Electronic Journal of Probability

Electron. J. Probab. **27** (2022), article no. 124, 1–44.
ISSN: 1083-6489 https://doi.org/10.1214/22-EJP853

# Convergence rates for the Vlasov-Fokker-Planck equation and uniform in time propagation of chaos in non convex cases*

Arnaud Guillin† Pierre Le Bris‡ Pierre Monmarché§

**Abstract**

We prove the existence of a contraction rate for Vlasov-Fokker-Planck equation in Wasserstein distance, provided the interaction potential is Lipschitz continuous and the confining potential is both (locally) Lipschitz continuous and greater than a quadratic function, thus requiring no convexity conditions. Our strategy relies on coupling methods suggested by A. Eberle [22] adapted to the kinetic setting enabling also to obtain uniform in time propagation of chaos in a non convex setting.

**Keywords:** Vlasov-Fokker-Planck equation; long-time convergence; propagation of chaos; coupling method.
**MSC2020 subject classifications:** 60J60; 35K58; 82B40.
Submitted to EJP on November 26, 2021, final version accepted on September 12, 2022.

## 1 Introduction

### 1.1 Framework

Let $U$ and $W$ be two functions in $\mathcal{C}^1\left(\mathbb{R}^d\right)$. We consider the Vlasov-Fokker-Planck equation:

$$\partial_t \nu_t\left(x,v\right) = -\nabla_x \cdot \left(v\nu_t\left(x,v\right)\right) + \nabla_v \cdot \left(\left(v + \nabla U\left(x\right) + \nabla W * \mu_t\left(x\right)\right)\nu_t\left(x,v\right) + \nabla_v \nu_t\left(x,v\right)\right), \tag{1.1}$$

where $\nu_t(x,v)$ is a probability density in the space of positions $x \in \mathbb{R}^d$ and velocities $v \in \mathbb{R}^d$,

$$\mu_t\left(x\right) = \int_{\mathbb{R}^d} \nu_t\left(x,dv\right)$$

*This work has been funded by the project EFI ANR-17-CE40-0030.
†Laboratoire de Mathématiques Blaise Pascal - Université Clermont-Auvergne.
E-mail: arnaud.guillin@uca.fr
‡Laboratoire Jacques-Louis Lions - Sorbonne Université.
E-mail: pierre.lebris@sorbonne-universite.fr
§Laboratoire Jacques-Louis Lions - Sorbonne Université.
E-mail: pierre.monmarche@sorbonne-universite.fr

is the space marginal of $\nu_t$ and

$$\nabla W * \mu_t(x) = \int_{\mathbb{R}^d} \nabla W(x-y)\mu_t(dy).$$

It has the following probabilistic counterpart, the non linear stochastic differential equation of *McKean-Vlasov* type, i.e. $\nu_t$ is the density of the law at time $t$ of the $\mathbb{R}^{2d}$-valued process $(X_t, V_t)_{t\geq 0}$ evolving as the mean field SDE (diffusive Newton's equations)

$$\begin{cases} dX_t = V_t dt \\ dV_t = \sqrt{2} dB_t - V_t dt - \nabla U\left(X_t\right) dt - \nabla W * \mu_t\left(X_t\right) dt \\ \mu_t = \text{Law}\left(X_t\right). \end{cases} \tag{1.2}$$

Here, $(X_t, V_t) \in \mathbb{R}^d \times \mathbb{R}^d$, $(B_t)_{t\geq 0}$ is a Brownian motion in dimension $d$ on a probability space $(\Omega, \mathcal{A}, \mathbb{P})$, and $\mu_t$ is the law of the position $X_t$. The symbol $\nabla$ refers to the gradient operator, and the symbol $*$ to the operation of convolution.

Both in the probability and in the partial differential equation community, existence and uniqueness of McKean-Vlasov processes have been well studied. See [36, 25, 44] for some historical milestones. In the specific case of (1.1) and (1.2), under the assumptions on $U$ and $W$ introduced in the next section, existence and uniqueness follow from [37] for square integrable initial data.

A related process is the $N$ particles system in $\mathbb{R}^d$ in mean field interaction

$$\forall i \in [\![1, N]\!], \qquad \begin{cases} dX_t^i &= V_t^i dt, \\ dV_t^i &= \sqrt{2} dB_t^i - V_t^i dt - \nabla U\left(X_t^i\right) dt - \dfrac{1}{N}\displaystyle\sum_{j=1}^N \nabla W\left(X_t^i - X_t^j\right) dt, \end{cases} \tag{1.3}$$

where $X_t^i$ and $V_t^i$ are respectively the position and the velocity of the i-th particle, and $\left(B_t^i, 1 \leq i \leq N\right)$ are independent Brownian motions in dimension $d$. One can see equation (1.3) as an approximation of equation (1.2), where the law $\mu_t$ is replaced by the empirical measure $\mu_t^N = \frac{1}{N}\sum_{i=1}^N \delta_{X_t^i}$.

It is well known, at least in a non kinetic setting [37, 44], that, under some weak conditions on $U$ and $W$, $\mu_t^N$ converges in some sense toward the law $\mu_t$ of $X_t$ solution of (1.2). This phenomenon has been stated under the name *propagation of chaos*, an idea motivated by M. Kac [32]), and greatly developed by A.S. Sznitman [44]. See the recent reviews on propagation of chaos [16, 15] and references therein for an overview on the subject.

In statistical physics, (1.3) is a Langevin equation that describes the motion of $N$ particles subject to damping, random collisions and a *confining potential* $U$ and interacting with one another through an *interaction potential* $W$, which can be polynomial (granular media), Newtonian (interacting stellar) or Coulombian (charged matter). See for instance [34] for an english translation of P. Langevin's landmark paper on the physics behind the standard underdamped Langevin dynamics. Therefore, Equation (1.1) has the following natural interpretation: the solution $\nu_t$ is the density of the law at time $t$ of the process $(X_t, V_t)_{t\geq 0}$ evolving according to (1.2), and thus describes the limit dynamic of a cloud of (charged) particles. In particular, it holds importance in plasma physics, see [47].

More recently, mean-field processes such as (1.3) have drawn much interest in the analysis of neuron networks in machine learning [18, 17]. In this context of stochastic algorithms, it is known that the underdamped Langevin dynamics (not necessarily with mean-field interactions) can converge faster than the overdamped (i.e non kinetic) Langevin dynamics [18, 28] toward its invariant measure. For example, the results on (1.2) could be applied to study the convergence of the Hamiltonian gradient descent algorithm for the overparametrized optimization as done in [33] for Generative Adversarial Network training.

The goal of the present work is twofold. We are interested, first, in the long-time convergence of the solution of (1.2) toward an equilibrium and, second, to a uniform in time convergence as $N \to +\infty$ of (1.3) toward (1.2). It is well known that such results cannot hold in full generality, as the non-linear equation (1.1) may have several equilibria. Here we will consider cases where the interaction is sufficiently small for the non-linear equilibrium to be unique and globally attractive, and for the propagation of chaos to be uniform in time.

There are various methods to study the long time behavior of kinetic type processes, such as Lyapunov conditions or hypocoercivity, and we will discuss these approaches and compare them with our results later on. We rely here on coupling methods following the guidelines of A. Eberle *et al.* in [23] where the convergence to equilibrium is established for (1.2) without interaction, and also extend the approach to handle only locally Lipschitz coefficient. In a second part, we also use reflection couplings (see [21]) for the propagation of chaos property.

Let us briefly describe the coupling method. The basic idea is that an upper bound on the Wasserstein distance between two probability distributions is given by the construction of any pair of random variables distributed respectively according to those. The goal is thus to construct simultaneously two solutions of (1.2) that have a trend to get closer with time. Have $(X_t, V_t)$ be a solution of (1.2) driven by some Brownian motion $(B_t)_{t\geqslant 0}$ and let $(X'_t, V'_t)$ solves

$$\left\{\begin{array}{l} dX'_t = V'_t dt \\ dV'_t = \sqrt{2} dB'_t - V'_t dt - \nabla U\left(X'_t\right) dt - \nabla W * \mu_t\left(X'_t\right) dt \\ \mu'_t = \text{Law}\left(X'_t\right). \end{array}\right.$$

with $(B'_t)_{t\geqslant 0}$ a d-dimensional Brownian motion. A coupling of $(X, V)$ and $(X', V')$ then follows from a coupling of the Brownian motions $B$ and $B'$. Choosing $B = B'$ yields the so-called *synchronous* coupling, for which the Brownian noise cancels out in the infinitesimal evolution of the difference $(Z_t, W_t) = (X_t - X'_t, V_t - V'_t)$. In that case the contraction of a distance between the processes can only be induced by the deterministic drift, as in [8]. Such a deterministic contraction only holds under very restrictive conditions, in particular $U$ should be strongly convex. Nevertheless, in more general cases, the calculation of the evolution of $Z_t$ and $W_t$ (see Section 3.1 below) shows that there is still some deterministic contraction when $Z_t + W_t = 0$. We can therefore use a synchronous coupling in the vicinity of this subspace.

Outside of $\{(z, w) \in \mathbb{R}^{2d}, z + w = 0\}$, it is necessary to make use of the noise to get the processes closer together, at least in the direction orthogonal to this space. In order to maximize the variance of this noise, we then use a so-called *reflection* coupling, which consists in $B$ and $B'$ being *antithetic* (i.e $B'_t = -B_t$) in the direction of space given by the difference of the processes, and synchronous in the orthogonal direction. In other words, writing

$$e_t = \left\{\begin{array}{ll} \frac{Z_t + W_t}{|Z_t + W_t|} & \text{if } Z_t + W_t \neq 0 \\ 0 & \text{otherwise} \end{array}\right.$$

we consider $dB'_t = \left(Id - 2e_t e_t^T\right) dB_t$. Levy's characterization then ensures that it is indeed a Brownian motion.

Finally we construct a Lyapunov function $H$ to take into account the trend of each process to come back to some compact set of $\mathbb{R}^{2d}$. We are then led to the study of a suitable distance between the two processes, which will be of the form $\rho_t := f(r_t)(1+\epsilon H(X_t,V_t)+\epsilon H(X'_t,V'_t))$, with $r_t = \alpha|Z_t| + |Z_t+W_t|$, where $\alpha, \epsilon > 0$ and the function $f$ are some parameters to choose. More precisely, we have to choose these parameters carefully in order for $\mathbb{E}\rho_t$ to decay exponentially fast. This leads to several constraints on $\alpha, \epsilon$ and on the parameters involved in the definition of $f$, and we have to prove that it is possible to meet all these conditions simultaneously. For the sake of clarity, in fact, we present the proof in a different order, namely we start by introducing very specific parameters and, throughout the proof, we check that our choice of parameters implies the needed constraints.

The study of the limit $N \to +\infty$ is based on a similar coupling, except that we couple a system of $N$ interacting particles (1.3) with $N$ independent non-linear processes (1.2).

The next subsections describe our main results and compare them to the few existing ones in the literature. Section 2 presents the precise construction of the aforementioned *ad hoc* Wasserstein distance. The proof of the long time behavior of the Vlasov-Fokker-Planck equation when confinement and interaction coefficient are Lipschitz continuous is done in Section 3, whereas the propagation of chaos property is proved in Section 4. An appendix gathers technical lemmas and the modifications of the main proofs when the confinement is only supposed locally Lipschitz continuous.

