyapunov condition: there are two positive constants $c_1^U$ and $c_2^U$ such that

$$\forall x \in \mathbb{R}^d, \quad x\cdot\nabla U(x) \geq c_1^U|x|^2 - c_2^U. \tag{6.27}$$

Thus, in our setting, the uniform log-Sobolev inequality will hold if these assumptions are true uniformly in $\mu$ for $x \mapsto \frac{\delta H}{\delta m}(\mu, x)$. Let us check these conditions: since for all $k \in \{2, \ldots, N\}$, $x \mapsto \nabla_{x_1} W^{(k)}(x)$ and $\nabla^2 W^{(k)}$ are bounded, for all $\mu \in \mathscr{P}_2(\mathbb{R}^d)$ and $x \in \mathbb{R}^d$,

$$\begin{aligned} x\cdot\nabla\frac{\delta H}{\delta m}(\mu, x) &\geq c_1|x|^2 - c_2 + \sum_{k=2}^N k\int x\cdot\nabla_{x_1} W^{(k)}(x, y)\mu^{\otimes k-1}(dy) \qquad (6.28)\\ &\geq c_1|x|^2 - c_2 - \Big(\sum_{k=2}^N k\|\nabla_{x_1} W^{(k)}\|_\infty\Big)|x|; \end{aligned}$$

and

$$\nabla^2\frac{\delta H}{\delta m}(\mu, \cdot) = \nabla^2 V + \sum_{k=2}^N k\int \nabla^2_{x_1} W^{(k)}(\cdot, y)\mu^{\otimes k-1}(dy) \tag{6.29}$$

is bounded from below. As $|x| \leq \varepsilon|x|^2 + C_\varepsilon$ for any $\varepsilon > 0$, there are two positive constants $c_1^\star$ and $c_2^\star$ such that

$$\forall (x, \mu) \in \mathbb{R}^d \times \mathscr{P}_2(\mathbb{R}^d), \quad x\cdot\nabla\frac{\delta H}{\delta m}(\mu, x) \geq c_1^\star|x|^2 - c_2^\star. \tag{6.30}$$

From what precedes, one can conclude $\Phi(\mu)$ satisfies a $\rho$ logarithmic Sobolev inequality with constant $\rho$ independent of $\mu$.

**Step 3: Uniform Poincaré inequality for $\mu_n^{i|-i}$.** First,

$$\frac{\mu_n^{i|-i}(\cdot|\mathbf{x}^{-i})}{dx_i} = \frac{e^{-n\mathrm{H}(\mu_{\mathbf{x}})}}{\int e^{-n\mathrm{H}(\mu_{\mathbf{x}})}dx_i} \tag{6.31}$$

which involves that its log-density satisfies

$$-\nabla_i \ln \mu_n^{i|-i}(x_i|\mathbf{x}^{-i}) = \mathscr{D}_m\mathrm{H}(\mu_{\mathbf{x}}, x_i). \tag{6.32}$$

Second, let us denote $\mu^{\Phi}_{\mathbf{x}^{-i}}(\cdot)$ the probability measure defined by

$$\frac{\mu^{\Phi}_{\mathbf{x}^{-i}}(\cdot)}{dx_i} := \frac{\Phi(\mu_{\mathbf{x}^{-i}})}{dx_i} = \frac{e^{-\frac{\delta\mathrm{H}}{\delta m}(\mu_{\mathbf{x}^{-i}},x_i)}}{\int e^{-\frac{\delta\mathrm{H}}{\delta m}(\mu_{\mathbf{x}^{-i}},x_i)}dx_i}. \tag{6.33}$$

Its log-density satisfies

$$-\nabla_i \ln \mu^{\Phi}_{\mathbf{x}^{-i}}(x_i) = \mathscr{D}_m\mathrm{H}(\mu_{\mathbf{x}^{-i}}, x_i). \tag{6.34}$$

We have (definition of the flat derivative)

$$\mathscr{D}_m\mathrm{H}(\mu_1,x) - \mathscr{D}_m\mathrm{H}(\mu_0,x) = \int_0^1\int_{\mathbb{R}^d}\nabla_x\frac{\delta^2\mathrm{H}}{\delta m^2}(t\mu_1+(1-t)\mu_0,x,y)(\mu_1-\mu_0)(dy)dt. \tag{6.35}$$

By boundedness property of $(\mu,x,y)\mapsto\nabla_x\frac{\delta^2\mathrm{H}}{\delta m^2}(\mu,x,y)$, we have

$$|\mathscr{D}_m\mathrm{H}(\mu_1,x) - \mathscr{D}_m\mathrm{H}(\mu_0,x)| \le ||\nabla_x\frac{\delta^2\mathrm{H}}{\delta m^2}||_\infty||\mu_1-\mu_0||_{\mathrm{TV}}. \tag{6.36}$$

Then, for some constant M, H verifies

$$|\mathscr{D}_m\mathrm{H}(\mu_{\mathbf{x}},x_i) - \mathscr{D}_m\mathrm{H}(\mu_{\mathbf{x}^{-i}},x_i)| \le \frac{\mathrm{M}}{n}; \tag{6.37}$$

then the derivatives can at most differ by $\frac{\mathrm{M}}{n}$. This means that, for $n$ large, the conditional measure $\mu_n^{i|-i}(\cdot|\mathbf{x}^{-i})$ is a *weak log-Lipschitz perturbation* of the measure $\mu^{\Phi}_{\mathbf{x}^{-i}}$. Furthermore, $\mu^{\Phi}_{\mathbf{x}^{-i}}$ satisfies a $\rho$-log-Sobolev, a fortiori, it satisfies a $\rho$-Poincaré inequality. By Aida and Shigekawa perturbation theorem [36, Theorem 2.7], the uniform Poincaré inequality for $\mu_n^{i|-i}$ follows.

**Conclusion.** By [35, Theorem 1], we have the desired uniform log-Sobolev inequality for $\mu_n$. □

**Lemma 6.1.** *Let $\mu\in\mathscr{P}(\mathbb{R}^d)$. Let $f:\mathbb{R}^d\to\mathbb{R}$ be a $\mathscr{C}^1$ function with bounded derivative. Let $\varphi$ be a unique solution the Poisson equation $f-\langle\Phi(\mu),f\rangle=\mathscr{L}_\mu\varphi$ where $\mathscr{L}_\mu$ is defined by* (6.18)*. Then,*

$$\|\nabla\varphi\|_\infty \le [f]_1\underline{\lambda}^{-1}$$

*where $[f]_1$ denotes the Lipschitz constant of $f$ and $\underline{\lambda}^{-1}$ is defined by* (4.6)*.*

*Proof.* It is well-known that $\varphi(x)=\int_0^{+\infty}\mathrm{P}_t^\mu f(x)-\langle\Phi(\mu),f\rangle dt$ where $(\mathrm{P}_t^\mu)_{t\ge0}$ denotes the semi-group associated with $\mathscr{L}_\mu$ so that under adequate derivation conditions (which will be satisfied in our setting),

$$\nabla\varphi(x)=\int_0^{+\infty}\nabla\mathrm{P}_t^\mu f(x)dt.$$

Now,

$$\nabla\mathrm{P}_t^\mu f(x)=\mathbb{E}[\nabla f(\mathrm{X}_t^{\mu,x})\partial_x\mathrm{X}_t^{\mu,x}]$$

where $(\mathrm{X}_t^{\mu,x})$ denotes the solution starting from $x$ of the SDE associated with $\mathscr{L}_\mu$ and $(\partial_x\mathrm{X}_t^\mu)$ its first variation process solution to

$$d\mathrm{Y}_t=-\nabla^2\left(\frac{\delta\mathrm{F}}{\delta m}(\mu,\mathrm{X}_t^{\mu,x})+\mathrm{V}(\mathrm{X}_t^{\mu,x})\right)\mathrm{Y}_t dt$$

with $\mathrm{Y}_0=\mathrm{I}_d$. Under the assumption (4.6), one easily deduces from a Gronwall argument that for any $z\in\mathbb{R}^d$,

$$|\mathrm{Y}_t z|^2\le e^{-2\underline{\lambda}t}|z|^2$$

which implies that

$$|\nabla\mathrm{P}_t^\mu f(x)|\le[f]_1\|\partial_x\mathrm{X}_t^{\mu,x}\|_{\mathbf{op}}\le[f]_1e^{-\underline{\lambda}t}.$$

The result follows. □

*Remark* 6.1. In the proof of the Lemma 6.1, we can replace $\underline{\lambda}$ by

$$\inf_{x\in\mathbb{R}^d}\Big(\underline{\lambda}_{\nabla^2 V(x)} + \inf_{\mu\in\mathscr{P}(\mathbb{R}^d)}\underline{\lambda}_{\nabla^2\frac{\delta F}{\delta m}(\mu,x)}\Big) \quad (\geq \underline{\lambda}). \tag{6.38}$$

*Proof of Proposition 4.4.* As G, $\nabla$G and $\nabla^2$G are bounded, given that the interaction potentials are combinations of tensor products of G, it is easy to check (**H1**), (**H3**), **VFP1**, for all $i = 1,2$, $\gamma_i < +\infty$ and $\mathscr{P}_2(\mathbb{R}^d) \subset \mathscr{P}_G(\mathbb{R}^d) = \mathscr{P}(\mathbb{R}^d)$. As for (**H4**) and (**H5**), they are significantly more difficult to establish. Let's go prove them.

For (**H4**), we will proceed as in the proof of Proposition 4.2: we will establish assumptions of [35, Theorem 1].

**Step 1: Regularity properties.** For any $\mu\in\mathscr{P}(\mathbb{R}^d)$, consider the standard Langevin process $(X_t^\mu)_{t\geq 0}$ of Hamiltonian $\frac{\delta H}{\delta m}(\mu,\cdot) = V + Q'(\int G d\mu)G$, i.e.

$$dX_t^\mu = \sqrt{2}dB_t^\mu - \nabla\frac{\delta H}{\delta m}(\mu, X_t^\mu)dt. \tag{6.39}$$

This Langevin process admits for invariant distribution $\Phi(\mu)(dx) = \frac{1}{Z_\mu}e^{-\frac{\delta H}{\delta m}(\mu,x)}dx$. The functional H admits first and second-order flat derivatives that are jointly continuous and are $\mathscr{C}^2$ in the spatial variables. $H_n$ the Hamiltonian of the particle system and its derivatives are related to the macroscopic Hamiltonian H and its derivatives by

$$\begin{aligned} H_n(\mathbf{x}) &= nH(\mu_{\mathbf{x}}) \\ &= nH\Big(\frac{1}{n}\sum_{p=1}^n \delta_{x_p}\Big) \\ &= \sum_{p=1}^n V(x_p) + nQ\Big(\frac{1}{n}\sum_{p=1}^n G(x_p)\Big); \end{aligned} \tag{6.40}$$

$$\nabla_{x_i}H_n(\mathbf{x}) = \mathscr{D}_m H(\mu_{\mathbf{x}}, x_i); \tag{6.41}$$

$$\nabla^2_{x_j x_i}H_n(\mathbf{x}) = \nabla\mathscr{D}_m H(\mu_{\mathbf{x}}, x_i)\delta_{ij} + \frac{1}{n}\mathscr{D}_m^2 H(\mu_{\mathbf{x}}, x_i, x_j). \tag{6.42}$$

As

$$\frac{\delta^2 H}{\delta m^2}(\mu, x_1, x_2) = Q''\Big(\int G d\mu\Big)G(x_1)G(x_2), \tag{6.43}$$

and G, $\nabla$G and $\nabla^2$G are bounded, the functions $(\mu,x_1,x_2)\mapsto \nabla_{x_1}\frac{\delta^2 H}{\delta m^2}(\mu,x_1,x_2)$ and $(\mu,x_1,x_2)\mapsto\nabla^2_{x_1x_2}\frac{\delta^2 H}{\delta m^2}(\mu,x_1,x_2)$ are bounded. Then, to apply [35, Theorem 1], the boundedness assumption on second order flat derivative and the convexity in flat interpolation sense assumption ($Q''\geq 0$) on $\mathscr{E}_W$ hold true. It remains to establish the assumption 3 concerning uniform logarithmic Sobolev inequality for $\Phi(\mu)$ and the assumption 4 concerning uniform Poincaré inequality for the i-th conditional marginal of the invariant distribution of the particle system given by $\mu_n(d\mathbf{x}) = \frac{1}{Z_n}e^{-H_n(\mathbf{x})}d\mathbf{x}$.

**Step 2: Uniform log-Sobolev inequality.** For this step, one uses [27, 28] which gives some sufficient conditions on a potential U to ensure a $\rho_U$-log-Sobolev inequality for $\alpha_U(dx) := \frac{e^{-U(x)}}{Z_U}dx$. More precisely, such a result holds if U is $\mathscr{C}^2$ with lower bounded Hessian and satisfies the following Lyapunov condition: there are two positive constants $c_1^U$ and $c_2^U$ such that

$$\forall x\in\mathbb{R}^d,\quad x\cdot\nabla U(x) \geq c_1^U|x|^2 - c_2^U. \tag{6.44}$$

Thus, in our setting, the uniform log-Sobolev inequality will hold if these assumptions are true uniformly in $\mu$ for $x\mapsto\frac{\delta H}{\delta m}(\mu,x)$. Let us check these conditions: since $\gamma_1 < +\infty$ ($|\gamma_3|\leq\gamma_1$), for all $\mu\in\mathscr{P}_2(\mathbb{R}^d)$ and $x\in\mathbb{R}^d$, we have

$$\begin{aligned} \nabla^2\frac{\delta H}{\delta m}(\mu,\cdot) &\geq \nabla^2 V + \gamma_3\nabla^2 G \\ &\geq \nabla^2 V - \gamma_1\nabla^2 G; \end{aligned} \tag{6.45}$$

$$\begin{aligned} x\cdot\nabla\frac{\delta H}{\delta m}(\mu,x) &\geq c_1|x|^2 - c_2 + \gamma_3 x\cdot\nabla G(x) \\ &\geq c_1|x|^2 - c_2 - \gamma_1||\nabla G||_\infty|x|. \end{aligned} \tag{6.46}$$

As $|x| \le \varepsilon |x|^2 + C_\varepsilon$ for any $\varepsilon > 0$, there are two positive constants $c_1^\star$ and $c_2^\star$ such that

$$\forall (x,\mu) \in \mathbb{R}^d \times \mathscr{P}_2(\mathbb{R}^d), \quad x \cdot \nabla \frac{\delta H}{\delta m}(\mu, x) \ge c_1^\star |x|^2 - c_2^\star. \tag{6.47}$$

From what precedes, one can conclude $\Phi(\mu)$ satisfies a $\rho$ logarithmic Sobolev inequality with constant $\rho$ independent of $\mu$.

**Step 3: Uniform Poincaré inequality for $\mu_n^{i|-i}$.** First,

$$\frac{\mu_n^{i|-i}(\cdot|\mathbf{x}^{-i})}{dx_i} = \frac{e^{-nH(\mu_\mathbf{x})}}{\int e^{-nH(\mu_\mathbf{x})} dx_i} \tag{6.48}$$

which involves that its log-density satisfies

$$-\nabla_i \ln \mu_n^{i|-i}(x_i|\mathbf{x}^{-i}) = \mathscr{D}_m H(\mu_\mathbf{x}, x_i). \tag{6.49}$$

Second, let us denote $\mu^\Phi_{\mathbf{x}^{-i}}(\cdot)$ the probability measure defined by

$$\frac{\mu^\Phi_{\mathbf{x}^{-i}}(\cdot)}{dx_i} := \frac{\Phi(\mu_{\mathbf{x}^{-i}})}{dx_i} = \frac{e^{-\frac{\delta H}{\delta m}(\mu_{\mathbf{x}^{-i}}, x_i)}}{\int e^{-\frac{\delta H}{\delta m}(\mu_{\mathbf{x}^{-i}}, x_i)} dx_i}. \tag{6.50}$$

Its log-density satisfies

$$-\nabla_i \ln \mu^\Phi_{\mathbf{x}^{-i}}(x_i) = \mathscr{D}_m H(\mu_{\mathbf{x}^{-i}}, x_i). \tag{6.51}$$

We have (definition of the flat derivative)

$$\mathscr{D}_m H(\mu_1, x) - \mathscr{D}_m H(\mu_0, x) = \int_0^1 \int_{\mathbb{R}^d} \nabla_x \frac{\delta^2 H}{\delta m^2}(t\mu_1 + (1-t)\mu_0, x, y)(\mu_1 - \mu_0)(dy)dt. \tag{6.52}$$

By boundedness property of $(\mu, x, y) \mapsto \nabla_x \frac{\delta^2 H}{\delta m^2}(\mu, x, y)$, we have

$$|\mathscr{D}_m H(\mu_1, x) - \mathscr{D}_m H(\mu_0, x)| \le ||\nabla_x \frac{\delta^2 H}{\delta m^2}||_\infty ||\mu_1 - \mu_0||_{TV}. \tag{6.53}$$

Then, for some constant M, H verifies

$$|\mathscr{D}_m H(\mu_\mathbf{x}, x_i) - \mathscr{D}_m H(\mu_{\mathbf{x}^{-i}}, x_i)| \le \frac{M}{n}; \tag{6.54}$$

then the derivatives can at most differ by $\frac{M}{n}$. This means that, for $n$ large, the conditional measure $\mu_n^{i|-i}(\cdot|\mathbf{x}^{-i})$ is a *weak log-Lipschitz perturbation* of the measure $\mu^\Phi_{\mathbf{x}^{-i}}$. Furthermore, $\mu^\Phi_{\mathbf{x}^{-i}}$ satisfies a $\rho$-log-Sobolev, a fortiori, it satisfies a $\rho$-Poincaré inequality. By Aida and Shigekawa perturbation theorem [36, Theorem 2.7], the uniform Poincaré inequality for $\mu_n^{i|-i}$ follows.

**Conclusion.** By [35, Theorem 1], we have the desired uniform log-Sobolev inequality for $\mu_n$.

For (**H5**), we follow the same strategy as in the proof of Proposition 4.1 (see (6.17)): for $\nu_0, \nu_1 \in \mathscr{P}_2(\mathbb{R}^d)$ and a 1-Lipschitz function $f : \mathbb{R}^d \to \mathbb{R}$, we write $\nu_t = (1-t)\nu_0 + t\nu_1$. First,

$$\langle \Phi(\nu_t), f \rangle = \frac{1}{Z_{\nu_t}} \int_{\mathbb{R}^d} f(x) e^{-V(x) - Q'(\int G(y)\nu_t(dy))G(x)} dx \quad \text{and} \quad \frac{d}{dt}\langle \Phi(\nu_t), f \rangle = \mathrm{Cov}_{\Phi(\nu_t)}(f, g_t) \tag{6.55}$$

with

$$g_t(x) = -\partial_t \Big( Q'\Big(\int G(y)\nu_t(dy)\Big) G(x)\Big) = -G(x) Q''\left(\int G(y)\nu_t(dy)\right) \int G(y)(\nu_1 - \nu_0)(dy).$$

Following carefully the proof of Proposition 4.1, we get

$$|\mathrm{Cov}_{\Phi(\nu_t)}(f, g_t)| \le \|\nabla \varphi_t\|_{L^2(\Phi(\nu_t))} \|\nabla g_t\|_{L^2(\Phi(\nu_t))} \le \frac{1}{\underline{\lambda}^\star} \|\nabla g_t\|_{L^2(\Phi(\nu_t))} \tag{6.56}$$

with $\underline{\lambda}^\star = \inf_{x\in\mathbb{R}^d}\left(\underline{\lambda}_{\nabla^2 V(x)} + \gamma_3\underline{\lambda}_{\nabla^2 G(x)}\right) > 0$ under the assumptions of Proposition 4.4. Now, since

$$\nabla g_t = -Q''\Big(\int G(y)\nu_t(dy)\Big)\int G(y)(\nu_1-\nu_0)(dy)\nabla G,$$

since $\gamma_2 < +\infty$ and $\nabla G$ is bounded, by the Kantorovich-Rubinstein duality relation, we have

$$\|\nabla g_t\|_{L^2(\Phi(\nu_t))} \le \|\nabla g_t\|_\infty = \left|Q''\Big(\int G(y)\nu_t(dy)\Big)\right|\cdot\left|\int G(y)(\nu_1-\nu_0)(dy)\right| ||\nabla G||_\infty \le \gamma_2||\nabla G||_\infty^2 \mathscr{W}_1(\nu_0,\nu_1)$$

and by (6.56), we get for any smooth 1-Lipschitz function

$$|\langle\Phi(\nu_1),f\rangle - \langle\Phi(\nu_0),f\rangle| \le \int_0^1 |\frac{d}{dt}\langle\Phi(\nu_t),f\rangle| dt \le \frac{\gamma_2||\nabla G||_\infty^2}{\underline{\lambda}^\star}\mathscr{W}_1(\nu_0,\nu_1).$$

As $\frac{\gamma_2}{\underline{\lambda}^\star}||\nabla G||_\infty^2 < 1$, by a density argument and the Kantorovich-Rubinstein duality relation, it follows that $\Phi$ is a contraction on $(\mathscr{P}_2(\mathbb{R}^d),\mathscr{W}_1)$.

□

# A Appendix and additional proofs

## A.1 Propagation of chaos

**Theorem A.1** (Moment control). *Let $m \in \mathbb{R}_{\ge 1}$, $||\cdot||_2$ be the standard Euclidean norm on $\mathbb{R}^d$ and the continuous function*

$$\Theta_t: \quad \mathbb{R} \longrightarrow [0,+\infty[ \qquad \text{(A.1)}$$
$$\theta \longmapsto \begin{cases} \frac{1-e^{-2\theta t}}{\theta} & \text{if } \theta \neq 0;\\ 2t & \text{if } \theta = 0.\end{cases}$$

*Suppose that there exist pairs of constants $(\theta,\vartheta),(\theta_2,\vartheta_2),,\dots,(\theta_N,\vartheta_N)$ such that for all $(\mu,x)\in\mathscr{P}(\mathbb{R}^d)\times\mathbb{R}^d$,*

$$x\cdot\nabla V(x) \ge \theta||x||^2 + \vartheta; \qquad \text{(A.2)}$$
$$\forall k\in\{2,\dots,N\}, \quad x\cdot\nabla_{x_1}W^{(k)} * \mu^{\otimes k-1}(x) \ge \theta_k||x||^2 + \vartheta_k. \qquad \text{(A.3)}$$

*Let $\overline{\omega} := \theta + \sum_{k=2}^N k\theta_k$ and $\overline{\vartheta} := \vartheta + \sum_{k=2}^N k\vartheta_k$. We have*

$$\forall\mu_0\in\mathscr{P}_{2m}(\mathbb{R}^d) \quad \forall t\ge 0, \quad \Big|\Big|||\cdot||_2^2\Big|\Big|_{L^m(\mu_t)} \le \Big|\Big|||\cdot||_2^2\Big|\Big|_{L^m(\mu_0)} e^{-2\overline{\omega}t} + \frac{2m+d-2-\overline{\vartheta}+|2m+d-2-\overline{\vartheta}|}{2}\Theta_t(\overline{\omega}). \qquad \text{(A.4)}$$

*In particular, if $\overline{\omega}>0$ or $\overline{\omega}=0$ and $2m+d-2-\overline{\vartheta}\le 0$, we have uniformity in time:*

$$\sup_{t\ge 0}\Big|\Big|||\cdot||_2^2\Big|\Big|_{L^m(\mu_t)} \le \Big|\Big|||\cdot||_2^2\Big|\Big|_{L^m(\mu_0)} + \begin{cases}\frac{2m+d-2-\overline{\vartheta}+|2m+d-2-\overline{\vartheta}|}{2\overline{\omega}} & \text{if } \overline{\omega}>0;\\ 0 & \text{if } \overline{\omega}=0 \text{ and } 2m+d-2-\overline{\vartheta}\le 0.\end{cases} \qquad \text{(A.5)}$$

*Proof of Theorem A.1.* By the transfer formula, for any function $\psi\in L^1(\mu_t)$ or of constant sign, we have

$$\mathbb{E}[\psi(X_t)] = \int_{\mathbb{R}^d}\psi(x)\mu_t(dx). \qquad \text{(A.6)}$$

For all $m\ge 1$, by Itô's formula, we have

$$\frac{d}{dt}\mathbb{E}[||X_t||^{2m}] = \mathbb{E}[\mathscr{L}_{\mu_t}||\cdot||_2^{2m}(X_t)]. \qquad \text{(A.7)}$$

On the other hand, we have

$$\nabla||\cdot||_2^{2m}(x) = 2m||x||^{2(m-1)}x; \tag{A.8}$$

$$\begin{aligned}\Delta||\cdot||_2^{2m}(x) &= \nabla\cdot\nabla||\cdot||_2^{2m}(x) \\ &= 2m\sum_{j=1}^{d}\partial_{x_j}\left(|x||^{2(m-1)}x_j\right) \\ &= 2m\sum_{j=1}^{d}\left(||x||^{2(m-1)}+2(m-1)||x||^{2(m-2))}x_j^2\right) \\ &= 2md||x||^{2(m-1)}+4m(m-1)||x||^{2(m-2)}||x||^2 \\ &= 2m(2m+d-2)||x||^{2(m-1)}.\end{aligned} \tag{A.9}$$

We deduce that

$$\begin{aligned}\mathscr{L}_\mu||\cdot||_2^{2m}(x) &= 2m(2m+d-2)||x||^{2(m-1)} - 2m||x||^{2(m-1)}x\cdot\left(\nabla\frac{\delta F}{\delta m}(\mu,x)+\nabla V\right) \\ &= 2m||x||^{2(m-1)}\Big(2m+d-2-x\cdot\nabla V(x)-\sum_{k=2}^{N}kx\cdot\nabla_{x_1}W^{(k)}*\mu^{\otimes k-1}(x)\Big).\end{aligned} \tag{A.10}$$

By setting $S_m(t) := \mathbb{E}[||X_t||^{2m}]$, we have

$$S_m'(t) \le 2m\Big(2m+d-2-\vartheta-\sum_{k=2}^{N}k\vartheta_k\Big)S_{m-1}(t) - 2m\Big(\theta+\sum_{k=2}^{N}k\theta_k\Big)S_m(t). \tag{A.11}$$

As for a finite measure $\mu$, we have

$$\forall p\le q \quad \forall f\in L^q(\mu), \quad ||\psi||_{L^p(\mu)} := \left(\int|\psi|^p d\mu\right)^{\frac{1}{p}} \le \left(\int 1 d\mu\right)^{\frac{1}{p}-\frac{1}{q}}||\psi||_{L^q(\mu)}, \tag{A.12}$$

we have

$$S_{m-1}(t) \le S_m(t)^{\frac{m-1}{m}}. \tag{A.13}$$

**Case 1: Using $L^p$ injection.** If $\mathbf{2m+d-2-\vartheta-\sum_{k=2}^{N}k\vartheta_k>0}$, we have

$$S_m'(t) \le 2m\Big(2m+d-2-\vartheta-\sum_{k=2}^{N}k\vartheta_k\Big)S_m(t)^{\frac{m-1}{m}} - 2m\Big(\theta+\sum_{k=2}^{N}k\theta_k\Big)S_m(t). \tag{A.14}$$

We deduce that

$$m\Big(S_m(t)^{\frac{1}{m}}\Big)' \le 2m\Big(2m+d-2-\vartheta-\sum_{k=2}^{N}k\vartheta_k\Big) - 2m\Big(\theta+\sum_{k=2}^{N}k\theta_k\Big)S_m(t)^{\frac{1}{m}}. \tag{A.15}$$

By setting $z_m(t) := S_m(t)^{\frac{1}{m}}$, we have

$$z_m'(t) \le 2\Big(2m+d-2-\vartheta-\sum_{k=2}^{N}k\vartheta_k\Big) - 2\Big(\theta+\sum_{k=2}^{N}k\theta_k\Big)z_m(t). \tag{A.16}$$