### 1.2 Main results

For $\mu$ and $\nu$ two probability measures on $\mathbb{R}^{2d}$, denote by $\Pi(\mu,\nu)$ the set of couplings of $\mu$ and $\nu$, i.e. the set of probability measures $\Gamma$ on $\mathbb{R}^{2d}\times\mathbb{R}^{2d}$ with $\Gamma(A\times\mathbb{R}^{2d}) = \mu(A)$ and $\Gamma(\mathbb{R}^{2d}\times A) = \nu(A)$ for all Borel set $A$ of $\mathbb{R}^{2d}$. We will define $L^1$ and $L^2$ Wasserstein distances as

$$\mathcal{W}_1(\mu,\nu) = \inf_{\Gamma\in\Pi(\mu,\nu)} \int \left(|x-\tilde{x}| + |v-\tilde{v}|\right)\Gamma\left(dxdvd\tilde{x}d\tilde{v}\right),$$
$$\mathcal{W}_2(\mu,\nu) = \left(\inf_{\Gamma\in\Pi(\mu,\nu)} \int \left(|x-\tilde{x}|^2 + |v-\tilde{v}|^2\right)\Gamma\left(dxdvd\tilde{x}d\tilde{v}\right)\right)^{1/2}.$$

Our main results will be stated in terms of these distances, even if we work and get contraction in the Wasserstein distance defined with the aformentioned $\rho$. Let us detail the assumptions on the potentials $U$ and $W$.

**Assumption 1.1.** *The potential $U$ is non-negative and there exist $\lambda > 0$ and $A \geq 0$ such that*

$$\forall x \in \mathbb{R}^d, \qquad \frac{1}{2}\nabla U(x)\cdot x \geq \lambda\left(U(x) + \frac{|x|^2}{4}\right) - A. \tag{1.4}$$

The condition (1.4) implies that the force $-\nabla U$ has a confining effect, bringing back particles toward some compact set. It implies the following:

**Lemma 1.2.** *If Assumption 1.1 holds, then there exists $\tilde{A} \geq 0$ such that for all $x \in \mathbb{R}^d$,*

$$U(x) \geq \frac{\lambda}{6}|x|^2 - \tilde{A}. \tag{1.5}$$

The proof is postponed to Appendix A.1. In particular, it implies that $U$ goes to infinity at infinity and is bounded below. Since only the gradient of $U$ is involved in the dynamics, the condition $U \geqslant 0$ is thus not restrictive as it can be enforced without loss of generality by adding a sufficient large constant to $U$. This condition is added in order to simplify some calculations.

We will also assume that the potential U satisfies one of the two following conditions:

**Assumption 1.3.** *There is a constant $L_U > 0$ such that*

$$\forall x, y \in \mathbb{R}^d \times \mathbb{R}^d, \qquad |\nabla U(x) - \nabla U(y)| \leq L_U|x-y|.$$

**Assumption 1.4.** *There exist $L_U > 0$ and a function $\psi : \mathbb{R}^d \mapsto \mathbb{R}$ such that*

$$\forall x, y \in \mathbb{R}^d \times \mathbb{R}^d, \qquad |\nabla U(x) - \nabla U(y)| \leq (L_U + \psi(x) + \psi(y))\, |x - y|,$$

*and*

$$\forall x \in \mathbb{R}^d, \qquad 0 \leq \psi(x) \leq L_\psi \sqrt{\lambda |x|^2 + 24U(x)},$$

*where $L_\psi > 0$ is sufficiently small in the sense that*

$$L_\psi \leq c_\psi(L_U, \lambda, \tilde{A}, d, a),$$

*where $c_\psi$ is an explicit function given below in* (5.9)*, $L_U$ is given in Assumption 1.3, $\lambda$ by Assumption 1.1, $\tilde{A}$ by Lemma 1.2, $d$ is the dimension and $a$ is a parameter such that* (5.1) *holds for some $\mathcal{C}^0$, namely is used to bound an initial moment.*

Obviously, Assumption 1.4 implies Assumption 1.3. We distinguish it as it yields simpler proofs. Actually, the proofs of our main results already rely on quite involved computations under Assumption 1.3, and thus for the convenience of the reader we present the proofs in this case with full details in a first step, and then in a second step we explain how the more general situation of Assumption 1.4 is tackled.

**Remark 1.5.** In the literature, see for instance [40] or the recent [10], it is common to find the assumption $U$ twice continuously differentiable with an hessian matrix satisfying

$$||\nabla_x^2 U(x)|| \leq C(1 + |\nabla_x U(x)|), \tag{1.6}$$

where $||\nabla_x^2 U(x)||$ denotes the matrix norm of the hessian. Here, Assumption 1.4 together with Assumption 1.1 yield a stronger version of (1.6). Indeed, in dimension one for instance, we have

$$\begin{aligned} |U''(x)| &= \lim_{y \to x} \frac{|U'(x) - U'(y)|}{|x - y|} \\ &\leq L_U + 2L_\psi \sqrt{\lambda |x|^2 + 24U(x)}. \end{aligned}$$

Using Assumption 1.1, we obtain the existence of a constant $\hat{A}$ such that

$$|U'(x)| \geq \frac{\lambda}{4}|x| - \hat{A},$$

which implies, once again using Assumption 1.1, that

$$|U'(x)| \left( \frac{4}{\lambda}|U'(x)| + \frac{4}{\lambda}\hat{A} \right) \geq |U'(x)||x| \geq U'(x)x \geq 2\lambda U(x) + \frac{\lambda}{2}|x|^2 - 2A.$$

In particular, there are constants $c_1$, $c_2$, $c_3$ and $c_4$ such that

$$c_1|U'(x)| + c_2 \geq \sqrt{c_3 U(x) + c_4 |x|^2},$$

and therefore we obtain, for some constants $C$ and $\eta$,

$$|U''(x)| \leq C(1 + \eta |U'(x)|),$$

where $\eta$ has to be sufficiently small. This is no surprise as, in our work, we consider the "local Lipschitz condition" to be a perturbation of the global Lipschitz Assumption 1.3.

**Example 1.6.** Assume d=1. The double-well potential given by

$$U(x) = \begin{cases} \left(x^2 - 1\right)^2 & \text{if } |x| \leq 1, \\ \left(|x| - 1\right)^2 & \text{otherwise.} \end{cases}$$

satisfies Assumptions 1.1 and 1.3.

**Example 1.7.** Likewise, we may consider $U(x) = \frac{1}{2}x^2 + \frac{3}{2}\cos(x)$ in dimension 1, which is neither strongly convex, nor strongly convex outside a ball, but satisfies Assumptions 1.1 and 1.3.

**Example 1.8.** Consider $U(x) = \frac{1}{4}x^2 + \frac{b}{4}x^4$ in dimension 1. We have

$$\nabla U(x) \cdot x = \frac{x^2}{2} + bx^4 \geq \left(\frac{x^2}{4} + \frac{x^2}{4} + \frac{b}{4}x^4\right) = \left(U(x) + \frac{x^2}{4}\right),$$

hence $U$ satisfies Assumption 1.1. $U$ is not Lipschitz continuous, however it satisfies

$$\begin{aligned}|\nabla U(x) - \nabla U(y)| =& \frac{1}{2}|x-y| + b\left|x^3 - y^3\right| \\ =& \frac{1}{2}|x-y| + b|x-y|\left|x^2 + xy + y^2\right| \\ \leq& \frac{1}{2}|x-y| + \frac{3b}{2}|x-y|\left|x^2 + y^2\right| \\ =& \left(\frac{1}{2} + \psi(x) + \psi(y)\right)|x-y|,\end{aligned}$$

where,

$$\psi(x) = \frac{3b}{2}x^2 \leq \sqrt{b}\sqrt{24\frac{b}{4}x^4} \leq \sqrt{b}\sqrt{\lambda|x|^2 + 24U(x)}.$$

We then require $b$ to be sufficiently small for Assumption 1.4 to hold.

Let us now give the assumption on the interaction potential.

**Assumption 1.9.** *The potential $W$ is even, i.e. $W(x) = W(-x)$ for all $x \in \mathbb{R}^d$, in particular $\nabla W(0) = 0$. Moreover, there exists $L_W < \lambda/8$ (where $\lambda$ is given in Assumption 1.1) such that*

$$\forall x, y \in \mathbb{R}^d \times \mathbb{R}^d, \qquad |\nabla W(x) - \nabla W(y)| \leq L_W|x-y|. \tag{1.7}$$

*In particular $|\nabla W(x)| \leq L_W|x|$ for all $x \in \mathbb{R}^d$.*

Here we consider an interaction force that is the gradient of a potential $W$, as we stick to the formalism of other related works (for instance [21]). Nevertheless, all the results and proofs still hold if $\nabla W$ is replaced by some $F : \mathbb{R}^d \mapsto \mathbb{R}^d$ satisfying the same conditions. The confinement potential may also be non gradient, however the fact that the confinement force $\nabla U$ is a gradient simplifies the construction of a Lyapunov function.

The condition $L_W \leq \lambda/8$ is related to the fact the interaction is considered as a perturbation of the non-interacting process studied in [23]. Therefore, $\nabla W$ has to be controlled by $\nabla U$ in some sense. Note that we immediately get the following bound on the non-linear drift:

**Lemma 1.10.** *Under Assumption 1.9, for all probability measures $\mu$ and $\nu$ on $\mathbb{R}^d$ and $x, \tilde{x} \in \mathbb{R}^d$,*

$$|\nabla W * \mu(x) - \nabla W * \nu(\tilde{x})| \leq L_W|x - \tilde{x}| + L_W\mathcal{W}_1(\mu, \nu).$$

See Appendix A.2 for the proof.

**Example 1.11.** Assumption 1.9 is satisfied for an harmonic interaction $W(x) = \pm L_W|x|^2/2$, or a mollified Coulomb interaction for $a, b > 0$ and $k \in \mathbb{N}^*$

$$W(x) = \pm\frac{a}{\left(|x|^k + b^k\right)^{\frac{1}{k}}}, \quad \text{i.e} \quad \nabla W(x) = \mp\frac{ax|x|^{k-2}}{\left(|x|^k + b^k\right)^{1+\frac{1}{k}}}.$$

The first of our main results concern the long-time convergence of the non-linear system (1.1).

**Theorem 1.12.** *Let $U$ be continuously differentiable and satisfy Assumption 1.1 and Assumption 1.4. There is an explicit $c^W > 0$ such that, for all $W$ continuously differentiable satisfying Assumption 1.9 with $L_W < c^W$, there is an explicit $\tau > 0$ such that for all probability measures $\nu_0^1$ and $\nu_0^2$ on $\mathbb{R}^{2d}$ with either a finite second moment (if Assumption 1.3 holds) or a finite Gaussian moment (if only Assumption 1.4 holds), there are explicit constants $C_1, C_2 > 0$ such that for all $t \geqslant 0$,*

$$\mathcal{W}_1\left(\nu_t^1, \nu_t^2\right) \leq e^{-\tau t} C_1\,, \qquad \mathcal{W}_2\left(\nu_t^1, \nu_t^2\right) \leq e^{-\tau t} C_2$$

*where $\nu_t^1$ and $\nu_t^2$ are solutions of* (1.1) *with respective initial distributions $\nu_0^1$ and $\nu_0^2$.*

*In particular, we have existence and uniqueness of, as well as convergence towards, a stationary solution.*

The second of our main results is a uniform in time convergence as $N \to +\infty$ of (1.3) toward (1.2).