And by Gronwall's lemma, if $\mathbf{\theta+\sum_{k=2}^{N}k\theta_k\neq 0}$, we have

$$z_m(t) - \frac{2m+d-2-\vartheta-\sum_{k=2}^{N}k\vartheta_k}{\theta+\sum_{k=2}^{N}k\theta_k} \le \left(z_m(0) - \frac{2m+d-2-\vartheta-\sum_{k=2}^{N}k\vartheta_k)}{\theta+\sum_{k=2}^{N}k\theta_k}\right)e^{-2t\left(\theta+\sum_{k=2}^{N}k\theta_k\right)} \tag{A.17}$$

If $\mathbf{\theta+\sum_{k=2}^{N}k\theta_k=0}$, by integration, we deduce that

$$z_m(t) \le z_m(0) + 2\Big(2m+d-2-\vartheta-\sum_{k=2}^{N}k\vartheta_k\Big)t \tag{A.18}$$

In conclusion of this disjunction of cases, whether $\theta+\sum_{k=2}^{N}k\theta_k=0$ is zero or not, we have

$$z_m(t)\le z_m(0)e^{-2t\left(\theta+\sum_{k=2}^{N}k\theta_k\right)}+\left(2m+d-2-\vartheta-\sum_{k=2}^{N}k\vartheta_k\right)\Theta_t\left(\theta+\sum_{k=2}^{N}k\theta_k\right); \tag{A.19}$$

$$\Theta_t:\quad \mathbb{R}\longrightarrow[0,+\infty[$$
$$\theta\longmapsto\begin{cases}\frac{1-e^{-2\theta t}}{\theta} & \text{if}\quad \theta\neq 0;\\ 2t & \text{if}\quad \theta=0.\end{cases}$$

**Case 2: Direct increase.** If $\mathbf{2m+d-2-\vartheta-\sum_{k=2}^{N}k\vartheta_k\le 0}$, we have

$$S_m'(t)\le -2m\left(\theta+\sum_{k=2}^{N}k\theta_k\right)S_m(t). \tag{A.20}$$

And by Gronwall's lemma, we have

$$S_m(t)\le S_m(0)e^{-2tm\left(\theta+\sum_{k=2}^{N}k\theta_k\right)}. \tag{A.21}$$

In conclusion, in all these cases, if $\mu_0\in\mathscr{P}_{2m}(\mathbb{R}^d)$, we have

$$\forall T\ge 0,\quad \sup_{0\le t\le T}\mathbb{E}[||X_t||^{2m}]\le M(m,\mu_0,T). \tag{A.22}$$

As for uniformity in time, it is ensured if one of the following conditions is verified

- ▷ $\theta+\sum_{k=2}^{N}k\theta_k>0$;
- ▷ $\theta+\sum_{k=2}^{N}k\theta_k=0$ and $2m+d-2-\vartheta-\sum_{k=2}^{N}k\vartheta_k\le 0$.

□

*Remark* A.1. Suppose that there exist pairs of constants $(\theta,\vartheta),(\theta_2,\vartheta_2),,\dots,(\theta_N,\vartheta_N)$ such that for all $(\mu,x)\in\mathscr{P}(\mathbb{R}^d)\times\mathbb{R}^d$,

$$x\cdot\nabla V(x)\ge\theta||x||^2+\vartheta; \tag{A.23}$$

$$\forall k\in\{2,\dots,N\},\quad x\cdot\nabla_{x_1}W^{(k)}*\mu^{\otimes k-1}(x)\ge\theta_k||x||^2+\vartheta_k. \tag{A.24}$$

If we add $\nabla V(0)=0$ and $\nabla_{x_1}W^{(k)}(0,\cdot)=0$ to the reduced Hessian conditions, we obtain the following hypotheses above. Indeed, if there exists $\beta\in\mathbb{R}$ such that $\nabla^2V\ge-\beta\mathbb{I}_d$ and for all $k\in\{2,\dots,N\}$, there exists $\beta_k\in\mathbb{R}$ such that $\nabla_{x_1}^2W^{(k)}\ge-\beta_k\mathbb{I}_d$, we have in particular

$$\langle\nabla V(x)-\nabla V(y),x-y\rangle\ge-\beta||x-y||^2; \tag{A.25}$$

$$\langle\nabla_{x_1}W^{(k)}(x,\cdot)-\nabla_{x_1}W^{(k)}(y,\cdot),x-y\rangle\ge-\beta_k||x-y||^2. \tag{A.26}$$

In this case, $\theta=-\beta$, $\vartheta=0$, for all $k\in\{2,\dots,N\}$, $\theta_k=-\beta_k$ and $\vartheta_k=0$.

*Proof of Theorem 5.1.* To show the result and for greater clarity of proof, we proceed in five steps described below.

**Step 1: Itô's formula and drift division.** By setting $G=\frac{2V}{(N-1)(N+2)}$, we have

$$dX_t^{(n),p}=\sqrt{2}dB_t^{(n),p}-\sum_{k=2}^{N}\frac{k}{\binom{n-1}{k-1}}\sum_{\substack{1\le i_1<i_2<\dots<i_{k-1}\le n\\ \forall j,\quad i_j\neq p}}\left(\nabla G(X_t^{(n),p})+\nabla_{x_1}W^{(k)}(X_t^{(n),p},X_t^{(n),i_1},\dots,X_t^{(n),i_{k-1}})\right)dt \tag{A.27}$$

$$=\sqrt{2}dB_t^{(n),p}-\sum_{k=2}^{N}\frac{k}{\binom{n-1}{k-1}}\sum_{\substack{1\le i_1<i_2<\dots<i_{k-1}\le n\\ \forall j,\quad i_j\neq p}}\int\left(\nabla G(X_t^{(n),p})+\nabla_{x_1}W^{(k)}(X_t^{(n),p},X_t^{(n),i_1},\dots,X_t^{(n),i_{k-1}})\right)d\mu^{\otimes(k-1)}dt.$$

$$dX_t^{(p)}=\sqrt{2}dB_t^{(n),p}-\sum_{k=2}^{N}k\int_{(\mathbb{R}^d)^{k-1}}\left(\nabla G(X_t^{(p)})+\nabla_{x_1}W^{(k)}(X_t^{(p)},y)\right)\mathbb{P}_{X_t}^{\otimes(k-1)}(dy)dt \tag{A.28}$$

$$=\sqrt{2}dB_t^{(n),p}-\sum_{k=2}^{N}\frac{k}{\binom{n-1}{k-1}}\sum_{\substack{1\le i_1<i_2<\dots<i_{k-1}\le n\\ \forall j,\quad i_j\neq p}}\int_{(\mathbb{R}^d)^{k-1}}\left(\nabla G(X_t^{(p)})+\nabla_{x_1}W^{(k)}(X_t^{(p)},y)\right)\mathbb{P}_{X_t}^{\otimes(k-1)}(dy)dt.$$

So

$$X_t^{(n),p} - X_t^{(p)} = X_r^{(n),p} - X_r^{(p)} + \int_r^t \sum_{k=2}^{N} \frac{k}{\binom{n-1}{k-1}} \sum_{\substack{1\le i_1<i_2<\ldots<i_{k-1}\le n \\ \forall j,\quad i_j\neq p}} \int_{(\mathbb{R}^d)^{k-1}} \Big(\nabla G(X_s^{(p)}) - \nabla G(X_s^{(n),p}) \tag{A.29}$$

$$+ \nabla_{x_1} W^{(k)}(X_s^{(p)}, y) - \nabla_{x_1} W^{(k)}(X_s^{(n),p}, X_s^{(n),i_1}, \ldots, X_s^{(n),i_{k-1}})\Big) \mathbb{P}_{X_s}^{\otimes(k-1)}(dy)ds.$$

By Itô's formula, we have

$$\sum_{p=1}^{n} ||X_t^{(n),p} - X_t^{(p)}||^2 = \sum_{p=1}^{n} ||X_r^{(n),p} - X_r^{(p)}||^2 \tag{A.30}$$

$$-2\sum_{p=1}^{n}\int_r^t \sum_{k=2}^{N} \frac{k}{\binom{n-1}{k-1}} \sum_{\substack{1\le i_1<i_2<\ldots<i_{k-1}\le n \\ \forall j,\quad i_j\neq p}} \int_{(\mathbb{R}^d)^{k-1}} \Big\langle \nabla G(X_s^{(n),p}) - \nabla G(X_s^{(p)}) - \nabla_{x_1} W^{(k)}(X_s^{(p)}, y)$$

$$+ \nabla_{x_1} W^{(k)}(X_s^{(n),p}, X_s^{(n),i_1}, \ldots, X_s^{(n),i_{k-1}}), X_s^{(n),p} - X_s^{(p)}\Big\rangle \mathbb{P}_{X_s}^{\otimes(k-1)}(dy)ds.$$

By setting $\mu_t = \mathbb{P}_{X_t}$ and $\nabla_{x_1} W^{(k)} * \mu_s^{\otimes(k-1)} := \int \nabla_{x_1} W^{(k)}(\cdot, y)\mu_s^{\otimes k-1}(dy)$, we have

$$\sum_{p=1}^{n} ||X_t^{(n),p} - X_t^{(p)}||^2 = \sum_{p=1}^{n} ||X_r^{(n),p} - X_r^{(p)}||^2 \tag{A.31}$$

$$-2\sum_{p=1}^{n}\int_r^t \Big\langle \nabla V(X_s^{(n),p}) - \nabla V(X_s^{(p)}), X_s^{(n),p} - X_s^{(p)}\Big\rangle ds$$

$$-2\sum_{p=1}^{n}\int_r^t \sum_{k=2}^{N} \frac{k}{\binom{n-1}{k-1}} \sum_{\substack{1\le i_1<i_2<\ldots<i_{k-1}\le n \\ \forall j,\quad i_j\neq p}} \Big\langle \nabla_{x_1} W^{(k)}(X_s^{(n),p}, X_s^{(n),i_1}, \ldots, X_s^{(n),i_{k-1}})$$

$$-\nabla_{x_1} W^{(k)} * \mu_s^{\otimes(k-1)}(X_s^{(p)}), X_s^{(n),p} - X_s^{(p)}\Big\rangle ds.$$

Let

$$\rho^{(1)}_{i_1,\ldots,i_{k-1}}(s) := \Big\langle \nabla_{x_1} W^{(k)}(X_s^{(n),p}, X_s^{(n),i_1}, \ldots, X_s^{(n),i_{k-1}}) - \nabla_{x_1} W^{(k)} * \mu_s^{\otimes(k-1)}(X_s^{(p)}), X_s^{(n),p} - X_s^{(p)}\Big\rangle \tag{A.32}$$

$$= \rho^{(2)}_{i_1,\ldots,i_{k-1}}(s) + \rho^{(3)}_{i_1,\ldots,i_{k-1}}(s);$$

$$\rho^{(2)}_{i_1,\ldots,i_{k-1}}(s) := \Big\langle \nabla_{x_1} W^{(k)}(X_s^{(n),p}, X_s^{(n),i_1}, \ldots, X_s^{(n),i_{k-1}}) - \nabla_{x_1} W^{(k)}(X_s^{(p)}, X_s^{(i_1)}, \ldots, X_s^{(i_{k-1})}), X_s^{(n),p} - X_s^{(p)}\Big\rangle; \tag{A.33}$$

$$\rho^{(3)}_{i_1,\ldots,i_{k-1}}(s) := \Big\langle \nabla_{x_1} W^{(k)}(X_s^{(p)}, X_s^{(i_1)}, \ldots, X_s^{(i_{k-1})}) - \nabla_{x_1} W^{(k)} * \mu_s^{\otimes(k-1)}(X_s^{(p)}), X_s^{(n),p} - X_s^{(p)}\Big\rangle. \tag{A.34}$$

**Step 2: Control of the confinement term.** As the Hessian matrix of the confinement potential is bounded from below, we have

$$-\sum_{p=1}^{n}\int_r^t \Big\langle \nabla V(X_s^{(n),p}) - \nabla V(X_s^{(p)}), X_s^{(n),p} - X_s^{(p)}\Big\rangle ds \le \beta \int_r^t \sum_{p=1}^{n} ||X_s^{(n),p} - X_s^{(p)}||^2 ds. \tag{A.35}$$

**Step 3: Control of the interaction term in $\rho^{(2)}$.** As the Hessian matrices of the interaction potentials in the first coordinate are uniformly bounded from below, we have

$$\forall k \in \{2,\ldots,N\}, \quad -\rho^{(2)}_{i_1,\ldots,i_{k-1}}(s) \le \beta_k ||X_s^{(n),p} - X_s^{(p)}||^2. \tag{A.36}$$

**Step 4: Control of the interaction term in $\rho^{(3)}$.** By Cauchy-Schwarz inequality, we have

$$-\mathbb{E}\sum_{\substack{1\le i_1<i_2<\ldots<i_{k-1}\le n\\ \forall j,\quad i_j\neq p}}\rho^{(3)}_{i_1,\ldots,i_{k-1}}(s)\le\sqrt{\zeta_{i_1,\ldots,i_{k-1}}(s)\mathbb{E}||X^{(n),p}_s-X^{(p)}_s||^2};\tag{A.37}$$

$$\zeta_{i_1,\ldots,i_{k-1}}(s):=\mathbb{E}\Bigg|\Bigg|\sum_{\substack{1\le i_1<i_2<\ldots<i_{k-1}\le n\\ \forall j,\quad i_j\neq p}}\nabla_{x_1}W^{(k)}(X^{(p)}_s,X^{(i_1)}_s,\ldots,X^{(i_{k-1})}_s)-\nabla_{x_1}W^{(k)}*\mu_s^{\otimes(k-1)}(X^{(p)}_s)\Bigg|\Bigg|^2\tag{A.38}$$

$$=\sum_{\substack{1\le i_1<i_2<\ldots<i_{k-1}\le n\\ \forall j,\quad i_j\neq p}}\mathbb{E}||\xi^{(p)}_{i_1,\ldots,i_{k-1}}(s)||^2+2\sum\mathbb{E}\langle\xi^{(p))}_{i_1,\ldots,i_{k-1}}(s),\xi^{(p)}_{j_1,\ldots,j_{k-1}}(s)\rangle;$$

$$\xi^{(p)}_{i_1,\ldots,i_{k-1}}(s):=\nabla_{x_1}W^{(k)}(X^{(p)}_s,X^{(i_1)}_s,\ldots,X^{(i_{k-1})}_s)-\nabla_{x_1}W^{(k)}*\mu_s^{\otimes(k-1)}(X^{(p)}_s).\tag{A.39}$$

As the $X^{(p)}_s$ are independent copies of $X_s$ with distribution $\mu_s$, we have

$$\mathbb{E}\langle\xi^{(p)}_{i_1,\ldots,i_{k-1}}(s),\xi^{(p)}_{j_1,\ldots,j_{k-1}}(s)\rangle=0.\tag{A.40}$$

Moreover, as the McKean-Vlasov flow admits bounded moments (see Theorem A.1), we have

$$\begin{aligned}\mathbb{E}||\xi^{(p)}_{i_1,\ldots,i_{k-1}}(s)||^2&\le\mathbb{E}||\nabla_{x_1}W^{(k)}(X^{(p)}_s,X^{(i_1)}_s,\ldots,X^{(i_{k-1})}_s)||^2 &\text{(A.41)}\\ &\le\Omega(\mathbb{E}||X^{(p)}_s||^{2m}+\mathbb{E}||X^{(i_1)}_s||^{2m}+\ldots+\mathbb{E}||X^{(i_{k-1})}_s||^{2m})\\ &=k\Omega\mathbb{E}[||X_s||^{2m}]\\ &\le k\Omega M_{2m}(T). &\text{(A.42)}\end{aligned}$$

It follows that

$$-\mathbb{E}\sum_{\substack{1\le i_1<i_2<\ldots<i_{k-1}\le n\\ \forall j,\quad i_j\neq p}}\rho^{(3)}_{i_1,\ldots,i_{k-1}}(s)\le\sqrt{\binom{n-1}{k-1}k\Omega M_{2m}(T)}\sqrt{\mathbb{E}||X^{(n),p}_s-X^{(p)}_s||^2}.\tag{A.43}$$

**Step 5: Gronwall's lemma.** For all $q\in\{1,\ldots,n\}$, let $\gamma_q(t)=\mathbb{E}||X^{(n),q}_t-X^{(q)}_t||^2$. By taking the expectation in equation A.31, by the previous controls and the exchangeability of the marginals of the particle system, we obtain

$$n\gamma_q(t)\le n\gamma_q(r)+2n\Big(\beta+\sum_{k=2}^{N}k\beta_k\Big)\int_r^t\gamma_q(s)ds+2n\sqrt{\Omega M_{2m}(T)}\Big(\sum_{k=2}^{N}\frac{k^{\frac{3}{2}}}{\sqrt{\binom{n-1}{k-1}}}\Big)\int_r^t\sqrt{\gamma_q(s)}ds.\tag{A.44}$$

As

$$\binom{n-1}{k-1}=\frac{k}{n}\binom{n}{k};\tag{A.45}$$

$$\Big(\frac{n}{k}\Big)^k\le\binom{n}{k}<e^k\Big(\frac{n}{k}\Big)^k;\tag{A.46}$$

we have

$$\sum_{k=2}^{N}\frac{k^{\frac{3}{2}}}{\sqrt{\binom{n-1}{k-1}}}\le\sum_{k=2}^{N}k^{\frac{3}{2}}\Big(\frac{k}{n}\Big)^{\frac{k-1}{2}}=\sum_{k=2}^{N}\frac{k^{\frac{k+2}{2}}}{n^{\frac{k-1}{2}}}\le(N-1)\frac{N^{\frac{N+2}{2}}}{\sqrt{n}}.\tag{A.47}$$

By setting $\kappa=(N-1)\sqrt{\Omega M_{2m}(T)N^{N+2}}$, we have

$$\gamma_q(t)\le\gamma_q(r)+2\Big(\beta+\sum_{k=2}^{N}k\beta_k\Big)\int_r^t\gamma_q(s)ds+2\frac{\kappa}{\sqrt{n}}\int_r^t\sqrt{\gamma_q(s)}ds.\tag{A.48}$$

Or equivalently

$$\frac{\gamma_q(t)-\gamma_q(r)}{t-r}\le2\Big(\beta+\sum_{k=2}^{N}k\beta_k\Big)\frac{\int_r^t\gamma_q(s)ds}{t-r}+2\frac{\kappa}{\sqrt{n}}\frac{\int_r^t\sqrt{\gamma_q(s)}ds}{t-r}.\tag{A.49}$$

And by passing to the limit, we obtain

$$\Upsilon_q'(t) \le 2\Big(\beta + \sum_{k=2}^{N} k\beta_k\Big)\Upsilon_q(t) + 2\frac{\kappa}{\sqrt{n}}\sqrt{\Upsilon_q(t)}. \tag{A.50}$$

If $\omega := \beta + \sum_{k=2}^{N} k\beta_k = 0$, we have ($\Upsilon_q(0) = 0$)

$$\Big(\sqrt{\Upsilon_q(t)}\Big)' = \frac{\Upsilon_q'(t)}{2\sqrt{\Upsilon_q(t)}} \le \frac{\kappa}{\sqrt{n}} \Longrightarrow \sqrt{\Upsilon_q(t)} \le \frac{\kappa}{\sqrt{n}}t. \tag{A.51}$$

By setting $y_q = \sqrt{\Upsilon_q}$, if $\omega \neq 0$, we have

$$\Bigg(y_q(t) + \frac{\kappa}{\Big(\beta+\sum_{k=2}^{N} k\beta_k\Big)\sqrt{n}}\Bigg)' = y_q'(t) \le \Big(\beta + \sum_{k=2}^{N} k\beta_k\Big)y_q(t) + \frac{\kappa}{\sqrt{n}} = \Big(\beta + \sum_{k=2}^{N} k\beta_k\Big)\Bigg(y_q(t) + \frac{\kappa}{\Big(\beta+\sum_{k=2}^{N} k\beta_k\Big)\sqrt{n}}\Bigg). \tag{A.52}$$

And by Gronwall's lemma, we deduce that ($\Upsilon_q(0) = 0$)

$$y_q(t) + \frac{\kappa}{\Big(\beta+\sum_{k=2}^{N} k\beta_k\Big)\sqrt{n}} \le \Bigg(y_q(0) + \frac{\kappa}{\Big(\beta+\sum_{k=2}^{N} k\beta_k\Big)\sqrt{n}}\Bigg)e^{t\left(\beta+\sum_{k=2}^{N} k\beta_k\right)}; \tag{A.53}$$

$$\sqrt{\Upsilon_q(t)} \le \frac{\kappa}{\Big(\beta+\sum_{k=2}^{N} k\beta_k\Big)\sqrt{n}}\Big(e^{t\left(\beta+\sum_{k=2}^{N} k\beta_k\right)} - 1\Big).$$

□

**Proposition A.1** (Corollary of the theorem 5.1)**.** *The result of the theorem 5.1 ensure with explicit rates,*

*(i) the weak convergence of the law $\mu_t^{(n),p}$ of a particle towards $\mu_t$: in fact, $X_t^{(n),p}$ has the law $\mu_t^{(n),p}$ by definition and $X_t^{(p)}$ has the law $\mu_t$ by construction, therefore for all $t \le T$ and all $n \ge 1$,*

$$\mathcal{W}_2^2(\mu_t^{(n),p}, \mu_t) \le \frac{M}{n} \tag{A.54}$$

*with M not depending on the number of particles. The uniformity in time of M is verified if $\omega < 0$ in theorem 5.1;*

*(ii) the propagation of chaos for the particle system: $q$ being a fixed integer- or more generally a $\mathbf{o}(n)$- and for the Wasserstein distance defined on $(\mathbb{R}^d)^q$, for $\{i_1, \dots, i_q\}$ a part of $\{1, \dots, n\}$, we have*

$$\mathcal{W}_2^2\Big(\mathbb{P}_{(X_t^{(n),i_1},\dots,X_t^{(n),i_q})}, \mu_t^{\otimes q}\Big) \le \frac{qM}{n}; \tag{A.55}$$

*(iii) the convergence of the empirical measurement of the particle system towards the McKean-Vlasov particle law: for any Lipschitzian function $\varphi : \mathbb{R}^d \to \mathbb{R}$, we have*

$$\mathbb{E}|\langle \Pi_t^{(n)}, \varphi\rangle - \langle \mu_t, \varphi\rangle|^2 \le \frac{M_{M,M_{2m},\varphi}}{n}. \tag{A.56}$$

*Proof of Proposition A.1.* By definition of the Wasserstein metric and empirical measurement, we have

(i)

$$\mathcal{W}_2^2(\mu_t^{(n),p}, \mu_t) \le \mathbb{E}||X_t^{(n),p} - X_t^{(p)}||^2 \le \frac{M}{n}; \tag{A.57}$$

(ii)

$$\mathcal{W}_2^2\Big(\mathbb{P}_{(X_t^{(n),i_1},\dots,X_t^{(n),i_q})}, \mu_t^{\otimes q}\Big) \le \mathbb{E}\Big|\Big|(X_t^{(n),i_1},\dots,X_t^{(n),i_q}) - (X^{(i_1)},\dots,X_t^{(i_q)})\Big|\Big|^2 \le q\mathbb{E}||X_t^{(n),p} - X_t^{(p)}||^2 \le \frac{qM}{n} = M\mathbf{o}(1); \tag{A.58}$$

(iii)

$$\mathbb{E}|\langle \Pi_t^{(n)}, \varphi\rangle - \langle \mu_t, \varphi\rangle|^2 \le \frac{2M + M_{2m}}{n}[\varphi]_1^2. \tag{A.59}$$

□

### A.2 McKean-Vlasov theory

**Theorem A.2** (Existence and uniqueness of solutions of Eq. (1.13))**.** *Let us assume that the functions* $b$ *and* $\sigma$ *are globally Lipschitz:* $\exists \mathrm{K} > 0\ \forall (x, y, \mu, \nu) \in \mathbb{R}^{\mathrm{D}} \times \mathbb{R}^{\mathrm{D}} \times \mathscr{P}_2(\mathbb{R}^{\mathrm{D}}) \times \mathscr{P}_2(\mathbb{R}^{\mathrm{D}})$,

$$||b(x,\mu) - b(y,\nu)|| + |||\sigma(x,\mu) - \sigma(y,\nu)||| \leq \mathrm{K}\Big(|x-y| + \mathscr{W}_2(\mu,\nu)\Big), \tag{A.60}$$

*where* $||\cdot||$ *denotes a vector norm,* $|||\cdot|||$ *is a matrix norm and* $\mathscr{W}_2$ *denotes the Wasserstein-2 distance. Assume that* $\mu_0 \in \mathscr{P}(\mathbb{R}^{\mathrm{D}})$*. Then for any* $\mathrm{T} \geq 0$ *the SDE Eq. (1.13) has a unique strong solution on* $[0, \mathrm{T}]$ *and consequently, its law is the unique weak solution to the Fokker-Planck equation Eq. (1.12) and the unique solution to the associated nonlinear martingale problem.*

The proof of this theorem is fairly classical. This proof is based on a *fixed point argument* that is sketched in [7, Proposition.1].

**Theorem A.3** (Polynomial Potential)**.** *Let* E *be a Polish measurable space. Let* $\alpha \in \mathscr{P}(\mathrm{E})$*. Let us consider a random vector* $\mathrm{X}^n$ *in* $\mathrm{E}^n$*, distributed according to the Gibbs measure:*

$$\mu_n(dx) := \frac{1}{\mathrm{Z}_n} e^{n\mathrm{F}(\mu_x)} \alpha^{\otimes n}(dx), \tag{A.61}$$

*where* $\mathrm{Z}_n$ *is a normalization constant and* F *is a polynomial function on* $\mathscr{P}(\mathrm{E})$ *(called the energy functional) of the form given by Eq. (1.17). Then (for some symmetric continuous bounded functions* $\mathrm{W}^{(k)}$*) the laws of* $\mu_{\mathrm{X}^n}$ *satisfy a large deviation principle in* $\mathscr{P}(\mathscr{P}(\mathrm{E}))$ *with speed* $\frac{1}{n}$ *and rate function*

$$\mu \longmapsto \mathbf{H}[\mu|\alpha] - \mathrm{F}(\mu) - \inf_{\eta \in \mathscr{P}(\mathrm{E})} \{\mathbf{H}[\eta|\alpha] - \mathrm{F}(\eta)\}. \tag{A.62}$$

### A.3 Gibbs-Laplace Variational Principle

**Definition A.1** (Distribution support)**.** Let $\mu$ be a probability measure on a Polish space E (or even a measure on a topological space!). We call support of $\mu$ noted $\mathbf{supp}(\mu)$ the closed set defined by

$$\bigcap_{\mathrm{F}\subset\mathrm{E}\text{ closed},\ \mu(\mathrm{F})=1} \mathrm{F} = \Big(\bigcup_{\mathrm{O}\subset\mathrm{E}\text{ open},\ \mu(\mathrm{O})=0} \mathrm{O}\Big)^{\mathrm{C}}. \tag{A.63}$$

In other words, the support of a distribution is the complement of the largest open set over which it is zero: the smallest closed set of maximum mass!

**Definition A.2** (Extremum essential)**.** Let $\mu$ be a probability measure on a Polish space E and $\mathrm{V} : \mathrm{E} \longrightarrow [-\infty, +\infty]$ measurable. We call infimum $\mu-$essential of V the quantity

$$\mu - \mathbf{essinf}\mathrm{V} := \inf\{v \in \mathbb{R}, \quad \mu(\{\mathrm{V} \leq v\}) > 0\} \tag{A.64}$$

**Theorem A.4** (Variational principle)**.** *For any probability measure* $\mu$ *on a topological space* $\Omega$ *and any measurable function* $\mathrm{V} : \Omega \longrightarrow \overline{\mathbb{R}}$*, we have*

$$\lim_{n\to+\infty} \frac{1}{n} \log \int e^{-n\mathrm{V}} d\mu = -\mu - \mathbf{essinf}\mathrm{V}. \tag{A.65}$$

*Moreover, if* V *is upper semicontinuous, then*

$$\inf_{\mathbf{supp}(\mu)} \mathrm{V} = \mu - \mathbf{essinf}\mathrm{V}. \tag{A.66}$$

*Proof Sketch:* Suppose $\mu - \mathbf{essinf}\mathrm{V}$ is finite. Check that we can assume without loss of generality that $\mathrm{V} \geq 0$ and $\mu - \mathbf{essinf}\mathrm{V} = 0$. Then check $\mathbb{I}_{\mathrm{V}\leq\varepsilon} e^{-n\varepsilon} \leq e^{-n\mathrm{V}} \leq 1$ and conclude. Show that the limit is $+\infty$ with the lower bound when $\mu - \mathbf{essinf}\mathrm{V} = -\infty$.