**Theorem 1.13.** *Let $\tilde{\mathcal{C}}^0 > 0$ and $\tilde{a} > 0$. Let $U$ be continuously differentiable and satisfy Assumptions 1.1 and 1.3. There is an explicit $c^W > 0$ such that, for all $W$ continuously differentiable satisfying Assumption 1.9 with $L_W < c^W$, there exist explicit $B_1, B_2 > 0$, such that for all probability measures $\nu_0$ on $\mathbb{R}^{2d}$ satisfying $\mathbb{E}_{\nu_0}\left(e^{\tilde{a}(|X|+|V|)}\right) \leq \tilde{\mathcal{C}}^0$,*

$$\mathcal{W}_1\left(\nu_t^{k,N}, \bar{\nu}_t^{\otimes k}\right) \leq \frac{kB_1}{\sqrt{N}}\,, \qquad \mathcal{W}_2^2\left(\nu_t^{k,N}, \bar{\nu}_t^{\otimes k}\right) \leq \frac{kB_2}{\sqrt{N}},$$

*for all $k \in \mathbb{N}$, where $\nu_t^{k,N}$ is the marginal distribution at time $t$ of the first $k$ particles $\left((X_t^1, V_t^1), ...., (X_t^k, V_t^k)\right)$ of an $N$ particle system* (1.3) *with initial distribution $(\nu_0)^{\otimes N}$, while $\bar{\nu}_t$ is a solution of* (1.1) *with initial distribution $\nu_0$.*

The organization of the article is as follows: in Section 2 we define the various tools involved in the construction of a good semimetrics. In Section 3 we study the long-time behavior of the Vlasov-Fokker-Planck equation (i.e Theorem 1.12) under the global Lipschitz Assumption 1.3 on $U$. Then, in Section 4, we prove propagation of chaos (i.e Theorem 1.13). Finally, in Section 5, we show how one may obtain the result of Theorem 1.12 under the local Lipschitz Assumption 1.4 on $U$.

We choose to present these proofs in this order, starting with the case in which the computations are the least cumbersome, in order to describe the method and motivate the construction of the semimetrics. Then, we add the tools to deal with the propagation of chaos. Finally, by combining the tools developed in Section 4 and the method of Section 3, we observe that it is possible to handle a small perturbation of the Lipschitz condition on U. In this way, we hope to gradually bring the difficulties and keep a form of clarity despite the sometimes involved calculations.

### 1.3 Comparison to existing works

Space homogeneous (i.e non kinetic) models of diffusive and interacting granular media, usually named McKean-Vlasov diffusions (see [5]), have attracted a lot of attention in the last twenty years. They have been treated by means of a stochastic interpretation and synchronous couplings as in [13] or in the recent [21] by reflection couplings enabling to get rid of convexity conditions, but limited to small interactions. Remark however that small interactions are natural to get uniform in time propagation of chaos as for large interactions the non linear limit equation may have several stationary measures (see [31] for example). The granular media equations were interpreted as gradient flows in the space of probability measures in [12], leading to explicit exponential (or algebraic for non uniformly convex cases) rates of convergence to equilibrium of the non linear equation. Another approach relying on the dissipation of the Wasserstein distance and

$WJ$ inequalities was introduced in [7] handling small non convex cases. This approach was implemented in [42] to get propagation of chaos, under roughly the same type of assumptions. Mean-field limit using $\Gamma$-convergence tools has also been obtained in [11] for $\lambda$-convex potentials in this non kinetic setting.

Results on the long time behavior of the non-linear equation (1.2), i.e. space inhomogeneous, are few, as they combine the difficulty of getting explicit contraction rates for hypoelliptic diffusions as well as a non linear term. Recent works have tackled the question of contraction rate for the underdamped Langevin diffusion when there are no interaction (i.e $W = 0$). Results were obtained using hypocoercivity [19] and recently functional inequalities [1, 10], all in an $L^2$ setting that is not well adapted to the interacting particle system. For singular potential $U$, still without interaction, convergence rate in $H^1$ were obtained in [4]. Concerning the uniform in time propagation of chaos, there are no results except in the strictly convex case (with very small perturbation). We however refer to [46] for a result on the torus with $W$ bounded with continuous derivative of all orders, see also [9]. Using functional inequalities (Poincaré or logarithmic Sobolev inequalities) for mean field models obtained in [27], other results were obtained provided the confining potential is a small perturbation of a quadratic function as in [38, 26, 29] which combines the hypocoercivity approach with independent of the number of particles constants appearing in the logarithmic Sobolev inequalities. The convergence of the Vlasov-Fokker-Planck equation to equilibrium for specific non-convex confining potentials and convex polynomial attractive interaction potentials using the free-energy approach has also been obtained in [20]. Our results generalize [29]. Indeed, we may consider non gradient interactions whereas it is crucial in their approach to know explicitly the invariant measure of the particles system, and also we may handle only locally Lipschitz confinement potential, whereas they impose at most quadratic growth of the potentials, and non strictly convex at infinity potential. It is however difficult to compare the smallness of the interaction potentials needed in both approaches. Note however that they obtain convergence to equilibrium in entropy whereas we get it in Wasserstein distance (controlled by entropy through a Talagrand inequality). Using a coupling strategy, and more precisely synchronous couplings, results under strict convexity assumption were obtained in [8] for contraction rates in Wasserstein distance, see also [33] but only for the nonlinear system.

As we mentioned, we adapt a proof from [23], which tackles (1.2) without interaction term. The article uses a Lyapunov condition that guarantees the recurrence of the process on a compact set. This idea is common when proving similar results through a probabilistic lens (see for instance [45] or [2]). Lyapunov conditions may also help to implement hypocoercivity techniques *à la Villani* to handle entropic convergence for non quadratic potentials, see [14]. Under the assumption $U$ "greater than a quadratic function" at infinity and $\nabla W$ Lipschitz continuous, we too consider a Lyapunov function that allows us to construct a specific semimetric improving the convergence speed. But, and this is to our knowledge something new, when proving propagation of chaos we add a form of non linearity in the quantity we consider to tackle a part of the non linearity appearing in the dynamic (see Section 4 below). Let us also mention the very recent preprint by Schuh [43], posterior to our work, which also aims at proving long time behavior for the second-order Langevin dynamics and its non linear limit as well as uniform in time propagation of chaos, by constructing two separate metrics for small and large distances and showing contraction for both these quantities.

## 2 Modified semimetrics

As mentioned in the introduction, the proofs rely on the construction of suitable semimetrics on $\mathbb{R}^{2d}$ and $\mathbb{R}^{2dN}$. They are introduced in this section, together with some useful properties. In all this section, $\lambda, A, \tilde{A}, L_U$ and $L_W$ are given by Assumptions 1.1, 1.3 and 1.9 and Lemma 1.2.

Before going into the details, let us highlight the main points of the construction of the semimetrics. It relies on the superposition of three ideas. The first idea is that, in order to deal with the kinetic process (1.2), the standard Euclidean norm $|x|^2 + |v|^2$ is not suitable and one should consider a linear change of variables, like $(x, v) \mapsto (x, x + \beta v)$ for some $\beta \in \mathbb{R}$. This is the case when using coupling methods as in [23, 8] but also when using hypocoercive modified entropies involving mixed derivatives as in [46, 45, 3, 14], the link being made in [39]. This motivates the definition of $r$ below. The second idea is a modification of this distance $r$ by some concave function $f$, which is related to the fact we are using, at least in some parts of the space, a reflection coupling. The concavity is well adapted to Itô's formula enabling the diffusion to provide a contraction effect (in a compact). This method has been considered for elliptic diffusions in [22], see also [24]. Intuitively, the contraction is produced by the fact that a random decrease in $r$ has more effect on $f(r)$ than a random increase of the same amount. Finally, the third idea is the multiplication of a distance by a Lyapunov function $G$, which has first been used for Wasserstein distances in [30]. That way, on average, $f(r)G$ tends to decay because, when $r$ is small, $f(r)$ tends to decay and, when $r$ is large, $G$ tends to decay.

### 2.1 A Lyapunov function

Let

$$\gamma = \frac{\lambda}{2(\lambda+1)}, \qquad B = 24\left(A + (\lambda - \gamma)\tilde{A} + d\right) \tag{2.1}$$

and, for $x, v \in \mathbb{R}^d$,

$$H(x,v) = 24U(x) + (6(1-\gamma) + \lambda)|x|^2 + 12x \cdot v + 12|v|^2 .$$

For $\mu$ a probability measure on $\mathbb{R}^d$ with finite first moment, $\nabla W$ being assumed Lipschitz continuous, denote by $\mathcal{L}_\mu$ the generator given by

$$\mathcal{L}_\mu \phi(x,v) = v \cdot \nabla_x \phi(x,v) - (v + \nabla U(x) + \nabla W * \mu(x)) \cdot \nabla_v \phi(x,v) + \Delta_v \phi(x,v) .$$

The main properties of $H$ are the following.

**Lemma 2.1.** *Under Assumptions 1.1, 1.3 and 1.9, for all $x, v \in \mathbb{R}^d$ and $\mu$,*

$$H(x,v) \geq 24U(x) + \lambda|x|^2 + 12\left|v + \frac{x}{2}\right|^2, \tag{2.2}$$

$$\mathcal{L}_\mu H(x,v) \leq B + L_W(6+8\lambda)\left(\int |y| d\mu(y)\right)^2 - \left(\frac{3}{4}\lambda + \lambda^2\right)|x|^2 - \gamma H(x,v), \tag{2.3}$$

$$\mathcal{L}_\mu H(x,v) \leq B + \left(\left(\int |y| d\mu(y)\right)^2 - |x|^2\right)\left(\frac{3}{4}\lambda + \lambda^2\right) - \gamma H(x,v). \tag{2.4}$$

*In particular $H$ is non-negative and goes to $+\infty$ at infinity.*

The proof follows from elementary computations and is detailed in Appendix A.3. Notice that the condition $L_W \leq \lambda/8$ is used here.

In the case of particular interest where $\mu = \mu_t$ is given by (1.2), taking the expectation in (2.4) and using Gronwall's lemma, we immediately get the following.

**Lemma 2.2.** *Under Assumptions 1.1, 1.3 and 1.9, let $(X_t, V_t)_{t\geqslant 0}$ be a solution of (1.2) with finite second moment at initial time. For all $t \geq 0$,*

$$\frac{d}{dt}\mathbb{E}H(X_t, V_t) \leq B - \gamma \mathbb{E}H(X_t, V_t), \tag{2.5}$$

$$\mathbb{E}H(X_t, V_t) - \frac{B}{\gamma} \leq \left(\mathbb{E}H(X_0, V_0) - \frac{B}{\gamma}\right)e^{-\gamma t}. \tag{2.6}$$

### 2.2 Change of variable and concave modification

We start by fixing the values of some parameters. The somewhat intricate expressions in this section are dictated by the computations arising in the proofs later on. Recall the definition of $\gamma$ and $B$ in (2.1). Set

$$\alpha = L_U + \frac{\lambda}{4}\,, \quad R_0 = \sqrt{\frac{24B}{5\gamma\min\left(3,\frac{\lambda}{3}\right)}}\,, \quad R_1 = \sqrt{\frac{24\left((1+\alpha)^2+\alpha^2\right)}{5\gamma\min\left(3,\frac{\lambda}{3}\right)}B}.$$

For $x,\tilde{x},v,\tilde{v}\in\mathbb{R}^d$, set

$$r(x,\tilde{x},v,\tilde{v}) = \alpha|x-\tilde{x}| + |x-\tilde{x}+v-\tilde{v}|\,.$$

**Lemma 2.3.** *Under Assumptions 1.1, 1.3 and 1.9, for all $x,\tilde{x},v,\tilde{v}\in\mathbb{R}^d$,*

$$r(x,\tilde{x},v,\tilde{v})^2 \le 2\frac{(1+\alpha)^2+\alpha^2}{\min\left(\frac{1}{3}\lambda,3\right)}\left(H(x,v)+H(\tilde{x},\tilde{v})\right)\,, \tag{2.7}$$

*so that, in particular,*

$$r(x,\tilde{x},v,\tilde{v}) \ge R_1 \qquad \Rightarrow \qquad \gamma H(x,v)+\gamma H(\tilde{x},\tilde{v}) \ge \frac{12}{5}B.$$

We refer to Appendix A.4 for the proof. Let

$$\begin{aligned} c = \min\Bigg\{ \frac{\gamma}{36}\,,\,\frac{B}{3}\,,\,\frac{1}{7}\min\left(\frac{1}{2}-\frac{L_U+L_W}{2\alpha}, 2\sqrt{\frac{L_U+L_W}{2\pi\alpha}}\right) \\ \times\exp\left(-\frac{1}{8}\left(\frac{L_U+L_W}{\alpha}+\alpha+96\max\left(\frac{1}{2\alpha},1\right)\right)R_1^2\right)\Bigg\}. \end{aligned} \tag{2.8}$$

Set

$$\epsilon = \frac{3c}{B}\,, \qquad \mathbf{C} = c+2\epsilon B$$

and, for $s\ge 0$,

$$\phi(s) = \exp\left(-\frac{1}{8}\left(\frac{1}{\alpha}(L_U+L_W)+\alpha+96\epsilon\max\left(\frac{1}{2\alpha},1\right)\right)s^2\right), \qquad \Phi(s) = \int_0^s \phi(u)\,du$$

$$g(s) = 1-\frac{\mathbf{C}}{4}\int_0^s\frac{\Phi(u)}{\phi(u)}du\,, \qquad f(s) = \int_0^{\min(s,R_1)}\phi(u)\,g(u)\,du.$$

**Remark 2.4.** The parameters above are far from being optimal. They are somewhat roughly chosen as we only wish to convey the fact that every constant is explicit.