*Proof.* ▷ $\mu - \mathbf{essinf}\mathrm{V}$ is finished:

$$\frac{1}{n} \log \int e^{-n\mathrm{V}} d\mu + \mu - \mathbf{essinf}\mathrm{V} = \frac{1}{n} \log\Big(\int e^{-n(\mathrm{V}-\mu-\mathbf{essinf}\mathrm{V})} d\mu\Big). \tag{A.67}$$

This implies that we can assume without loss of generality that $\mathrm{V} \geq 0$ and $\mu - \mathbf{essinf}\mathrm{V} = 0$ because $\mathrm{V} - \mu - \mathbf{essinf}\mathrm{V} \geq 0$ almost surely, its essential infimum under $\mu$ is zero and the convergence that interests us is equivalent to

$$\frac{1}{n} \log\Big(\int e^{-n(\mathrm{V}-\mu-\mathbf{essinf}\mathrm{V})} d\mu\Big) \overset{n\to+\infty}{\longrightarrow} 0 \tag{A.68}$$

But for all $\varepsilon > 0 = \mu - \mathbf{essinf}V$,

$$\mathbb{1}_{V\leq\varepsilon} e^{-n\varepsilon} \leq e^{-nV} \leq 1 \Longleftrightarrow \frac{\log \mu(V \leq \varepsilon)}{n} - \varepsilon \leq \frac{1}{n}\log\int e^{-nV} d\mu \leq 0. \tag{A.69}$$

We deduce that by the bounding limit theorem, we have

$$\limsup_n \frac{1}{n}\log\int e^{-nV} d\mu = \liminf_n \frac{1}{n}\log\int e^{-nV} d\mu = 0. \tag{A.70}$$

▷ $\mu - \mathbf{essinf}V = -\infty$: In this case, for all $v \in \mathbb{R}$, we have $\mu(V \leq v) > 0$ and

$$\int_\Omega e^{-nV} d\mu \geq \int_{\{V\leq v\}} e^{-nV} d\mu \geq e^{-nv}\mu(\{V \leq v\}). \tag{A.71}$$

It is deduced that

$$\forall v \in \mathbb{R}, \quad \frac{1}{n}\log\int_\Omega e^{-nV} d\mu \geq -v + + \frac{\log \mu(V \leq v)}{n} \tag{A.72}$$

$$\Longrightarrow \lim_{n\to+\infty} \frac{1}{n}\log\int e^{-nV} d\mu = +\infty = -\mu - \mathbf{essinf}V.$$

□

**Theorem A.5** (Gibbs measures and deviations)**.** *Let* E *be a Polish space,* $\mu$ *a probability measure on* E *and* $V : E \longrightarrow \overline{\mathbb{R}}$ *a measurable function. We have:*

▷

$$\inf_{\mathbf{supp}(\mu)} V \leq \mu - \mathbf{essinf}V. \tag{A.73}$$

▷ *If* V *is upper semicontinuous, then*

$$\inf_{\mathbf{supp}(\mu)} V \geq \mu - \mathbf{essinf}V \Longrightarrow \inf_{\mathbf{supp}(\mu)} V = \mu - \mathbf{essinf}V. \tag{A.74}$$

*In particular, if* V *is continuous, then the principle of large deviations holds for*

$$\mu_n(dx) := \frac{1}{\int e^{-nV} d\mu} e^{-nV(x)}\mu(dx) \tag{A.75}$$

*with rate function* $I^V := V + I_0 - \inf\{V + I_0\}$ *with*

$$I_0(x) := \begin{cases} 0 & \text{if } x \in \mathbf{supp}(\mu), \\ +\infty & \text{else.} \end{cases} \tag{A.76}$$

*Proof sketch:*

▷ Show that

$$\left\{x, \quad V(x) < \inf_{\mathbf{supp}(\mu)} V\right\} \bigcap \mathbf{supp}(\mu) = \emptyset. \tag{A.77}$$

Then conclude.

▷ For all $\varepsilon > 0$, show that

$$\left\{x, \quad V(x) < \inf_{\mathbf{supp}(\mu)} V + \varepsilon\right\} \quad \text{is an open containing a support element:} \tag{A.78}$$

their intersection is non-empty; then conclude.

### A.4 Principle of contraction and tensorization

Let $f : X \longrightarrow G$ be continuous between two Polish spaces and $(X_N)$ a random variable sequence of X satisfying the principle of large deviations of rate function $I : X \longrightarrow [0, +\infty]$. Then $((f(X_N))$ satisfies the principle of large deviations of rate function $J : G \longrightarrow [0, +\infty]$ such that

$$J(g) := \inf_{f^{-1}(\{g\})} I. \tag{A.79}$$

Let $(X_n)_{n\geq 1}$ and $(Y_n)_{n\geq 1}$ be sequences with values respectively in $E_1$ and $E_2$, independent ($\mathbb{P}_{(X_n,Y_n)} = \mathbb{P}_{X_n} \bigotimes \mathbb{P}_{Y_n}$) and both satisfying the principle of large deviations of the respective good rate functions $I_1$ and $I_2$. Then $((X_n, Y_n))_{n\geq 1}$ satisfies the principle of large deviations on the product space and of good rate function I defined by

$$I(x, y) := I_1(x) + I_2(y). \tag{A.80}$$

### A.5 Entropy and Chaos

**Theorem A.6** (Characterization of relative entropy: Sanov's theorem)**.** *Let* $\mu$ *and* $\nu$ *be probability measures (even finite!) on a Polish space* $E$ *and* $(\varphi_j)_{j\in\mathbb{N}}$ *a dense sequence of functions bounded uniformly continuous. we have*

$$\lim_{k\to\infty}\lim_{\varepsilon\to 0}\lim_{n\to\infty}\frac{1}{n}\log\mu^{\otimes n}\left(\left\{y\in E^n;\quad \forall j\in\{1,\dots,k\},\quad \left|\int_E \varphi_j d\nu-\frac{1}{n}\sum_{i=1}^n\varphi_j(y_i)\right|\le\varepsilon\right\}\right)=-\mathbf{H}[\nu|\mu]. \tag{A.81}$$

*We interpret n as the number of particles; the* $\varphi_j$ *a sequence of observables whose mean value is measured; and* $\varepsilon$ *as the precision of the measurements. This formula concisely summarizes the essential information contained in the Boltzmann function* $\mathbf{H}$.

**Theorem A.7** ((strict) convexity of relative entropy)**.** *Let* $\mu\in\mathscr{P}(\Omega)$. $\mathbf{H}[\cdot|\mu]$ *has values in* $\overline{\mathbb{R}_+}$*, convex, strictly convex on* $\{\nu,\quad \mathbf{H}[\nu|\mu]<+\infty\}$ *and is zero only in* $\mu$.

**Theorem A.8** (Tensorization property)**.** *Let* $\mu\in\mathscr{P}(\Omega)$, $\nu\in\mathscr{P}(\Omega^n)$ *with* $\nu_i$ *its* $i-th$ *marginal. So*

$$\mathbf{H}[\nu|\mu^{\otimes n}]=\mathbf{H}[\nu|\bigotimes_{i=1}^n\nu_i]+\sum_{i=1}^n\mathbf{H}[\nu_i|\mu] \tag{A.82}$$

**Theorem A.9** (Villani)**.** *Let* $(X:=(X_1,\dots,X_n)$ *be a random variable on* $E^n$ *with* $E$ *a Polish space,* $\mu_n:=\mathbb{P}_X\in\mathscr{P}(E^n)$, $\delta_X:=\frac{1}{n}\sum\delta_{X_i}$ *and* $\mu\in\mathscr{P}(E)$. *The following assertions are equivalent:*

▷ $\delta_X$ *converges in law to* $\mu$*:*

$$\forall\varphi\in\mathscr{C}_b(E),\qquad \int\varphi d\delta_X\overset{n\to+\infty}{\longrightarrow}\int\varphi d\mu\quad\text{almost surely.} \tag{A.83}$$

▷

$$\forall\varphi\in \boldsymbol{Lip}_b(E),\qquad \lim_{n\to+\infty}\mathbb{E}_{\mu_n}\left[\left|\int\varphi d(\delta_X-\mu)\right|\right]=0. \tag{A.84}$$

Without repeating the proof, we can say that this result is obtained by defining a metric on $\mathscr{P}(E)$ from a dense sequence of Lipschitz functions and then by defining the transport distance Wasserstein's $\mathscr{W}_1$ on $\mathscr{P}(\mathscr{P}(E))$ associated with this metric. Using this result, we can more formally prove the propagation of chaos.

**Definition A.3** (U-statistics)**.** Let $E$ be a set, $k\in\mathbb{N}^*$ and $\Phi:E^k\longrightarrow\mathbb{R}$ a symmetric function. Then the application: $(n\ge k)$

$$X:=(x_j)_{j=1,\dots,n}\in E^n\longmapsto U(X):=\frac{k!(n-k)!}{n!}\sum_{1\le i_1<i_2<\dots<i_k\le n}\Phi(x_{i_1},\dots,x_{i_k}) \tag{A.85}$$

is called U−statistic of order $k$ and kernel $\Phi$. $U(X)$ is called U−statistic of order $k$ and kernel $\Phi$ associated with the sample X. This statistic corresponds to the arithmetic mean of the kernel over all the parts at $k$ elements of the set of sample values. we often write $U_n(\Phi)(X):=U(X)$. If E is a measurable space, we generalize this definition to the space of probabilities by the functional $\mu\longmapsto\mathbb{E}_{\mu^{\otimes k}}[\Phi]$.

### A.6 Proofs

*Proof of Proposition 5.1.* Let $\mathfrak{S}_n$ be the group of permutations of $\{1,\dots,n\}$ and $\mathfrak{B}_n$ the $\sigma$−algebra defined by

$$\mathfrak{B}_n:=\sigma\left\{B_n\times C_n|C_n\in\mathscr{B}(E^{[n+1,+\infty[}),\quad B_n\in\mathscr{B}(E^n),\quad\forall\tau\in\mathfrak{S}_n,\quad \tau\mathbb{1}_{B_n}=\mathbb{1}_{B_n}\right\}. \tag{A.86}$$

This $\sigma$−algebra is invariant under permutations and verifies for all $n\ge1$,

$$\mathfrak{B}_{n+1}\subset\mathfrak{B}_n. \tag{A.87}$$

By integrability,

$$\forall(i_1,\dots,i_k)\in I_n^k,\quad \mathbb{E}[\Phi(X_{i_1},\dots,X_{i_k})|\mathfrak{B}_n]=\mathbb{E}[\Phi(X_1,\dots,X_k)|\mathfrak{B}_n], \tag{A.88}$$

which implies that

$$U_n(\Phi)=\mathbb{E}[\Phi(X_1,\dots,X_k)|\mathfrak{B}_n]. \tag{A.89}$$

According to the limit theorems on martingales (closed martingale) and the law of $0-1$ applied to the asymptotic tribe $\mathfrak{B}_\infty:=\bigcap_{n\ge1}\mathfrak{B}_n$, we deduce that we almost surely have

$$U_n(\Phi)\overset{n\to+\infty}{\longrightarrow}\mathbb{E}[\Phi(X_1,\dots,X_k)|\mathfrak{B}_\infty]=\mathbb{E}[\Phi(X_1,\dots,X_k)]. \tag{A.90}$$

□

*Proof of Proposition 5.3.* We prove this result by induction. Indeed, for $k=1$ the inequality is verified since we have equality of the two members. Suppose that for $k-1$ the inequality holds. Denote by $B_k$ the left side of this inequality. We have

$$B_k = \log \mathbb{E}^{X^k}\Big[\mathbb{E}\Big[\exp\Big(\frac{1}{|I_n^{k-1}|}\sum_{(i_1,\dots,i_{k-1})\in I_n^{k-1}}\sum_{i_k\notin\{i_1,\dots,i_{k-1}\}}\frac{1}{n-k+1}\Phi_{i_1,\dots,i_k}(X_{i_1}^1,\dots,X_{i_k}^k)\Big)\Big|X^k\Big]\Big] \tag{A.91}$$

with $X^k := (X_1^k,\dots,X_n^k)$. Let us set

$$\tilde\Phi_{i_1,\dots,i_{k-1}} := \frac{1}{n-k+1}\sum_{i_k\notin\{i_1,\dots,i_{k-1}\}}\Phi_{i_1,\dots,i_k}(X_{i_1}^1,\dots,X_{i_k}^k). \tag{A.92}$$

By induction hypothesis, we deduce that

$$B_k \le \log \mathbb{E}^{X^k}\Big[\exp\Big(\frac{n-k+2}{|I_n^{k-1}|}\sum_{(i_1,\dots,i_{k-1})\in I_n^{k-1}}\log\mathbb{E}\Big[\exp\Big(\frac{1}{n-k+2}\tilde\Phi_{i_1,\dots,i_{k-1}}\Big)\Big|X^k\Big]\Big)\Big]. \tag{A.93}$$

Since

$$\log \mathbb{E}^{X^k}\Big[\exp\Big(\frac{n-k+2}{|I_n^{k-1}|}\sum_{(i_1,\dots,i_{k-1})\in I_n^{k-1}}\log\mathbb{E}\Big[\exp\Big(\frac{1}{n-k+2}\tilde\Phi_{i_1,\dots,i_{k-1}}\Big)\Big|X^k\Big]\Big)\Big] \tag{A.94}$$

is upper bounded by

$$\log \mathbb{E}^{X^k}\Big[\exp\Big(\frac{1}{|I_n^{k-1}|}\sum_{(i_1,\dots,i_{k-1})\in I_n^{k-1}}\log\Big(\mathbb{E}\Big[\exp\Big(\frac{1}{n-k+2}\tilde\Phi_{i_1,\dots,i_{k-1}}\Big)\Big|X^k\Big]\Big)^{n-k+2}\Big)\Big], \tag{A.95}$$

by convexity of $X\longmapsto \log\mathbb{E}[e^X]$ (consequence of Holder's inequality), we have

$$B_k \le \frac{1}{|I_n^{k-1}|}\sum_{(i_1,\dots,i_{k-1})\in I_n^{k-1}}\log\mathbb{E}^{X^k}\Big[\Big(\mathbb{E}\Big[\exp\Big(\frac{1}{n-k+2}\tilde\Phi_{i_1,\dots,i_{k-1}}\Big)\Big|X^k\Big]\Big)^{n-k+2}\Big]. \tag{A.96}$$

In this last inequality, for all $(i_1,\dots,i_{k-1})$, the logarithmic term verifies

$$\begin{aligned}&\mathbb{E}^{X^k}\Big[\Big(\mathbb{E}\Big[\exp\Big(\frac{1}{n-k+2}\tilde\Phi_{i_1,\dots,i_{k-1}}\Big)\Big|X^k\Big]\Big)^{n-k+2}\Big]\\ &=\mathbb{E}^{X^k}\Big[\Big(\mathbb{E}\Big[\exp\Big(\frac{1}{(n-k+2)(n-k+1)}\sum_{i_k\notin\{i_1,\dots,i_{k-1}\}}\Phi_i(X_{i_1}^1,\dots,X_{i_k}^k)\Big)\Big|X^k\Big]\Big)^{n-k+2}\Big],\end{aligned} \tag{A.97}$$

and by Holder's inequality, we have

$$\begin{aligned}&\mathbb{E}^{X^k}\Big[\Big(\mathbb{E}\Big[\exp\Big(\frac{1}{n-k+2}\tilde\Phi_{i_1,\dots,i_{k-1}}\Big)\Big|X^k\Big]\Big)^{n-k+2}\Big]\\ &\le\mathbb{E}^{X^k}\Big[\Big(\prod_{i_k\notin\{i_1,\dots,i_{k-1}\}}\mathbb{E}\Big[\exp\Big(\frac{1}{n-k+2}\Phi_{i_1,\dots,i_k}(X_{i_1}^1,\dots,X_{i_k}^k)\Big)\Big|X^k\Big]\Big)^{\frac{n-k+2}{n-k+1}}\Big].\end{aligned} \tag{A.98}$$

By Jensen's inequality, we also have the upper bound of the right-hand side of this last inequality by

$$\mathbb{E}^{X^k}\Big[\prod_{i_k\notin\{i_1,\dots,i_{k-1}\}}\mathbb{E}\Big[\exp\Big(\frac{1}{n-k+1}\Phi_{i_1,\dots,i_k}(X_{i_1}^1,\dots,X_{i_k}^k)\Big)\Big|X^k\Big]\Big],$$

and by independence, this quantity is equal to

$$\prod_{i_k\notin\{i_1,\dots,i_{k-1}\}}\mathbb{E}\Big[\exp\Big(\frac{1}{n-k+1}\Phi_{i_1,\dots,i_k}(X_{i_1}^1,\dots,X_{i_k}^k)\Big)\Big]. \tag{A.99}$$

□

*Proof of Proposition 5.4.* Let $((X_1^j,\dots,X_n^j))_{j=1,\dots,k}$ be independent copies of $(X_1,\dots,X_n)$. By the two propositions above, setting for all $i\in I_n^k$, $\Phi_{i_1,\dots,i_k}\equiv W^{(k)}$, we have for all $\lambda>0$

$$\Lambda_n(\lambda,W^{(k)}) = \frac{1}{n}\log\mathbb{E}\Big[\exp\Big(\frac{\lambda n}{|I_n^k|}\sum_{i\in I_n^k}W^{(k)}(X_{i_1},\dots,X_{i_k})\Big)\Big], \tag{A.100}$$

and it follows that

$$\Lambda_n(\lambda, W^{(k)}) \leq \frac{1}{n}\log\mathbb{E}\Big[\exp\Big(\frac{1}{|I_n^k|}\sum_{i\in I_n^k}\lambda n C_k |W^{(k)}|(X_{i_1}^1,\dots,X_{i_k}^k)\Big)\Big] \tag{A.101}$$

$$\leq \frac{1}{n|I_n^{k-1}|}\sum_{i\in I_n^k}\log\mathbb{E}\Big[\exp\Big(\frac{\lambda n C_k}{n-k+1}|W^{(k)}|(X_{i_1}^1,\dots,X_{i_k}^k)\Big)\Big].$$

Gold

$$\frac{1}{n|I_n^{k-1}|}\sum_{i\in I_n^k}\log\mathbb{E}\Big[\exp\Big(\frac{\lambda n C_k}{n-k+1}|W^{(k)}|(X_{i_1}^1,\dots,X_{i_k}^k)\Big)\Big] = \frac{n-k+1}{n}\log\mathbb{E}\Big[\exp\Big(\frac{\lambda n C_k}{n-k+1}|W^{(k)}|(X_{i_1},\dots,X_{i_k})\Big)\Big]. \tag{A.102}$$

It is deduced that

$$\Lambda_n(\lambda, W^{(k)}) \leq \frac{n-k+1}{n}\log\mathbb{E}\Big[\exp\Big(\frac{\lambda n C_k}{n-k+1}|W^{(k)}|(X_{i_1},\dots,X_{i_k})\Big)\Big] \tag{A.103}$$

$$\leq \frac{1}{k}\log\mathbb{E}\Big[\exp\Big(kC_k\lambda|W^{(k)}|(X_1,\dots,X_k)\Big)\Big],$$

and this last inequality is obtained by growth on $(0,+\infty)$ of $a \longmapsto \frac{1}{a}\log\mathbb{E}[e^{aX}]$ and from the fact that for all $n$ and $k$ such that $n \geq k$, we have $\frac{n}{n-k+1} \leq k$. □

*Proof of Proposition 5.5.* To do this, we will show that for any probability measure $\mu$ such that $\mathbf{H}[\mu|\alpha] < +\infty$ and for any $k$, $W^{(k)} \in L^1(\mu^{\otimes k})$, we have $l^*(\mathcal{O}) \geq -\mathbf{E}_W[\mu]$. Let $\mathbb{B}(\mu,\delta)$ be the open ball with center $\mu$ and of radius $\delta > 0$ in $\mathcal{M}_1(\mathbb{R}^d)$ endowed with the Lévy-Prokhorov metric $d_{LP}$ such that $\mathbb{B}(\mu,\delta) \subset \mathcal{O}$. Let us introduce the events

- ▷ $$A_n := \Big\{x \in (\mathbb{R}^d)^n \Big| \quad L_n := L_n(x,\cdot) \in \mathbb{B}(\mu,\delta)\Big\}; \tag{A.104}$$
- ▷ $$B_n := \Big\{x \in (\mathbb{R}^d)^n \Big| \quad \frac{1}{n}\sum_{i=1}^n \log\frac{d\mu}{d\alpha}(x_i) = \frac{1}{n}\log\Big(\frac{d\mu}{d\alpha}\Big)^{\otimes n}(x) \leq \mathbf{H}[\mu|\alpha] + \varepsilon\Big\}; \tag{A.105}$$
- ▷ $$C_n := \Big\{x \in (\mathbb{R}^d)^n \Big| \sum_{k=2}^N U_n(W^{(k)}) \leq \sum_{k=2}^N \mathbf{W}^{(k)}[\mu] + \varepsilon\Big\}. \tag{A.106}$$

We deduce that for all $\varepsilon > 0$, we have

$$\mu_n^*(L_n \in \mathbb{B}(\mu,\delta)) \geq \int_{A_n}\Big(\frac{d\mu^{\otimes n}}{d\mu_n^*}(x)\Big)^{-1}\mu^{\otimes n}(dx) = \int_{A_n} e^{-\sum_{i=1}^n \log\frac{d\mu}{d\alpha}(x_i)}e^{-n\sum_{k=2}^N U_n(W^{(k)})}\mu^{\otimes n}(dx) \tag{A.107}$$

and

$$\int_{A_n} e^{-\sum_{i=1}^n \log\frac{d\mu}{d\alpha}(x_i)}e^{-n\sum_{k=2}^N U_n(W^{(k)})}\mu^{\otimes n}(dx) \geq \mu^{\otimes n}(A_n\cap B_n\cap C_n)e^{-n(\mathbf{H}[\mu|\alpha]+\varepsilon)-\gamma} \quad \text{with } \gamma := n\Big(\sum_{k=2}^N \mathbf{W}^{(k)}[\mu]+\varepsilon\Big). \tag{A.108}$$

Thereby

$$\mu_n^*(L_n \in \mathbb{B}(\mu,\delta)) \geq \mu^{\otimes n}(A_n\cap B_n\cap C_n)e^{-n\mathbf{E}_W[\mu]-2n\varepsilon}. \tag{A.109}$$

We will prove that

$$\mu^{\otimes n}(A_n\cap B_n\cap C_n) \overset{n\to+\infty}{\longrightarrow} 1. \tag{A.110}$$

Indeed, by the law of large numbers, we have

$$\mu^{\otimes n}(A_n) \overset{n\to+\infty}{\longrightarrow} 1, \quad \mu^{\otimes n}(B_n) \overset{n\to+\infty}{\longrightarrow} 1. \tag{A.111}$$

Moreover, by the law of large numbers for U−statistics (Section 5.2), we also have

$$\mu^{\otimes n}(C_n) \overset{n\to+\infty}{\longrightarrow} 1. \tag{A.112}$$

It is deduced that

$$l^*(\mathcal{O}) \geq \liminf_{n\to+\infty}\frac{1}{n}\mu_n^*(L_n \in \mathbb{B}(\mu,\delta)) \geq -\mathbf{E}_W[\mu] - 2\varepsilon, \tag{A.113}$$

and we conclude by letting $\varepsilon$ tend to zero. □

*Proof of Proposition 5.6.* To do this, consider the truncation function

$$W^{(k),L} := \max(-L, \min(W^{(k)}, L)). \tag{A.114}$$

We have by Lebesgue's dominated convergence theorem

$$\log \mathbb{E}[\exp(m|W^{(k),L} - W^{(k)}|(X_1,\dots,X_k))] \overset{L\to+\infty}{\longrightarrow} 0. \tag{A.115}$$

So we can choose $L = L(m)$ so that

$$\log \mathbb{E}[\exp(m|W^{(k),L(m)} - W^{(k)}|(X_1,\dots,X_k))] \le \frac{1}{m}. \tag{A.116}$$

For $m \ge 1$ and $L(m) > 0$ fixed, we can find a sequence $(W_l^{(k),L})_{l\ge1}$ of continuous functions bounded such that

$$W_l^{(k),L}(X_1,\dots,X_k) \overset{l\to+\infty, L^1}{\longrightarrow} W^{(k),L}(X_1,\dots,X_k), \quad \forall l \ge 1, \quad |W_l^{(k),L}(X_1,\dots,X_k)| \le L, \tag{A.117}$$

otherwise, we consider the truncation $\max(-L, \min(W_l^{(k),L}, L))$. Seen that $\forall l \ge 1$,

$$\exp\Big(m\Big(|W^{(k)} - W^{(k),L}|(X_1,\dots,X_k) + |W^{(k)} - W_l^{(k),L}|(X_1,\dots,X_k)\Big)\Big) \le \exp\Big(m|W^{(k)} - W^{(k),L}|(X_1,\dots,X_k) + 2mL\Big), \tag{A.118}$$

by dominated convergence, we have

$$\mathbb{E}\Big[\exp\Big(m\Big(|W^{(k)} - W^{(k),L}|(X_1,\dots,X_k) + |W^{(k)} - W_l^{(k),L}|(X_1,\dots,X_k)\Big)\Big)\Big] \overset{l\to+\infty}{\longrightarrow} \mathbb{E}\Big[\exp\Big(m|W^{(k)} - W^{(k),L}|(X_1,\dots,X_k)\Big)\Big]. \tag{A.119}$$

For $L = L(m)$, we can choose $l = l(m)$ so that

$$\log \mathbb{E}\Big[\exp\Big(m\Big(|W^{(k)} - W^{(k),L}|(X_1,\dots,X_k) + |W^{(k)} - W_l^{(k),L}|(X_1,\dots,X_k)\Big)\Big)\Big] \le \frac{2}{m}. \tag{A.120}$$

By setting $W_m^{(k)} = W_{l(m)}^{(k),L(m)}$ bounded continuous function, we have by triangular inequality

$$\log \mathbb{E}\Big[\exp\Big(m|W^{(k)} - W_m^{(k)}|(X_1,\dots,X_k)\Big)\Big] \le \frac{2}{m}. \tag{A.121}$$

Since by Jensen's inequality, we have for all $\lambda > 0$,

$$\forall m \ge \lambda, \quad \log \mathbb{E}\Big[\exp\Big(\lambda|W^{(k)} - W_m^{(k)}|(X_1,\dots,X_k)\Big)\Big] \le \frac{\lambda}{m}\mathbb{E}\Big[\exp\Big(m|W^{(k)} - W_m^{(k)}|(X_1,\dots,X_k)\Big)\Big], \tag{A.122}$$

we deduce that

$$\log \mathbb{E}\Big[\exp\Big(\lambda|W^{(k)} - W_m^{(k)}|(X_1,\dots,X_k)\Big)\Big] \overset{m\to+\infty}{\longrightarrow} 0. \tag{A.123}$$

For all $\delta > 0$ and $\lambda > 0$, by the Markov-Tchebychev inequality, we have

$$\mathbb{P}(|U_n(W^{(k)}) - U_n(W_m^{(k)})| > \delta) \le e^{-n\lambda\delta}\mathbb{E}\Big[\exp\Big(n\lambda U_n(|W^{(k)} - W_m^{(k)}|)\Big)\Big]. \tag{A.124}$$