The next lemma, proved in Appendix B, gathers the intermediary bounds that will be useful in the proofs of the main results.

**Lemma 2.5.** *Under Assumptions 1.1, 1.3 and 1.9,*

$$c \le \frac{\gamma}{6}\left(1-\frac{\frac{5\gamma}{6}}{2\epsilon B+\frac{5\gamma}{6}}\right), \tag{2.9}$$

$$L_U+L_W < \alpha, \tag{2.10}$$

$$c+2\epsilon B \le \frac{1}{2}\left(1-\frac{L_U+L_W}{\alpha}\right)\inf_{r\in]0,R_1]}\frac{r\phi(r)}{\Phi(r)}, \tag{2.11}$$

$$c+2\epsilon B \le 2\left(\int_0^{R_1}\Phi(s)\,\phi(s)^{-1}\,ds\right)^{-1}, \tag{2.12}$$

$$\forall s\ge 0, \qquad 0 = 4\phi'(s)+\left(\frac{1}{\alpha}(L_U+L_W)+\alpha+96\epsilon\max\left(\frac{1}{2\alpha},1\right)\right)s\phi(s)\,. \tag{2.13}$$

The main properties of $f$ are the following.

**Lemma 2.6.** *The function $f$ is twice continuously differentiable on $(0,R_1)$ with $f'_+(0)=1$ and $f'_-(R_1)>0$, and constant on $[R_1,\infty)$. Moreover, it is non-negative, non-decreasing and concave, and for all $s\geq 0$,*

$$\min(s,R_1)\, f'_-(R_1) \leq f(s) \leq \min(s, f(R_1)) \leq \min(s,R_1)\,.$$

*Proof.* First, notice that (2.12) ensures that $g(s)\geq \frac{1}{2}$ for all $s\geq 0$. Then, all the points immediately follow from the fact the functions $\phi$ and $g$ are twice continuously differentiable, positive and decreasing, with $\phi(0)=g(0)=1$. ☐

### 2.3 The modified semimetrics

For $x,\tilde{x},v,\tilde{v}\in\mathbb{R}^d$, set

$$\begin{aligned} G(x,v,\tilde{x},\tilde{v}) &= 1+\epsilon H(x,v)+\epsilon H(\tilde{x},\tilde{v})\,,\\ \rho(x,v,\tilde{x},\tilde{v}) &= f(r(x,v,\tilde{x},\tilde{v}))\, G(x,v,\tilde{x},\tilde{v})\,. \end{aligned}$$

An immediate corollary of Lemmas 2.3 and 2.6 is that $\rho$ is a semimetric on $\mathbb{R}^{2d}$ which controls the usual L1 and L2 distances:

**Lemma 2.7.** *There are explicit constants $\mathcal{C}_1,\mathcal{C}_2,\mathcal{C}_r,\mathcal{C}_z>0$ such that for all $x,x',v,v'\in\mathbb{R}^d$,*

$$\begin{aligned} |x-x'|+|v-v'| &\leq \mathcal{C}_1\rho((x,v),(x',v'))\\ |x-x'|^2+|v-v'|^2 &\leq \mathcal{C}_2\rho((x,v),(x',v'))\\ r(x,v,x',v') &\leq \mathcal{C}_r\rho((x,v),(x',v'))\\ |x-x'| &\leq \mathcal{C}_z f(r(x,v,x',v'))\left(1+\epsilon\sqrt{H(x,v)}+\epsilon\sqrt{H(x',v')}\right). \end{aligned}$$

We also mention a technical lemma, see Appendix A.6 for proof.

**Lemma 2.8.** *For all $x,v,\tilde{x},\tilde{v}\in\mathbb{R}^d$*

$$|H(x,v)-H(\tilde{x},\tilde{v})| \leq C_{dH,1}r(x,\tilde{x},v,\tilde{v})+C_{dH,2}r(x,\tilde{x},v,\tilde{v})\left(\sqrt{H(x,v)}+\sqrt{H(\tilde{x},\tilde{v})}\right), \tag{2.14}$$

*where*

$$C_{dH,1}:=\frac{24|\nabla U(0)|}{\alpha}\quad\text{and}\quad C_{dH,2}:=\frac{24L_U}{\alpha\sqrt{\lambda}}+\frac{6(1-\gamma)+\lambda-3}{\alpha\sqrt{\lambda}}+2\sqrt{3}\max\left(1,\frac{1}{2\alpha}\right).$$

Finally, for $\mu$ and $\nu$ two probability measures on $\mathbb{R}^{2d}$ and a measurable function $h:\mathbb{R}^{2d}\times\mathbb{R}^{2d}\to\mathbb{R}$, we define

$$\mathcal{W}_h(\mu,\nu)=\inf_{\Gamma\in\Pi(\mu,\nu)}\int h(x,v,\tilde{x},\tilde{v})\,\Gamma(d(x,v)\,d(\tilde{x},\tilde{v}))\,.$$

## 3 Proof of Theorem 1.12

In this section, for the sake of clarity, we only assume the potential $U$ satisfies Assumption 1.1 and Assumption 1.3. We refer to Section 5 for the adjustment of the proof in the case $\nabla U$ locally Lipschitz continuous.

Our goal is to prove the following result

**Theorem 3.1.** *Let $\mathcal{C}^0>0$. Let $U$ be continuously differentiable and satisfy Assumption 1.1 and Assumption 1.3. Let*

$$\tilde{\mathcal{C}}_K:=\mathcal{C}_1\left(1+\frac{2\epsilon B}{\gamma}+2\epsilon\mathcal{C}^0\right)+2\epsilon\left(\frac{B}{\gamma}+\mathcal{C}^0\right)\frac{6+8\lambda}{\lambda}.$$

*For all $W$ twice continuously differentiable satisfying Assumption 1.9 with $L_W < c/\tilde{\mathcal{C}}_K$, for all probability measures $\nu_0^1$ and $\nu_0^2$ on $\mathbb{R}^{2d}$ satisfying $\mathbb{E}_{\nu_0^1} H \leq \mathcal{C}^0$ and $\mathbb{E}_{\nu_0^2} H \leq \mathcal{C}^0$*

$$\forall t \geq 0, \qquad \mathcal{W}_\rho\left(\nu_t^1, \nu_t^2\right) \leq e^{-(c - L_W \tilde{\mathcal{C}}_K)t} \mathcal{W}_\rho\left(\nu_0^1, \nu_0^2\right),$$

*where $\nu_t^1$ (resp. $\nu_t^2$) is a solution of* (1.1) *with initial distribution $\nu_0^1$ (resp. $\nu_0^2$).*

### 3.1 Step one: coupling and evolution of the coupling semimetric

Let $\xi > 0$, and let $rc,\ sc : \mathbb{R}^{2d} \mapsto [0,1]$ be two Lipschitz continuous functions such that:

$$\begin{aligned}
&rc^2 + sc^2 = 1, \\
&rc\,(z,w) = 0 \text{ if } |z+w| \leq \frac{\xi}{2} \text{ or } \alpha|z| + |z+w| \geq R_1 + \xi, \\
&rc\,(z,w) = 1 \text{ if } |z+w| \geq \xi \text{ and } \alpha|z| + |z+w| \leq R_1.
\end{aligned}$$

These two functions translate into mathematical terms the regions in which we use a reflection coupling (represented by $rc = 1$) and the ones where we use a synchronous coupling (represented by $sc = 1$). Finally, $\xi$ is a parameter that will vanish to zero in the end. We therefore consider the following coupling:

$$\begin{cases}
dX_t = V_t dt \\
dV_t = -V_t dt - \nabla U\left(X_t\right) dt - \nabla W * \mu_t\left(X_t\right) dt + \sqrt{2} rc\left(Z_t, W_t\right) dB_t^{rc} \\
\qquad\quad + \sqrt{2} sc\left(Z_t, W_t\right) dB_t^{sc} \\
\mu_t = \text{Law}\left(X_t\right) \\
d\tilde{X}_t = \tilde{V}_t dt \\
d\tilde{V}_t = -\tilde{V}_t dt - \nabla U(\tilde{X}_t) dt - \nabla W * \tilde{\mu}_t(\tilde{X}_t) dt + \sqrt{2} rc\left(Z_t, W_t\right)\left(Id - 2 e_t e_t^T\right) dB_t^{rc} \\
\qquad\quad + \sqrt{2} sc\left(Z_t, W_t\right) dB_t^{sc} \\
\tilde{\mu}_t = \text{Law}(\tilde{X}_t),
\end{cases}$$

where $B^{rc}$ and $B^{sc}$ are independent Brownian motions, and

$$Z_t = X_t - \tilde{X}_t, \quad W_t = V_t - \tilde{V}_t, \quad Q_t = Z_t + W_t, \quad e_t = \begin{cases} \frac{Q_t}{|Q_t|} & \text{if } Q_t \neq 0, \\ 0 & \text{otherwise,} \end{cases}$$

and $e_t^T$ is the transpose of $e_t$. Then

$$\frac{dZ_t}{dt} = W_t = Q_t - Z_t. \tag{3.1}$$

So $\frac{d|Z_t|}{dt} = \frac{Z_t}{|Z_t|}\left(Q_t - Z_t\right)$ for every $t$ such that $Z_t \neq 0$, and $\frac{d|Z_t|}{dt} \leq |Q_t|$ for every $t$ such that $Z_t = 0$. In particular

$$\frac{d|Z_t|}{dt} \leq |Q_t| - |Z_t|.$$

We start by using Itô's formula to compute the evolution of $|Q_t|$. The following lemma is Lemma 7 of A. Durmus *et al.* [21] of which, for the sake of completeness, we give the proof.

**Lemma 3.2.** *Under Assumption 1.1, Assumption 1.3 and Assumption 1.9, we have almost surely for all $t \geq 0$.*

$$\begin{aligned}
d|Q_t| = &- e_t \cdot \left(\nabla U\left(X_t\right) - \nabla U(\tilde{X}_t)\right) dt - e_t \cdot \left(\nabla W * \mu_t\left(X_t\right) - \nabla W * \tilde{\mu}_t(\tilde{X}_t)\right) dt \\
&+ 2\sqrt{2} rc\left(Z_t, W_t\right) e_t \cdot dB_t^{rc}
\end{aligned} \tag{3.2}$$

*Proof.* Let $t \geq 0$. We begin by considering the dynamics of $Z_t$, $W_t$ and $Q_t$. We have

$$\begin{aligned}
dZ_t &= W_t dt \\
dW_t &= -W_t dt - \left(\nabla U\left(X_t\right) - \nabla U(\tilde{X}_t)\right) dt - \left(\nabla W * \mu_t\left(X_t\right) - \nabla W * \tilde{\mu}_t(\tilde{X}_t)\right) dt \\
&\quad + 2\sqrt{2} rc\left(Z_t, W_t\right) e_t e_t \cdot dB_t^{rc} \\
dQ_t &= -\left(\nabla U\left(X_t\right) - \nabla U(\tilde{X}_t)\right) dt - \left(\nabla W * \mu_t\left(X_t\right) - \nabla W * \tilde{\mu}_t(\tilde{X}_t)\right) dt \\
&\qquad + 2\sqrt{2} rc\left(Z_t, W_t\right) e_t e_t \cdot dB_t^{rc}.
\end{aligned}$$

Therefore

$$\begin{aligned}
d|Q_t|^2 &= -2 Q_t \cdot \left(\nabla U\left(X_t\right) - \nabla U(\tilde{X}_t)\right) dt - 2 Q_t \cdot \left(\nabla W * \mu_t\left(X_t\right) - \nabla W * \tilde{\mu}_t(\tilde{X}_t)\right) dt \\
&\quad + 4\sqrt{2} rc\left(Z_t, W_t\right)\left(Q_t \cdot e_t\right) e_t \cdot dB_t^{rc} + 8 rc^2\left(Z_t, W_t\right) dt.
\end{aligned}$$

We consider, for $\eta > 0$, the function $\psi_\eta(r) = (r+\eta)^{1/2}$ which is $\mathcal{C}^\infty$ on $]0, \infty[$ and satisfies

$$\begin{aligned}
&\forall r \geq 0, \quad \lim_{\eta \to 0} \psi_\eta(r) = r^{1/2}, \quad \lim_{\eta \to 0} 2\psi_\eta'(r) = r^{-1/2}, \quad \lim_{\eta \to 0} 4\psi_\eta''(r) = -r^{-3/2}, \\
&\qquad \text{and thus } \lim_{\eta \to 0} 2r\psi_\eta''(r) + \psi_\eta'(r) = 0.
\end{aligned}$$

Then

$$\begin{aligned}
d\psi_\eta\left(|Q_t|^2\right) =& -2\psi_\eta'\left(|Q_t|^2\right) Q_t \cdot \left(\nabla U\left(X_t\right) - \nabla U(\tilde{X}_t)\right) dt \\
&- 2\psi_\eta'\left(|Q_t|^2\right) Q_t \cdot \left(\nabla W * \mu_t\left(X_t\right) - \nabla W * \tilde{\mu}_t(\tilde{X}_t)\right) dt \\
&+ 4\psi_\eta'\left(|Q_t|^2\right) \sqrt{2} rc\left(Z_t, W_t\right)\left(Q_t \cdot e_t\right) e_t \cdot dB_t^{rc} + 8\psi_\eta'\left(|Q_t|^2\right) rc^2\left(Z_t, W_t\right) dt \\
&+ 16\psi_\eta''\left(|Q_t|^2\right) rc^2\left(Z_t, W_t\right) |Q_t|^2 dt.
\end{aligned}$$

We make sure each individual term converges almost surely as $\eta \to 0$. First, we notice that

$$2|Q_t|\psi_\eta'\left(|Q_t|^2\right) = \frac{|Q_t|}{\left(|Q_t|^2 + \eta\right)^{1/2}} \leq 1.$$

So

$$2\psi_\eta'\left(|Q_t|^2\right) Q_t \cdot \left(\nabla U\left(X_t\right) - \nabla U(\tilde{X}_t)\right) \leq |\nabla U\left(X_t\right) - \nabla U(\tilde{X}_t)| \leq L_U |Z_t|.$$

Then, by dominated convergence, for all $T \geq 0$ almost surely

$$\begin{aligned}
\lim_{\eta \to 0} \int_0^T 2\psi_\eta'\left(|Q_t|^2\right) Q_t \cdot \left(\nabla U\left(X_t\right) - \nabla U(\tilde{X}_t)\right) dt &= \int_0^T \frac{Q_t}{|Q_t|} \cdot \left(\nabla U\left(X_t\right) - \nabla U(\tilde{X}_t)\right) dt \\
&= \int_0^T e_t \cdot \left(\nabla U\left(X_t\right) - \nabla U(\tilde{X}_t)\right) dt.
\end{aligned}$$

Likewise for all $T \geq 0$

$$\begin{aligned}
2\psi_\eta'\left(|Q_t|^2\right) Q_t \cdot \left(\nabla W * \mu_t\left(X_t\right) - \nabla W * \tilde{\mu}_t(\tilde{X}_t)\right) &\leq \left|\nabla W * \mu_t\left(X_t\right) - \nabla W * \tilde{\mu}_t(\tilde{X}_t)\right| \\
&\leq L_W |Z_t| + L_W \mathbb{E}|Z_t|,
\end{aligned}$$

hence

$$\begin{aligned}
&\lim_{\eta \to 0} \int_0^T 2\psi_\eta'\left(|Q_t|^2\right) Q_t \cdot \left(\nabla W * \mu_t\left(X_t\right) - \nabla W * \tilde{\mu}_t(\tilde{X}_t)\right) dt \\
&\qquad = \int_0^T e_t \cdot \left(\nabla W * \mu_t\left(X_t\right) - \nabla W * \tilde{\mu}_t(\tilde{X}_t)\right) dt.
\end{aligned}$$

Then, since $rc\left(Z_t, W_t\right) = 0$ for $|Q_t| \leq \frac{\xi}{2}$ and

$$8\psi_\eta'\left(|Q_t|^2\right) + 16\psi_\eta''\left(|Q_t|^2\right)|Q_t|^2 = 4\left(\frac{1}{\left(|Q_t|^2+\eta\right)^{1/2}} - \frac{|Q_t|^2}{\left(|Q_t|^2+\eta\right)^{3/2}}\right)$$
$$= 4\frac{\eta}{\left(|Q_t|^2+\eta\right)^{3/2}} \leq \frac{4\eta}{|Q_t|^3},$$

we have by dominated convergence

$$\lim_{\eta\to 0}\int_0^T \left(8d\psi_\eta'\left(|Q_t|^2\right) rc^2\left(Z_t, W_t\right) + 16\psi_\eta''\left(|Q_t|^2\right) rc^2\left(Z_t, W_t\right)|Q_t|^2\right) dt = 0.$$

Finally, by Theorem 2.12 chapter 4 of [41]

$$\lim_{\eta\to 0}\int_0^T 4\sqrt{2}\psi_\eta'\left(|Q_t|^2\right) rc\left(Z_t, W_t\right)\left(Q_t\cdot e_t\right) e_t\cdot dB_t^{rc} = \int_0^T 2\sqrt{2}rc\left(Z_t, W_t\right) e_t\cdot dB^{rc}.$$

For any $t$, we obtain the desired result almost surely. The continuity of $t\mapsto |Q_t|$ then allows us to conclude that (3.2) is almost surely true for all $t$. $\square$

We denote

$$r_t := \alpha|X_t - \tilde{X}_t| + |X_t - \tilde{X}_t + V_t - \tilde{V}_t| = \alpha|Z_t| + |Q_t|, \tag{3.3}$$
$$\rho_t := f\left(r_t\right) G_t \text{ where } G_t = 1 + \epsilon H\left(X_t, V_t\right) + \epsilon H(\tilde{X}_t, \tilde{V}_t). \tag{3.4}$$

Since $H\left(x, v\right) \geq 0$ we have $G_t \geq 1$. We now state the main lemma of this section.

**Lemma 3.3.** *Under Assumption 1.1, Assumption 1.3 and Assumption 1.9, let $c \in ]0, \infty[$. Then almost surely for all $t \geq 0$*

$$\forall t \geq 0,\ e^{ct}\rho_t \leq \rho_0 + \int_0^t e^{cs}K_s ds + M_t, \tag{3.5}$$

*where $\left(M_t\right)_t$ is a continuous local martingale and*

$$K_t = 4f''\left(r_t\right) rc\left(Z_t, W_t\right)^2 G_t + cf\left(r_t\right) G_t + 96\epsilon \max\left(1, \frac{1}{2\alpha}\right) r_t f'\left(r_t\right) rc\left(Z_t, W_t\right)^2$$
$$+ \left(\alpha\frac{d|Z_t|}{dt} + \left(L_U + L_W\right)|Z_t| + L_W\mathbb{E}|Z_t|\right) f'\left(r_t\right) G_t$$
$$+ \epsilon\left(2B - \gamma H\left(X_t, V_t\right) - \gamma H(\tilde{X}_t, \tilde{V}_t)\right) f\left(r_t\right)$$
$$+ \epsilon L_W\left(6 + 8\lambda\right)\left(\mathbb{E}\left(|X_t|\right)^2 + \mathbb{E}(|\tilde{X}_t|)^2\right) f\left(r_t\right).$$

*Proof.* Using (3.2)

$$|Q_t| = |Q_0| + A_t^Q + M_t^Q \text{ with}$$
$$dA_t^Q = -e_t\cdot\left(\nabla U\left(X_t\right) - \nabla U(\tilde{X}_t)\right) dt - e_t\cdot\left(\nabla W * \mu_t\left(X_t\right) - \nabla W * \tilde{\mu}_t(\tilde{X}_t)\right) dt$$
$$dM_t^Q = 2\sqrt{2}rc\left(Z_t, W_t\right) e_t\cdot dB_t^{rc}.$$

Therefore $r_t = |Q_0| + \alpha|Z_t| + A_t^Q + M_t^Q$. Let $c > 0$. By Itô's formula

$$d\left(e^{ct}f\left(r_t\right)\right) = ce^{ct}f\left(r_t\right) dt + e^{ct}f'\left(r_t\right) dr_t + \frac{1}{2}e^{ct}f''\left(r_t\right) 8rc^2\left(Z_t, W_t\right) dt.$$

Hence

$$\begin{aligned}
e^{ct}f\left(r_t\right) =& f\left(r_0\right)+\hat{A}_t+\hat{M}_t \text{ with}\\
d\hat{A}_t =& \Big(cf\left(r_t\right)+\alpha f'\left(r_t\right)\frac{d|Z_t|}{dt}-f'\left(r_t\right)e_t\cdot\Big(\nabla U\left(X_t\right)-\nabla U(\tilde{X}_t)\Big)\\
&-f'\left(r_t\right)e_t\cdot\Big(\nabla W*\mu_t\left(X_t\right)-\nabla W*\tilde{\mu}_t(\tilde{X}_t)\Big)+4f''\left(r_t\right)rc^2\left(Z_t,W_t\right)\Big)e^{ct}dt\\
d\hat{M}_t =& e^{ct}2\sqrt{2}f'\left(r_t\right)rc\left(Z_t,W_t\right)e_t\cdot dB_t^{rc}.
\end{aligned}$$

We now consider the evolution of $G_t=1+\epsilon H\left(X_t,V_t\right)+\epsilon H(\tilde{X}_t,\tilde{V}_t)$

$$\begin{aligned}
dG_t =& \epsilon\left(\mathcal{L}_{\mu_t}H\left(X_t,V_t\right)+\mathcal{L}_{\tilde{\mu}_t}H(\tilde{X}_t,\tilde{V}_t)\right)dt\\
&+\epsilon\sqrt{2}rc\left(Z_t,W_t\right)\left(\nabla_vH\left(X_t,V_t\right)-\nabla_vH(\tilde{X}_t,\tilde{V}_t)\right)\cdot e_te_t^TdB_t^{rc}\\
&+\epsilon\sqrt{2}rc\left(Z_t,W_t\right)\left(\nabla_vH\left(X_t,V_t\right)+\nabla_vH(\tilde{X}_t,\tilde{V}_t)\right)\cdot\left(Id-e_te_t^T\right)dB_t^{rc}\\
&+\epsilon\sqrt{2}sc\left(Z_t,W_t\right)\left(\nabla_vH\left(X_t,V_t\right)+\nabla_vH(\tilde{X}_t,\tilde{V}_t)\right)\cdot dB_t^{sc}.
\end{aligned}$$