From the above (Section 5.2), we deduce that

$$\frac{1}{n}\log \mathbb{P}(|U_n(W^{(k)}) - U_n(W_m^{(k)})| > \delta) \le -\lambda\delta + \frac{1}{k}\log \mathbb{E}\Big[\exp\Big(kC_k\lambda|W^{(k)} - W_m^{(k)}|(X_1,\dots,X_k)\Big)\Big]. \tag{A.125}$$

We conclude that we have the expected result when $m \to +\infty$ since $\lambda$ is arbitrary. □

*Proof of Proposition 5.12.* Let $Q^i(\cdot|x_{[1,i-1]})$ be the conditional distribution of $x_i$ knowing $x_{[1,i-1]} := (x_1,\cdots,x_{i-1})$ (not knowing if $i = 1$). We have:

$$\mathbf{H}[Q|\prod_{i=1}^N \alpha_i] = \mathbb{E}_Q\Big[\log\Big(\frac{dQ}{d\prod_{i=1}^N \alpha_i}\Big)\Big] = \mathbb{E}_Q\Big[\sum_{i=1}^N \log\Big(\frac{Q^i(dx_i|x_{[1,i-1]})}{\alpha_i(dx_i)}\Big)\Big] = \mathbb{E}_Q\Big[\sum_{i=1}^N \mathbf{H}[Q^i(\cdot|x_{[1,i-1]})|\alpha_i]\Big]. \tag{A.126}$$

Since

$$\mathbb{E}_Q[Q^i(\cdot|x_{[1,i-1]})] = Q_i(\cdot), \tag{A.127}$$

we obtain by convexity of the relative entropy (Jensen's inequality):

$$\mathbb{E}_Q\Big[\mathbf{H}[Q^i(\cdot|x_{[1,i-1]})|\alpha_i]\Big] \ge \mathbf{H}[Q_i|\alpha_i] \tag{A.128}$$

Which shows that we have the result of the proposition. □

*Proof of Proposition 5.13.* For $f$ a measurable function on E, we define:

$$\Lambda_\mu(f) := \log(\mathbb{E}_\mu[e^f]) = \log \int e^f \, d\mu \in (-\infty, +\infty] \tag{A.129}$$

the log-Laplace transformation under μ which is convex in $f$ by Holder's inequality. We have:

$$\Lambda_{\mu_U}(f) = \log \int e^f \, d\mu_U = \Lambda_\mu(f - U) - \Lambda_\mu(-U) \le \frac{1}{p}\Lambda_\mu(-pU) + \frac{1}{q}\Lambda_\mu(qf) - \Lambda_\mu(-U) \tag{A.130}$$

by Holder's inequality considering the conjugate exponent $q := \frac{p}{p-1}$ of $p$. By the variational formula of Donsker-Varadhan, we deduce that:

$$\mathbf{H}[\nu|\mu_U] = \sup_{f \in \mathscr{M}_b(E)} \left\{ \int f \, d\nu - \Lambda_{\mu_U}(f) \right\} \ge \sup_{f \in \mathscr{M}_b(E)} \left\{ \int f \, d\nu - \frac{1}{q}\Lambda_\mu(qf) \right\} + \Lambda_\mu(-U) - \frac{1}{p}\Lambda_\mu(-pU). \tag{A.131}$$

Gold:

$$\sup_{f \in \mathscr{M}_b(E)} \left\{ \int f \, d\nu - \frac{1}{q}\Lambda_\mu(qf) \right\} + \Lambda_\mu(-U) - \frac{1}{p}\Lambda_\mu(-pU) = \frac{1}{q}\mathbf{H}[\nu|\mu] + \Lambda_\mu(-U) - \frac{1}{p}\Lambda_\mu(-pU). \tag{A.132}$$

So if $\mathbf{H}[\nu|\mu_U] < +\infty$, $\mathbf{H}[\nu|\mu] < +\infty$ or equivalently, $\log(\frac{d\nu}{d\mu}) \in L^1(\nu)$ and:

$$\log\left(\frac{d\nu}{d\mu_U}\right) = \log\left(\frac{d\nu}{d\mu}\right) + U + \Lambda_\mu(-U) \in L^1(\nu). \tag{A.133}$$

This proves the proposition. □

## References


[1] Wei Liu and Liming Wu. Large deviations for empirical measures of mean-field Gibbs measures. *Stochastic Process. Appl.*, 130(2):503–520, 2020. *6 quotes pages 1, 6, 20, 21, 23, and 25*

[2] Arnaud Guillin, Wei Liu, Liming Wu, and Chaoen Zhang. Uniform Poincaré and logarithmic Sobolev inequalities for mean field particle systems. *Ann. Appl. Probab.*, 32(3):1590–1614, 2022. *9 quotes pages 1, 2, 4, 7, 8, 10, 12, 13, and 27*

[3] Arnaud Guillin, Wei Liu, Liming Wu, and Chaoen Zhang. The kinetic Fokker-Planck equation with mean field interaction. *J. Math. Pures Appl. (9)*, 150:1–23, 2021. *8 quotes pages 1, 2, 4, 8, 13, 18, 29, and 31*

[4] Arnaud Guillin and Pierre Monmarché. Uniform long-time and propagation of chaos estimates for mean field kinetic particles in non-convex landscapes. *J. Stat. Phys.*, 185(2):Paper No. 15, 20, 2021. *2 quotes pages 1 and 7*

[5] Pierre Monmarché. Long-time behaviour and propagation of chaos for mean field kinetic particles. *Stochastic Process. Appl.*, 127(6):1721–1737, 2017. *5 quotes pages 1, 7, 8, 18, and 31*

[6] Pierre Monmarché. Generalized Γ calculus and application to interacting particles on a graph. *Potential Anal.*, 50(3):439–466, 2019. *4 quotes pages 1, 8, 18, and 31*

[7] Louis-Pierre Chaintron and Antoine Diez. Propagation of chaos: A review of models, methods and applications. I. Models and methods. *Kinet. Relat. Models*, 15(6):895–, 2022. *4 quotes pages 4, 6, 17, and 43*

[8] T. H. R. Skyrme. A non-linear field theory. *Proc. Roy. Soc. London Ser. A*, 260:127–138, 1961. *2 quotes pages 4 and 14*

[9] Kaitong Hu, Zhenjie Ren, David Šiška, and $\mathscr{L}$ ukasz Szpruch. Mean-field Langevin dynamics and energy landscape of neural networks. *Ann. Inst. Henri Poincaré Probab. Stat.*, 57(4):2043–2065, 2021. *quoted page 5*

[10] René Carmona and François Delarue. *Probabilistic theory of mean field games with applications. I*, volume 83 of *Probability Theory and Stochastic Modelling*. Springer, Cham, 2018. Mean field FBSDEs, control, and games. *quoted page 5*

[11] Luigi Ambrosio, Nicola Gigli, and Giuseppe Savaré. *Gradient flows in metric spaces and in the space of probability measures*. Lectures in Mathematics ETH Zürich. Birkhäuser Verlag, Basel, 2005. *3 quotes pages 5, 6, and 7*

[12] Gérard Ben Arous and Marc Brunaud. Méthode de Laplace: étude variationnelle des fluctuations de diffusions de type "champ moyen". *Stochastics Stochastics Rep.*, 31(1-4):79–144, 1990. *quoted page 6*

[13] Shigeo Kusuoka and Yozo Tamura. Gibbs measures for mean field potentials. *J. Fac. Sci. Univ. Tokyo Sect. IA Math.*, 31(1):223–245, 1984. *quoted page 6*

[14] E. Bolthausen. Laplace approximations for sums of independent random vectors. *Probab. Theory Relat. Fields*, 72(2):305–318, 1986. *quoted page 6*

[15] Felix Otto. The geometry of dissipative evolution equations: the porous medium equation. *Comm. Partial Differential Equations*, 26(1-2):101–174, 2001. *3 quotes pages 6, 7, and 12*

[16] Ansgar Jüngel. *Entropy methods for diffusive partial differential equations*. SpringerBriefs in Mathematics. Springer, [Cham], 2016. *quoted page 6*

[17] Dominique Bakry, Ivan Gentil, and Michel Ledoux. *Analysis and geometry of Markov diffusion operators*, volume 348 of *Grundlehren der mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer, Cham, 2014. *3 quotes pages 6, 7, and 17*

[18] Jean Dolbeault. Functional inequalities: nonlinear flows and entropy methods as a tool for obtaining sharp and constructive results. *Milan J. Math.*, 89(2):355–386, 2021. *quoted page 6*

[19] Linshan Liu, Mateusz B. Majka, and Łukasz Szpruch. Polyak-lojasiewicz inequality on the space of measures and convergence of mean-field birth-death processes. *arXiv*, 2022. *quoted page 7*

[20] Cédric Villani. *Optimal transport*, volume 338 of *Grundlehren der mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, 2009. Old and new. *2 quotes pages 7 and 28*

[21] Stéphane Mischler and Clément Mouhot. Kac's program in kinetic theory. *Invent. Math.*, 193(1):1–147, 2013. *quoted page 7*

[22] D. Bakry and Michel Émery. Diffusions hypercontractives. In *Séminaire de probabilités, XIX, 1983/84*, volume 1123 of *Lecture Notes in Math.*, pages 177–206. Springer, Berlin, 1985. *2 quotes pages 7 and 17*

[23] José A. Carrillo, Robert J. McCann, and Cédric Villani. Kinetic equilibration rates for granular media and related equations: entropy dissipation and mass transportation estimates. *Rev. Mat. Iberoamericana*, 19(3):971–1018, 2003. *3 quotes pages 7, 10, and 12*

[24] Cédric Villani. Hypocoercivity. *Mem. Amer. Math. Soc.*, 202(950):iv+141, 2009. *4 quotes pages 8, 13, 18, and 31*

[25] Alain-Sol Sznitman. Topics in propagation of chaos. In *École d'Été de Probabilités de Saint-Flour XIX—1989*, volume 1464 of *Lecture Notes in Math.*, pages 165–251. Springer, Berlin, 1991. *quoted page 10*

[26] Robert J. McCann. A convexity principle for interacting gases. *Adv. Math.*, 128(1):153–179, 1997. *quoted page 10*

[27] Patrick Cattiaux, Arnaud Guillin, and Li-Ming Wu. A note on Talagrand's transportation inequality and logarithmic Sobolev inequality. *Probab. Theory Related Fields*, 148(1-2):285–304, 2010. *3 quotes pages 12, 33, and 35*

[28] Patrick Cattiaux, Arnaud Guillin, and Li-Ming Wu. Some remarks on weighted logarithmic Sobolev inequality. *Indiana Univ. Math. J.*, 60(6):1885–1904, 2011. *3 quotes pages 12, 33, and 35*

[29] Boguslaw Zegarlinski. The strong decay to equilibrium for the stochastic dynamics of unbounded spin systems on a lattice. *Comm. Math. Phys.*, 175(2):401–432, 1996. *quoted page 12*

[30] Bogusł aw Zegarliński. Dobrushin uniqueness theorem and logarithmic Sobolev inequalities. *J. Funct. Anal.*, 105(1):77–111, 1992. *quoted page 12*

[31] Gilles Royer. *Une initiation aux inégalités de Sobolev logarithmiques*, volume 5 of *Cours Spécialisés [Specialized Courses]*. Société Mathématique de France, Paris, 1999. *quoted page 12*

[32] V. S. Koroljuk and Yu. V. Borovskich. *Theory of* U*-statistics*, volume 273 of *Mathematics and its Applications*. Kluwer Academic Publishers Group, Dordrecht, 1994. Translated from the 1989 Russian original by P. V. Malyshev and D. V. Malyshev and revised by the authors. *quoted page 20*

[33] Victor H. de la Peña. Decoupling and Khintchine's inequalities for U-statistics. *Ann. Probab.*, 20(4):1877–1892, 1992. *quoted page 20*

[34] F. Otto and C. Villani. Generalization of an inequality by Talagrand and links with the logarithmic Sobolev inequality. *J. Funct. Anal.*, 173(2):361–400, 2000. *quoted page 28*

[35] Songbo Wang. Uniform log-sobolev inequalities for mean field particles with flat-convex energy, 2024. *4 quotes pages 33, 34, 35, and 36*

[36] Shigeki Aida and Ichiro Shigekawa. Logarithmic Sobolev inequalities and spectral gaps: perturbation theory. *J. Funct. Anal.*, 126(2):448–475, 1994. *2 quotes pages 34 and 36*

[37] C. Léonard. Large deviations and law of large numbers for a mean field type interacting particle system. *Stochastic Process. Appl.*, 25(2):215–235, 1987. *no quotes*

[38] S. Mischler. Introduction aux limites de champ moyen pour des systèmes de particules. In *DEA*, pp.166. cel-00576329. Université Paris-Dauphine, 2011. *no quotes*

[39] Patrick Cattiaux and Arnaud Guillin. Hitting times, functional inequalities, Lyapunov conditions and uniform ergodicity. *J. Funct. Anal.*, 272(6):2361–2391, 2017. *no quotes*

[40] S. Mischler. Kac's chaos and Kac's program. In *Séminaire Laurent Schwartz—Équations aux dérivées partielles et applications. Année 2012–2013*, Sémin. Équ. Dériv. Partielles, pages Exp. No. XXII, 17. École Polytech., Palaiseau, 2014. *no quotes*

[41] Dominique Bakry, Franck Barthe, Patrick Cattiaux, and Arnaud Guillin. A simple proof of the Poincaré inequality for a large class of probability measures including the log-concave case. *Electron. Commun. Probab.*, 13:60–66, 2008. *no quotes*

[42] Dominique Bakry, Patrick Cattiaux, and Arnaud Guillin. Rate of convergence for ergodic continuous Markov processes: Lyapunov versus Poincaré. *J. Funct. Anal.*, 254(3):727–759, 2008. *no quotes*

[43] Roland Bauerschmidt and Thierry Bodineau. A very simple proof of the LSI for high temperature spin systems. *J. Funct. Anal.*, 276(8):2582–2588, 2019. *no quotes*

[44] S. Bobkov and M. Ledoux. Poincaré's inequalities and Talagrand's concentration phenomenon for the exponential distribution. *Probab. Theory Related Fields*, 107(3):383–400, 1997. *no quotes*