Therefore $e^{ct}\rho_t=e^{ct}f\left(r_t\right)G_t=\rho_0+A_t+M_t$, where

$$\begin{aligned}
dA_t =& G_td\hat{A}_t+\epsilon e^{ct}f\left(r_t\right)\left(\mathcal{L}_{\mu_t}H\left(X_t,V_t\right)+\mathcal{L}_{\tilde{\mu}_t}H(\tilde{X}_t,\tilde{V}_t)\right)dt\\
&+4\epsilon e^{ct}f'\left(r_t\right)rc^2\left(Z_t,W_t\right)\left(\nabla_vH\left(X_t,V_t\right)-\nabla_vH(\tilde{X}_t,\tilde{V}_t)\right)\cdot e_tdt,
\end{aligned}$$

and $M_t$ is a continuous local martingale. This last equality uses the fact that $B^{rc}$ and $B^{sc}$ are independent Brownian motion and that $e_t\cdot\left(Id-e_te_t^T\right)=0$. Furthermore

$$\begin{aligned}
|\nabla_vH\left(X_t,V_t\right)-\nabla_vH(\tilde{X}_t,\tilde{V}_t)| &= 12|X_t+2V_t-\tilde{X}_t-2\tilde{V}_t| = 12|2Q_t-Z_t|\\
&\le 24\left(\frac{1}{2}|Z_t|+|Q_t|\right)\\
&\le 24\max\left(1,\frac{1}{2\alpha}\right)r_t,
\end{aligned}$$

so that $dA_t\le e^{ct}\tilde{K}_tdt$, where

$$\begin{aligned}
\tilde{K}_t =& \Big(cf\left(r_t\right)+\alpha f'\left(r_t\right)\frac{d|Z_t|}{dt}-f'\left(r_t\right)e_t\cdot\Big(\nabla U\left(X_t\right)-\nabla U(\tilde{X}_t)\Big)\\
&-f'\left(r_t\right)e_t\cdot\Big(\nabla W*\mu_t\left(X_t\right)-\nabla W*\tilde{\mu}_t(\tilde{X}_t)\Big)+4f''\left(r_t\right)rc^2\left(Z_t,W_t\right)\Big)G_t\\
&+\epsilon\left(\mathcal{L}_tH\left(X_t,V_t\right)+\mathcal{L}_tH(\tilde{X}_t,\tilde{V}_t)\right)f\left(r_t\right)+96\epsilon\max\left(1,\frac{1}{2\alpha}\right)r_tf'\left(r_t\right)rc^2\left(Z_t,W_t\right).
\end{aligned}$$

And we conclude using Lemma 1.10, and Lemma 2.1. □

### 3.2 Step two : contractivity in various regions of space

At this point, we have

$$\forall t\ge 0,\ e^{ct}\rho_t\le\rho_0+\int_0^t e^{cs}K_sds+M_t,$$

where $M_t$ is a continuous local martingale and, by regrouping the terms according to how we will use them

$$K_t = \left(cf\left(r_t\right) + \left(\alpha\frac{d|Z_t|}{dt} + (L_U + L_W)|Z_t|\right)f'\left(r_t\right)\right)G_t \tag{3.6}$$

$$+ 4\left(f''\left(r_t\right)G_t + 24\epsilon\max\left(1, \frac{1}{2\alpha}\right)r_t f'\left(r_t\right)\right)rc\left(Z_t, W_t\right)^2 \tag{3.7}$$

$$+ \epsilon\left(2B - \gamma H\left(X_t, V_t\right) - \gamma H(\tilde{X}_t, \tilde{V}_t)\right)f\left(r_t\right) \tag{3.8}$$

$$+ L_W f'\left(r_t\right)\mathbb{E}\left(|Z_t|\right)G_t + \epsilon L_W\left(6 + 8\lambda\right)\left(\mathbb{E}\left(|X_t|\right)^2 + \mathbb{E}(|\tilde{X}_t|)^2\right)f\left(r_t\right). \tag{3.9}$$

Briefly,

- lines (3.6) and (3.7) will be non positive thanks to the construction of the function $f$ when using the reflection coupling,
- when only using the synchronous coupling, i.e when the deterministic drift is contracting, line (3.6) alone will be sufficiently small,
- line (3.8) translates the effect the Lyapunov function has in bringing back processes that would have ventured at infinity,
- finally, line (3.9) contains the non linearity and will be tackled by taking $L_W$ sufficiently small.

In this section, we thus prove the following lemma

**Lemma 3.4.** *Assume the confining potential $U$ satisfies Assumption 1.1 and Assumption 1.3. Then there is a constant $c^W > 0$ such that for all interaction potential $W$ satisfying Assumption 1.9 with $L_W < c^W$, the following holds for $K_t$ defined in* (3.6)-(3.9)

$$\mathbb{E}K_t \leq (1+\alpha)\,\xi\mathbb{E}G_t + L_W\left(\mathcal{C}_K + \mathcal{C}_K^0 e^{-\gamma t}\right)\mathbb{E}\rho_t,$$

*with*

$$\mathcal{C}_K = \mathcal{C}_1\left(1 + \frac{2\epsilon B}{\gamma}\right) + \frac{2\epsilon B}{\gamma\lambda}\left(6 + 8\lambda\right), \tag{3.10}$$

$$\mathcal{C}_K^0 = \epsilon\left(\mathcal{C}_1 + \frac{6+8\lambda}{\lambda}\right)\left(\mathbb{E}H\left(X_0, V_0\right) + \mathbb{E}H(\tilde{X}_0, \tilde{V}_0)\right).$$

*The constant $c^W$ is explicit, as it will be shown in Appendix B.*

To this end, we divide the space into three regions

$$\mathrm{Reg}_1 = \left\{(X_t, V_t, \tilde{X}_t, \tilde{V}_t) \text{ s.t. } |Q_t| \geq \xi \text{ and } r_t \leq R_1\right\},$$
$$\mathrm{Reg}_2 = \left\{(X_t, V_t, \tilde{X}_t, \tilde{V}_t) \text{ s.t. } |Q_t| < \xi \text{ and } r_t \leq R_1\right\},$$
$$\mathrm{Reg}_3 = \left\{(X_t, V_t, \tilde{X}_t, \tilde{V}_t) \text{ s.t. } r_t > R_1\right\},$$

and consider

$$\mathbb{E}K_t = \mathbb{E}(K_t\mathbb{1}_{\mathrm{Reg}_1}) + \mathbb{E}(K_t\mathbb{1}_{\mathrm{Reg}_2}) + \mathbb{E}(K_t\mathbb{1}_{\mathrm{Reg}_3}).$$

### 3.2.1 First region : $|Q_t| \geq \xi$ and $r_t \leq R_1$

In this region of space, we use the Brownian motion through the reflection coupling and the construction of the function $f$ to bring the processes closer together. Here we have $rc\left(Z_t, W_t\right) = 1$. Recall $\alpha|Z_t| + |Q_t| = r_t$ and $G_t \geq 1$.

- We have

$$\begin{aligned}\alpha\frac{d|Z_t|}{dt}+\left(L_U+L_W\right)|Z_t| &\leq \alpha|Q_t|-\alpha|Z_t|+\left(L_U+L_W\right)|Z_t| \\ &= \alpha r_t-\alpha^2|Z_t|-\alpha|Z_t|+\left(L_U+L_W\right)|Z_t| \\ &\leq \left(\frac{1}{\alpha}\left(L_U+L_W\right)+\alpha\right)r_t.\end{aligned}$$

- Since $G_t=1+\epsilon H\left(X_t,V_t\right)+\epsilon H(\tilde{X}_t,\tilde{V}_t)\geq 1$,

$$cG_t+\epsilon\left(2B-\gamma H\left(X_t,V_t\right)-\gamma H(\tilde{X}_t,\tilde{V}_t)\right)\leq cG_t+2\epsilon BG_t=\mathbf{C}G_t. \tag{3.11}$$

- We then have, by (2.13),

$$4\phi'\left(r_t\right)+\left(\frac{1}{\alpha}\left(L_U+L_W\right)+\alpha+96\epsilon\max\left(\frac{1}{2\alpha},1\right)\right)r_t\phi\left(r_t\right)=0.$$

  Hence

$$\begin{aligned}&4f''\left(r_t\right)+\left(\frac{1}{\alpha}\left(L_U+L_W\right)+\alpha+96\epsilon\max\left(\frac{1}{2\alpha},1\right)\right)r_tf'\left(r_t\right)\\ =&4\phi'\left(r_t\right)g\left(r_t\right)+4\phi\left(r_t\right)g'\left(r_t\right)\\ &+\left(\frac{1}{\alpha}\left(L_U+L_W\right)+\alpha+96\epsilon\max\left(\frac{1}{2\alpha},1\right)\right)r_t\phi\left(r_t\right)g\left(r_t\right)\\ =&4\phi\left(r_t\right)g'\left(r_t\right),\end{aligned}$$

  and

$$4\phi\left(r_t\right)g'\left(r_t\right)+\mathbf{C}f\left(r_t\right)\leq -4\frac{\mathbf{C}}{4}\Phi\left(r_t\right)+\mathbf{C}\Phi\left(r_t\right)=0.$$

- At this point, through this choice of function $f$, we are left with

$$K_t\mathbb{1}_{\mathrm{Reg}_1}\leq L_Wf'\left(r_t\right)\mathbb{E}\left(|Z_t|\right)G_t+\epsilon L_W\left(6+8\lambda\right)\left(\mathbb{E}\left(|X_t|\right)^2+\mathbb{E}(|\tilde{X}_t|)^2\right)f\left(r_t\right).$$

  Using Lemma 2.7, $f'\left(r_t\right)\leq 1$ and $G_t\geq 1$,

$$\mathbb{E}\left(K_t\mathbb{1}_{\mathrm{Reg}_1}\right)\leq L_W\mathcal{C}_1\mathbb{E}\left(\rho_t\right)\mathbb{E}\left(G_t\right)+\epsilon L_W\left(6+8\lambda\right)\left(\mathbb{E}\left(|X_t|\right)^2+\mathbb{E}(|\tilde{X}_t|)^2\right)\mathbb{E}\left(\rho_t\right).$$

  Recall (2.6)

$$\begin{aligned}\mathbb{E}\left(G_t\right)=&1+\epsilon\mathbb{E}H\left(X_t,V_t\right)+\epsilon\mathbb{E}H(\tilde{X}_t,\tilde{V}_t),\\ \leq&1+\frac{2\epsilon B}{\gamma}+\epsilon\left(\mathbb{E}H\left(X_0,V_0\right)+\mathbb{E}H(\tilde{X}_0,\tilde{V}_0)\right)e^{-\gamma t},\end{aligned}$$

  and, since $H(x,v)\geq\lambda|x|^2$,

$$\begin{aligned}\mathbb{E}\left(|X_t|\right)^2+\mathbb{E}(|\tilde{X}_t|)^2\leq&\frac{1}{\lambda}\mathbb{E}H\left(X_t,V_t\right)+\frac{1}{\lambda}\mathbb{E}H(\tilde{X}_t,\tilde{V}_t),\\ \leq&\frac{2B}{\gamma\lambda}+\frac{1}{\lambda}\left(\mathbb{E}H\left(X_0,V_0\right)+\mathbb{E}H(\tilde{X}_0,\tilde{V}_0)\right)e^{-\gamma t}.\end{aligned}$$

  Hence

$$\begin{aligned}\mathbb{E}\left(K_t\mathbb{1}_{\mathrm{Reg}_1}\right)\leq&L_W\left(\mathcal{C}_1\left(1+\frac{2\epsilon B}{\gamma}\right)+\frac{2\epsilon B}{\gamma\lambda}\left(6+8\lambda\right)\right)\mathbb{E}\left(\rho_t\right)\\ &+L_W\epsilon\left(\mathcal{C}_1+\frac{6+8\lambda}{\lambda}\right)\left(\mathbb{E}H\left(X_0,V_0\right)+\mathbb{E}H(\tilde{X}_0,\tilde{V}_0)\right)\mathbb{E}\left(\rho_t\right)e^{-\gamma t}.\end{aligned}$$

  We thus obtain $\mathbb{E}\left(K_t\mathbb{1}_{\mathrm{Reg}_1}\right)\leq L_W\left(\mathcal{C}_K+\mathcal{C}_K^0e^{-\gamma t}\right)\mathbb{E}\rho_t$.

### 3.2.2 Second region : $|Q_t| < \xi$ and $r_t \leq R_1$

In this region of space, we use the naturally contracting deterministic drift thanks to a synchronous coupling. Here $R_1 \geq r_t \geq \alpha |Z_t| \geq r_t - \xi$ so that

$$\begin{aligned} K_t \leq & \mathbf{C} f(r_t) G_t + \left( \alpha\xi - r_t + \xi + \frac{1}{\alpha}(L_U + L_W) r_t \right) f'(r_t) G_t \\ & + \left( 4f''(r_t) G_t + 96\epsilon \max\left(\frac{1}{2\alpha}, 1\right) r_t f'(r_t) \right) rc(Z_t, W_t)^2 \\ & + L_W f'(r_t) \mathbb{E}(|Z_t|) G_t + \epsilon L_W (6 + 8\lambda) \left( \mathbb{E}(|X_t|)^2 + \mathbb{E}(|\tilde{X}_t|)^2 \right) f(r_t), \end{aligned}$$

where we use (3.11). First

$$4f''(r_t) G_t + 96\epsilon \max\left(\frac{1}{2\alpha}, 1\right) r_t f'(r_t) \leq 0.$$

We use (2.10) to obtain, since $f(r_t) \leq \Phi(r_t)$ and $\frac{1}{2}\phi(r_t) \leq f'(r_t) = \phi(r_t) g(r_t) \leq \phi(r_t)$ by (2.12),

$$\begin{aligned} K_t \leq & \xi(1+\alpha)\phi(r_t) g(r_t) G_t + G_t \left( \mathbf{C}\Phi(r_t) + \frac{1}{2}\left( \frac{1}{\alpha}(L_U + L_W) - 1 \right) r_t \phi(r_t) \right) \\ & + L_W f'(r_t) \mathbb{E}(|Z_t|) G_t + \epsilon L_W (6 + 8\lambda) \left( \mathbb{E}(|X_t|)^2 + \mathbb{E}(|\tilde{X}_t|)^2 \right) f(r_t). \end{aligned}$$

Then, thanks to (2.11), like in the first region of space

$$\begin{aligned} K_t \mathbb{1}_{\text{Reg}_2} \leq & \xi(1+\alpha)\phi(r_t) g(r_t) G_t + L_W f'(r_t) \mathbb{E}(|Z_t|) G_t \\ & + \epsilon L_W (6 + 8\lambda) \left( \mathbb{E}(|X_t|)^2 + \mathbb{E}(|\tilde{X}_t|)^2 \right) f(r_t). \end{aligned}$$

Hence, since $\phi(r_t) g(r_t) \leq 1$

$$\mathbb{E} K_t \mathbb{1}_{\text{Reg}_2} \leq \xi(1+\alpha)\mathbb{E}(G_t) + L_W \left( \mathcal{C}_K + \mathcal{C}_K^0 e^{-\gamma t} \right) \mathbb{E}\rho_t.$$

### 3.2.3 Third region : $r_t > R_1$

In this region, we use the Lyapunov function. Here $f'(r_t) = f''(r_t) = 0$ so that

$$\begin{aligned} K_t \mathbb{1}_{\text{Reg}_3} = & \left( cG_t + \epsilon \left( 2B - \gamma H(X_t, V_t) - \gamma H(\tilde{X}_t, \tilde{V}_t) \right) \right) f(r_t) \mathbb{1}_{\text{Reg}_3} \\ & + \epsilon L_W (6 + 8\lambda) \left( \mathbb{E}(|X_t|)^2 + \mathbb{E}(|\tilde{X}_t|)^2 \right) f(r_t) \mathbb{1}_{\text{Reg}_3} \\ = & \Big[ \epsilon(c - \gamma) \left( H(X_t, V_t) + H(\tilde{X}_t, \tilde{V}_t) \right) + 2\epsilon B + c \\ & + \epsilon L_W (6 + 8\lambda) \left( \mathbb{E}(|X_t|)^2 + \mathbb{E}(|\tilde{X}_t|)^2 \right) \Big] f(r_t) \mathbb{1}_{\text{Reg}_3} \end{aligned}$$

Since $c - \gamma < 0$ as a consequence of (2.9), and using Lemma 2.3

$$\begin{aligned} K_t & \leq \left( (c-\gamma)\epsilon \frac{12}{5}\frac{B}{\gamma} + 2\epsilon B + c \right) f(r_t) + \epsilon L_W (6 + 8\lambda) \left( \mathbb{E}(|X_t|)^2 + \mathbb{E}(|\tilde{X}_t|)^2 \right) f(r_t) \\ & \leq \left( c\left( \frac{12\epsilon}{5}\frac{B}{\gamma} + 1 \right) - \frac{2}{5}\epsilon B \right) f(r_t) + \epsilon L_W (6 + 8\lambda) \left( \mathbb{E}(|X_t|)^2 + \mathbb{E}(|\tilde{X}_t|)^2 \right) f(r_t). \end{aligned}$$

Then, using (2.9), $\mathbb{E} K_t \mathbb{1}_{\text{Reg}_3} \leq L_W \mathcal{C}_K \mathbb{E}\rho_t + L_W \mathcal{C}_K^0 \mathbb{E}\rho_t e^{-\gamma t}$.

### 3.3 Step three: convergence

Let $\Gamma$ be a coupling of $\nu_0^1$ and $\nu_0^2$ such that $\mathbb{E}_\Gamma \rho\left((x,v),(\tilde{x},\tilde{v})\right)) < \infty$. We consider the coupling of $(X_t, V_t)$ and $(\tilde{X}_t, \tilde{V}_t)$, with initial distribution $((X_0,V_0),(\tilde{X}_0,\tilde{V}_0)) \sim \Gamma$, previously introduced. Using Lemma 3.3 and Lemma 3.4, by taking the expectation in (3.5) at stopping times $\tau_n$ increasingly converging to $t$, we have by Fatou's lemma for $n \to \infty$, $\forall \xi > 0, \forall t \geq 0$,

$$e^{ct}\mathbb{E}\rho_t \leq \mathbb{E}\rho_0 + (1+\alpha)\,\xi \int_0^t e^{cs}\mathbb{E}\left(G_s\right)ds + L_W \mathcal{C}_K \int_0^t e^{cs}\mathbb{E}\rho_s ds + L_W \mathcal{C}_K^0 \int_0^t e^{(c-\gamma)s}\mathbb{E}\rho_s ds. \tag{3.12}$$

Moreover, using Lemma 2.2 and the fact $\gamma > c$, for all $t \geqslant 0$,

$$\mathbb{E}\left(G_t\right) \leq \left(1 + \epsilon\mathcal{C}_H^0 + \epsilon\mathcal{C}_{\tilde{H}}^0\right), \qquad \int_0^t e^{(c-\gamma)s}\mathbb{E}\rho_s ds \leq \frac{f(R_1)\left(1 + \epsilon\mathcal{C}_H^0 + \epsilon\mathcal{C}_{\tilde{H}}^0\right)}{\gamma - c},$$
$$\int_0^t e^{cs} ds = \frac{c}{c - L_W\mathcal{C}_K}\frac{e^{ct}-1}{c} - \frac{L_W\mathcal{C}_K}{c - L_W\mathcal{C}_K}\int_0^t e^{cs}ds.$$

We get

$$\begin{aligned}
&e^{ct}\left(\mathbb{E}\rho_t - \frac{(1+\alpha)\,\xi}{c - L_W\mathcal{C}_K}\left(1 + \epsilon\mathcal{C}_H^0 + \epsilon\mathcal{C}_{\tilde{H}}^0\right)\right) \\
&\leq \mathbb{E}\rho_0 - \frac{(1+\alpha)\,\xi}{c - L_W\mathcal{C}_K}\left(1 + \epsilon\mathcal{C}_H^0 + \epsilon\mathcal{C}_{\tilde{H}}^0\right) + L_W\mathcal{C}_K^0\frac{f(R_1)\left(1 + \epsilon\mathcal{C}_H^0 + \epsilon\mathcal{C}_{\tilde{H}}^0\right)}{\gamma - c} \\
&\quad + L_W\mathcal{C}_K\int_0^t e^{cs}\left(\mathbb{E}\rho_s - \frac{(1+\alpha)\,\xi}{c - L_W\mathcal{C}_K}\left(1 + \epsilon\mathcal{C}_H^0 + \epsilon\mathcal{C}_{\tilde{H}}^0\right)\right)ds.
\end{aligned}$$

Gronwall's lemma yields, for all $t \geq 0$

$$\begin{aligned}
&e^{ct}\left(\mathbb{E}\left(\rho_t\right) - \frac{(1+\alpha)\,\xi}{c - L_W\mathcal{C}_K}\left(1 + \epsilon\mathcal{C}_H^0 + \epsilon\mathcal{C}_{\tilde{H}}^0\right)\right) \\
&\quad \leq \left(\mathbb{E}\left(\rho_0\right) + L_W\mathcal{C}_K^0\frac{f(R_1)\left(1 + \epsilon\mathcal{C}_H^0 + \epsilon\mathcal{C}_{\tilde{H}}^0\right)}{\gamma - c}\right)e^{L_W\mathcal{C}_K t}.
\end{aligned}$$

Since $\mathcal{W}_\rho\left(\mu_t, \nu_t\right) \leq \mathbb{E}\left(\rho_t\right)$, we have thus obtained for all $t \geq 0$

$$\begin{aligned}
\mathcal{W}_\rho\left(\nu_t^1, \nu_t^2\right) \leq &\frac{(1+\alpha)\,\xi}{c - L_W\mathcal{C}_K}\left(1 + \epsilon\mathcal{C}_H^0 + \epsilon\mathcal{C}_{\tilde{H}}^0\right) \\
&+ \left(\mathbb{E}\left(\rho_0\right) + L_W\mathcal{C}_K^0\frac{f(R_1)\left(1 + \epsilon\mathcal{C}_H^0 + \epsilon\mathcal{C}_{\tilde{H}}^0\right)}{\gamma - c}\right)e^{(L_W\mathcal{C}_K - c)t}
\end{aligned}$$

Taking the infimum over all couplings $\Gamma$ of the initial conditions and using the fact that the left hand side does not depend on $\xi$, so that we may take $\xi = 0$, we get finally that for all $t \geqslant 0$,

$$\mathcal{W}_\rho\left(\nu_t^1, \nu_t^2\right) \leq \left(\mathcal{W}_\rho\left(\nu_0^1, \nu_0^2\right) + L_W\mathcal{C}_K^0\frac{f(R_1)\left(1 + \epsilon\mathcal{C}_H^0 + \epsilon\mathcal{C}_{\tilde{H}}^0\right)}{\gamma - c}\right)e^{(L_W\mathcal{C}_K - c)t}, \tag{3.13}$$

and since, by Lemma 2.7, $\mathcal{C}_1\mathcal{W}_\rho\left(\nu_t^1, \nu_t^2\right) \geq \mathcal{W}_1\left(\nu_t^1, \nu_t^2\right)$ and $\mathcal{C}_2\mathcal{W}_\rho\left(\nu_t^1, \nu_t^2\right) \geq \mathcal{W}_2^2\left(\nu_t^1, \nu_t^2\right)$,

$$\begin{aligned}
\mathcal{W}_1\left(\nu_t^1, \nu_t^2\right) &\leq e^{-(c - L_W\mathcal{C}_K)t}C^1_{\nu_0^1,\nu_0^2}, \\
\mathcal{W}_2^2\left(\nu_t^1, \nu_t^2\right) &\leq e^{-(c - L_W\mathcal{C}_K)t}C^2_{\nu_0^1,\nu_0^2}.
\end{aligned}$$

Then, for all $W$ such that $L_W < c/\mathcal{C}_K$, there will be contraction at rate $\tau := c - L_W\mathcal{C}_K > 0$. So, it only remains for $L_W$ to satisfy

$$L_W \leq \frac{c}{\mathcal{C}_1\left(1+\frac{2\epsilon B}{\gamma}\right)+\frac{2\epsilon B}{\gamma\lambda}\left(6+8\lambda\right)}, \tag{3.14}$$

with

$$\mathcal{C}_1 = \max\left(\frac{2}{\alpha},1\right)\max\left(\frac{4\left(\left(1+\alpha\right)^2+\alpha^2\right)}{\epsilon\min\left(\frac{2}{3}\lambda,6\right)f\left(1\right)},\frac{1}{\phi\left(R_1\right)g\left(R_1\right)}\right).$$

**Remark 3.5.** We draw the reader's attention to the fact that Theorem 3.1 is then a consequence of everything we have done so far: if we have an upper bound on $\mathbb{E}H\left(X_0,V_0\right)+\mathbb{E}H(\tilde{X}_0,\tilde{V}_0)$, the constant $\mathcal{C}_K^0$ in Lemma 3.4 can be chosen equal to 0 provided we modify $\mathcal{C}_K$.