[45] Felix Otto and Cédric Villani. Comment on: "Hypercontractivity of Hamilton-Jacobi equations" [J. Math. Pures Appl. (9) **80** (2001), no. 7, 669–696; MR1846020 (2003b:47073)] by S. G. Bobkov, I. Gentil and M. Ledoux. *J. Math. Pures Appl. (9)*, 80(7):697–700, 2001. *no quotes*

[46] T. Bodineau and B. Helffer. The log-Sobolev inequality for unbounded spin systems. *J. Funct. Anal.*, 166(1):168–178, 1999. *no quotes*

[47] T. Bodineau and B. Helffer. Correlations, spectral gap and log-Sobolev inequalities for unbounded spins systems. In *Differential equations and mathematical physics (Birmingham, AL, 1999)*, volume 16 of *AMS/IP Stud. Adv. Math.*, pages 51–66. Amer. Math. Soc., Providence, RI, 2000. *no quotes*

[48] François Bolley, Ivan Gentil, and Arnaud Guillin. Uniform convergence to equilibrium for granular media. *Arch. Ration. Mech. Anal.*, 208(2):429–445, 2013. *no quotes*

[49] P. Cattiaux, A. Guillin, and F. Malrieu. Probabilistic approach for granular media equations in the non-uniformly convex case. *Probab. Theory Related Fields*, 140(1-2):19–40, 2008. *no quotes*

[50] Alain Durmus, Andreas Eberle, Arnaud Guillin, and Raphael Zimmer. An elementary approach to uniform in time propagation of chaos. *Proc. Amer. Math. Soc.*, 148(12):5387–5398, 2020. *no quotes*

[51] Andreas Eberle. Reflection couplings and contraction rates for diffusions. *Probab. Theory Related Fields*, 166(3-4):851–886, 2016. *no quotes*

[52] Andreas Eberle, Arnaud Guillin, and Raphael Zimmer. Quantitative Harris-type theorems for diffusions and McKean-Vlasov processes. *Trans. Amer. Math. Soc.*, 371(10):7135–7173, 2019. *no quotes*

[53] Arnaud Guillin, Christian Léonard, Liming Wu, and Nian Yao. Transportation-information inequalities for Markov processes. *Probab. Theory Related Fields*, 144(3-4):669–695, 2009. *no quotes*

[54] A. Guionnet and B. Zegarlinski. Lectures on logarithmic Sobolev inequalities. In *Séminaire de Probabilités, XXXVI*, volume 1801 of *Lecture Notes in Math.*, pages 1–134. Springer, Berlin, 2003. *no quotes*

[55] M. Ledoux. Logarithmic Sobolev inequalities for unbounded spin systems revisited. In *Séminaire de Probabilités, XXXV*, volume 1755 of *Lecture Notes in Math.*, pages 167–194. Springer, Berlin, 2001. *no quotes*

[56] F. Malrieu. Logarithmic Sobolev inequalities for some nonlinear PDE's. *Stochastic Process. Appl.*, 95(1):109–132, 2001. *no quotes*

[57] Florent Malrieu. Convergence to equilibrium for granular media equations and their Euler schemes. *Ann. Appl. Probab.*, 13(2):540–560, 2003. *no quotes*

[58] Daniel W. Stroock and Bogusł aw Zegarliński. The equivalence of the logarithmic Sobolev inequality and the Dobrushin-Shlosman mixing condition. *Comm. Math. Phys.*, 144(2):303–323, 1992. *no quotes*

[59] Daniel W. Stroock and Bogusł aw Zegarliński. The logarithmic Sobolev inequality for continuous spin systems on a lattice. *J. Funct. Anal.*, 104(2):299–326, 1992. *no quotes*

[60] Daniel W. Stroock and Bogusł aw Zegarliński. The logarithmic Sobolev inequality for discrete spin systems on a lattice. *Comm. Math. Phys.*, 149(1):175–193, 1992. *no quotes*

[61] Liming Wu. Gradient estimates of Poisson equations on Riemannian manifolds and applications. *J. Funct. Anal.*, 257(12):4015–4033, 2009. *no quotes*

[62] Nobuo Yoshida. Application of log-Sobolev inequality to the stochastic dynamics of unbounded spin systems on the lattice. *J. Funct. Anal.*, 173(1):74–102, 2000. *no quotes*

[63] Nobuo Yoshida. The equivalence of the log-Sobolev inequality and a mixing condition for unbounded spin systems on the lattice. *Ann. Inst. H. Poincaré Probab. Statist.*, 37(2):223–243, 2001. *no quotes*

[64] Kaitong Hu, Zhenjie Ren, David Siska, and Lukasz Szpruch. Mean-field Langevin dynamics and energy landscape of neural networks. *arXiv preprint arXiv:1905.07769*, 2019. *no quotes*

[65] Wuchen Li Erhan Bayraktar, Qi Feng. Exponential entropy dissipation for weakly self-consistent vlasov-fokker-planck equations. *arXiv:2204.12049v1*, 2022. *no quotes*

[66] René Carmona and François Delarue. *Probabilistic theory of mean field games with applications. II*, volume 84 of *Probability Theory and Stochastic Modelling*. Springer, Cham, 2018. Mean field games with common noise and master equations. *no quotes*

[67] Ludwig Boltzmann. Journey of a German professor to Eldorado. *Transport Theory Statist. Phys.*, 20(5-6):499–523, 1991. Translated from the German. *no quotes*

[68] Ludwig Boltzmann. *Wissenschaftliche Abhandlungen von Ludwig Boltzmann. I. Band (1865–1874); II. Band (1875–1881); III. Band (1882–1905)*. Chelsea Publishing Co., New York, 1968. Herausgegeben von Fritz Hasenöhrl. *no quotes*

[69] Ludwig Boltzmann. *Lectures on gas theory*. University of California Press, Berkeley-Los Angeles, Calif., 1964. Translated by Stephen G. Brush. *no quotes*

[70] M. Kac. Foundations of kinetic theory. In *Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability, 1954–1955, vol. III*, pages 171–197. University of California Press, Berkeley-Los Angeles, Calif., 1956. *no quotes*

[71] M. Kac. The Boltzmann equation: theory and applications. In E. G. D. Cohen and W. Thirring, editors, *Proceedings of the International Symposium "100 Years Boltzmann Equation" in Vienna, 4th-8th September 1972*, pages xii+642. Springer-Verlag, Vienna-New York, 1973. Acta Physica Austriaca Supplementum, X. *no quotes*

[72] M. Kac. Fluctuations near and far from equilibrium. In *Statistical mechanics and statistical methods in theory and application (Proc. Sympos., Univ. Rochester, Rochester, N.Y., 1976)*, pages 203–218. Plenum, New York, 1977. *no quotes*

[73] Mark Kac. *Aspects probabilistes de la théorie du potentiel*, volume 1968 of *Séminaire de Mathématiques Supérieures, No. 32 (Été*. Les Presses de l'Université de Montréal, Montreal, Que., 1970. *no quotes*

[74] H. P. McKean, Jr. Propagation of chaos for a class of non-linear parabolic equations. In *Stochastic Differential Equations (Lecture Series in Differential Equations, Session 7, Catholic Univ., 1967)*, pages 41–57. Air Force Office Sci. Res., Arlington, Va., 1967. *no quotes*

[75] H. P. McKean, Jr. A class of Markov processes associated with nonlinear parabolic equations. *Proc. Nat. Acad. Sci. U.S.A.*, 56:1907–1911, 1966. *no quotes*

[76] Thierry Bodineau, Isabelle Gallagher, and Laure Saint-Raymond. The Brownian motion as the limit of a deterministic system of hard-spheres. *Invent. Math.*, 203(2):493–553, 2016. *no quotes*

[77] Isabelle Gallagher, Laure Saint-Raymond, and Benjamin Texier. *From Newton to Boltzmann: hard spheres and short-range potentials*. Zurich Lectures in Advanced Mathematics. European Mathematical Society (EMS), Zürich, 2013. *no quotes*

[78] Nicola Bellomo, Pierre Degond, and Eitan Tadmor, editors. *Active particles. Vol. 1. Advances in theory, models, and applications*. Modeling and Simulation in Science, Engineering and Technology. Birkhäuser/Springer, Cham, 2017. *no quotes*

[79] Nicola Bellomo, Pierre Degond, and Eitan Tadmor, editors. *Active particles. Vol. 2. Advances in theory, models, and applications*. Modeling and Simulation in Science, Engineering and Technology. Birkhäuser/Springer, Cham, 2019. *no quotes*

[80] Nicola Bellomo, José Antonio Carrillo, and Eitan Tadmor, editors. *Active particles. Vol. 3. Advances in theory, models, and applications*. Modeling and Simulation in Science, Engineering and Technology. Birkhäuser/Springer, Cham, [2022] ©2022. *no quotes*

[81] Adrian Muntean and Federico Toschi, editors. *Collective dynamics from bacteria to crowds*, volume 553 of *CISM International Centre for Mechanical Sciences. Courses and Lectures*. Springer, Vienna, 2014. An excursion through modeling, analysis and simulation. *no quotes*

[82] Giovanni Naldi, Lorenzo Pareschi, and Giuseppe Toscani, editors. *Mathematical modeling of collective behavior in socio-economic and life sciences*. Modeling and Simulation in Science, Engineering and Technology. Birkhäuser Boston, Ltd., Boston, MA, 2010. *no quotes*

[83] Pierre Cardaliaguet, François Delarue, Jean-Michel Lasry, and Pierre-Louis Lions. *The master equation and the convergence problem in mean field games*, volume 201 of *Annals of Mathematics Studies*. Princeton University Press, Princeton, NJ, 2019. *no quotes*

[84] P. Del Moral. Measure-valued processes and interacting particle systems. Application to nonlinear filtering problems. *Ann. Appl. Probab.*, 8(2):438–495, 1998. *no quotes*

[85] Pierre Del Moral. *Feynman-Kac formulae*. Probability and its Applications (New York). Springer-Verlag, New York, 2004. Genealogical and interacting particle systems with applications. *no quotes*

[86] Pierre Del Moral. *Mean field simulation for Monte Carlo integration*, volume 126 of *Monographs on Statistics and Applied Probability*. CRC Press, Boca Raton, FL, 2013. *no quotes*

[87] René Pinnau, Claudia Totzeck, Oliver Tse, and Stephan Martin. A consensus-based model for global optimization and its mean-field limit. *Math. Models Methods Appl. Sci.*, 27(1):183–204, 2017. *no quotes*

[88] Claudia Totzeck. Trends in consensus-based optimization. In *Active particles. Vol. 3. Advances in theory, models, and applications*, Model. Simul. Sci. Eng. Technol., pages 201–226. Birkhäuser/Springer, Cham, [2022] ©2022. *no quotes*

[89] Sara Grassi and Lorenzo Pareschi. From particle swarm optimization to consensus based optimization: stochastic modeling and mean-field limit. *Math. Models Methods Appl. Sci.*, 31(8):1625–1657, 2021. *no quotes*

[90] José A. Carrillo, Shi Jin, Lei Li, and Yuhua Zhu. A consensus-based global optimization method for high dimensional machine learning problems. *ESAIM Control Optim. Calc. Var.*, 27(suppl.):Paper No. S5, 22, 2021. *no quotes*

[91] Song Mei, Andrea Montanari, and Phan-Minh Nguyen. A mean field view of the landscape of two-layer neural networks. *Proc. Natl. Acad. Sci. USA*, 115(33):E7665–E7671, 2018. *no quotes*

[92] Justin Sirignano and Konstantinos Spiliopoulos. Mean field analysis of neural networks: a law of large numbers. *SIAM J. Appl. Math.*, 80(2):725–752, 2020. *no quotes*

[93] C. Graham, Th. G. Kurtz, S. Méléard, Ph. E. Protter, M. Pulvirenti, and D. Talay. *Probabilistic models for nonlinear partial differential equations*, volume 1627 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin; Centro Internazionale Matematico Estivo (C.I.M.E.), Florence, 1996. Lectures given at the 1st Session and Summer School held in Montecatini Terme, May 22–30, 1995, Edited by Talay and L. Tubaro, Fondazione CIME/CIME Foundation Subseries. *no quotes*

[94] Sylvie Méléard. Asymptotic behaviour of some interacting particle systems; McKean-Vlasov and Boltzmann models. In *Probabilistic models for nonlinear partial differential equations (Montecatini Terme, 1995)*, volume 1627 of *Lecture Notes in Math.*, pages 42–95. Springer, Berlin, 1996. *no quotes*

[95] Mario Pulvirenti. Kinetic limits for stochastic particle systems. In *Probabilistic models for nonlinear partial differential equations (Montecatini Terme, 1995)*, volume 1627 of *Lecture Notes in Math.*, pages 96–126. Springer, Berlin, 1996. *no quotes*

[96] Pierre-Emmanuel Jabin and Zhenfu Wang. Mean field limit for stochastic particle systems. In *Active particles. Vol. 1. Advances in theory, models, and applications*, Model. Simul. Sci. Eng. Technol., pages 379–402. Birkhäuser/Springer, Cham, 2017. *no quotes*

[97] Pierre-Emmanuel Jabin and Zhenfu Wang. Quantitative estimates of propagation of chaos for stochastic systems with $W^{-1,\infty}$ kernels. *Invent. Math.*, 214(1):523–591, 2018. *no quotes*

[98] Maxime Hauray and Stéphane Mischler. On Kac's chaos and related problems. *J. Funct. Anal.*, 266(10):6055–6157, 2014. *no quotes*

[99] Pierre-Emmanuel Jabin. A review of the mean field limits for Vlasov equations. *Kinet. Relat. Models*, 7(4):661–711, 2014. *no quotes*

[100] François Golse. On the dynamics of large particle systems in the mean field limit. In *Macroscopic and large scale phenomena: coarse graining, mean field limits and ergodicity*, volume 3 of *Lect. Notes Appl. Math. Mech.*, pages 1–144. Springer, [Cham], 2016. *no quotes*

[101] Arnaud Guillin and Feng-Yu Wang. Degenerate Fokker-Planck equations: Bismut formula, gradient estimate and Harnack inequality. *J. Differential Equations*, 253(1):20–40, 2012. *no quotes*