Let us now show that there is existence and uniqueness of a stationary measure. Let $\mathcal{C}^0 > \frac{B}{\gamma}$ and $\mu_t$ a solution of (1.1) such that $\mathbb{E}_{\mu_0}H \leq \mathcal{C}^0$. Using (2.6), for all $t \geq 0$, $\mathbb{E}_{\mu_t}H \leq \mathcal{C}^0$. Thanks to Theorem 3.1, for $L_W$ sufficiently small, there is $\tau > 0$ such that for all $t \geq s \geq 0$

$$\mathcal{W}_\rho\left(\mu_t,\mu_s\right) \leq e^{-\tau s}\mathcal{W}_\rho\left(\mu_{t-s},\mu_0\right) \leq f(R_1)\left(1+2\epsilon\mathcal{C}^0\right)e^{-\tau s},$$

and thus

$$\mathcal{W}_1\left(\mu_t,\mu_s\right) \leq \mathcal{C}_1 f(R_1)\left(1+2\epsilon\mathcal{C}^0\right)e^{-\tau s}.$$

The space of probability measure with first moments, equipped with the $\mathcal{W}_1$ distance, being a complete metric space (see for instance [6]), and $\mu_t$ being a Cauchy sequence, there exists $\mu_\infty$ such that

$$\mathcal{W}_1\left(\mu_t,\mu_\infty\right) \to 0 \text{ as } t\to\infty,$$

and $\mu_\infty$ stationary. Theorem 1.12 then ensures uniqueness and convergence towards this stationary measure.

# 4 Proof of Theorem 1.13

In this section, we show how we obtain similar results for the convergence of the particle system to the non-linear kinetic Langevin diffusion using the same tools. We start by introducing the coupling, the new Lyapunov function, we give a new definition for the various quantities we consider, and then prove contraction of the coupling semimetric.

## 4.1 Coupling

We consider the following coupling

$$\begin{cases} d\bar{X}_t^i = \bar{V}_t^i dt \\ d\bar{V}_t^i = -\bar{V}_t^i dt - \nabla U\left(\bar{X}_t^i\right)dt - \nabla W * \bar{\mu}_t\left(\bar{X}_t^i\right)dt \\ \qquad\qquad +\sqrt{2}\left(rc\left(Z_t^i,W_t^i\right)dB_t^{rc,i}+sc\left(Z_t^i,W_t^i\right)dB_t^{sc,i}\right) \\ \bar{\mu}_t = \mathcal{L}\left(\bar{X}_t^i\right) \\ dX_t^{i,N} = V_t^{i,N}dt \\ dV_t^{i,N} = -V_t^{i,N}dt - \nabla U(X_t^{i,N})dt - \frac{1}{N}\sum_{j=1}^N \nabla W(X_t^{i,N}-X_t^{j,N})dt \\ \qquad\qquad +\sqrt{2}\left(rc\left(Z_t^i,W_t^i\right)\left(Id-2e_t^ie_t^{i,T}\right)dB_t^{rc,i}+sc\left(Z_t^i,W_t^i\right)dB_t^{sc,i}\right), \end{cases}$$

with, similarly as before,

$$Z_t^i = \bar{X}_t^i - X_t^{i,N}, \qquad W_t^i = \bar{V}_t^i - V_t^{i,N}, \qquad Q_t^i = Z_t^i + W_t^i, \quad e_t^i = \begin{cases} \frac{Q_t^i}{|Q_t^i|} \text{ if } Q_t^i \neq 0, \\ 0 \text{ otherwise.} \end{cases}$$

Let $\mu_t^N := \frac{1}{N}\sum_{i=1}^N \delta_{X_t^{i,N}}$ be the empirical distribution of the particle system, with i.i.d initial conditions $X_0^{i,N} \sim \nu_0$. We first notice that the particles are exchangeable. The generator of the process given by the particle system (1.3) is, for a function $\phi$ of $(x_1, ..., x_N, v_1, ..., v_N)$

$$\mathcal{L}^N \phi = \sum_{i=1}^N \mathcal{L}^{i,N} \phi,$$

with

$$\mathcal{L}^{i,N}\phi = v_i \cdot \nabla_{x_i}\phi - v_i \cdot \nabla_{v_i}\phi - \nabla U(x_i) \cdot \nabla_{v_i}\phi - \frac{1}{N}\sum_{j=1}^N \nabla W(x_i - x_j) \cdot \nabla_{v_i}\phi + \Delta_{v_i}\phi.$$

We define

$$r_t^i = \alpha |Z_t^i| + |Q_t^i|, \tag{4.1}$$

$$\tilde{H}(x,v) = \int_0^{H(x,v)} \exp\left(a\sqrt{u}\right) du, \tag{4.2}$$

$$G_t^i = 1 + \epsilon \tilde{H}\left(\bar{X}_t^i, \bar{V}_t^i\right) + \epsilon \tilde{H}(X_t^{i,N}, V_t^{i,N}) + \frac{\epsilon}{N}\sum_{j=1}^N \tilde{H}\left(\bar{X}_t^j, \bar{V}_t^j\right) + \frac{\epsilon}{N}\sum_{j=1}^N \tilde{H}(X_t^{j,N}, V_t^{j,N}), \tag{4.3}$$

$$\rho_t = \frac{1}{N}\sum_{i=1}^N f\left(r_t^i\right) G_t^i. \tag{4.4}$$

There are two ideas when constructing this new $G_t^i$ compared to the previous section. First, we consider a modification of the Lyapunov function $\tilde{H}$, which we will describe in the next subsection. Second, we add these empirical means of the form $\frac{1}{N}\sum \tilde{H}$. This will allow us to deal with the non linearity appearing in the calculations. Recall the expectation in (3.9): this term will become an empirical mean, see (4.23) and (4.24) below. When taking the expectation of what we will denote $K_t^i$ (similar to $K_t$ given in Lemma 3.3), we no longer have a product of expectations, which we were able to deal with using the uniform in time bounds, but an expectation of the product. We will therefore have to control a quantity on the particle $i$ multiplied by a quantity on the particle $j$, and we do not have independence within the particle system. Hence the necessity, in the calculations, of adding this empirical mean of Lyapunov functions.

## 4.2 A modified Lyapunov function

Notice how in the expression of $G^i$ above we did not consider the Lyapunov function $H$, but instead $\tilde{H}$. Let us assume there exist $\mathcal{C}_0, a > 0$ such that $\mathbb{E}_{\nu_0}\left(\tilde{H}(X,V)^2\right) \leq (\mathcal{C}_0)^2$ (which is equivalent to the existence of $\tilde{\mathcal{C}^0}, \tilde{a} > 0$ such that $\mathbb{E}_{\nu_0}\left(e^{\tilde{a}(|X|+|V|)}\right) \leq \tilde{\mathcal{C}^0}$, as it was stated in Theorem 1.13). First, notice

$$\tilde{H}(x,v) = \int_0^{H(x,v)} \exp\left(a\sqrt{u}\right) du = \frac{2}{a^2}\exp\left(a\sqrt{H(x,v)}\right)\left(a\sqrt{H(x,v)} - 1\right) + \frac{2}{a^2}.$$

The idea of considering the exponential of the Lyapunov function is common when trying to obtain a greater restoring force, see for instance [35].

Here, for technical reasons (more precisely when dealing with the last term of $A_t^i$ given below in (4.18)) we have to ensure, when writing $\tilde{H} = f(H)$, that $f'$ is of order $e^{\sqrt{x}}$ instead of $e^x$.

Direct calculations yield the following technical lemma.

**Lemma 4.1.** *We have, for all* $x, v \in \mathbb{R}^d$

$$H(x,v)\exp\left(a\sqrt{H(x,v)}\right) \geq \tilde{H}(x,v) \geq \exp\left(a\sqrt{H(x,v)}\right) - \frac{2}{a^2}\left(\exp\left(\frac{a^2}{2}\right) - 1\right), \tag{4.5}$$

$$\frac{2}{a}\sqrt{H(x,v)}\exp\left(a\sqrt{H(x,v)}\right) \geq \tilde{H}(x,v) \geq \frac{1}{a}\sqrt{H(x,v)}\exp\left(a\sqrt{H(x,v)}\right) - \frac{1}{a^2}(e-2), \tag{4.6}$$

$$\tilde{H}(x,v) \geq H(x,v) \tag{4.7}$$

We may calculate, using (2.2) and (2.3)

$$\begin{aligned}
\mathcal{L}_\mu\left(\tilde{H}\right) &= \exp\left(a\sqrt{H}\right)\mathcal{L}_\mu H + \frac{a}{2\sqrt{H}}\exp\left(a\sqrt{H}\right)|\nabla_v H|^2 \\
&= \exp\left(a\sqrt{H}\right)\mathcal{L}_\mu H + 24^2\frac{a}{2\sqrt{H}}\exp\left(a\sqrt{H}\right)\left|\frac{x}{2}+v\right|^2 \\
&\leq \exp\left(a\sqrt{H}\right)\left(B + L_W(6+8\lambda)\mathbb{E}_\mu(|x|)^2 - \left(\frac{3}{4}\lambda+\lambda^2\right)|x|^2 - \gamma H\right) \qquad (4.8)\\
&\quad + 24a\sqrt{H}\exp\left(a\sqrt{H}\right) \\
&\leq \exp\left(a\sqrt{H}\right)\left(B + \frac{288a^2}{\gamma} + L_W(6+8\lambda)\mathbb{E}_\mu(|x|)^2 - \frac{\gamma}{2}H\right), \qquad (4.9)
\end{aligned}$$

where for this last inequality we used Young's inequality $24a\sqrt{H} \leq \frac{\gamma}{2}H + 288\frac{a^2}{\gamma}$.

Notice that (4.9) ensures that this new Lyapunov function also tends to bring back particle which ventured at infinity, and at an even greater rate. This new rate $H\exp(\sqrt{H})$ however comes at a cost: the initial condition must have a finite exponential moment, and not just a finite second moment as in Section 3.

First, by (2.6) and (4.7),

$$\mathbb{E}(|\bar{X}_t^i|)^2 \leq \frac{1}{\lambda}\mathbb{E}\left(H\left(\bar{X}_t^i,\bar{V}_t^i\right)\right) \leq \frac{1}{\lambda}\left(\frac{B}{\gamma} + \mathbb{E}H\left(\bar{X}_0^i,\bar{V}_0^i\right)\right) \leq \frac{1}{\lambda}\left(\frac{B}{\gamma} + \mathcal{C}^0\right).$$

Furthermore, the function $h \mapsto \exp\left(a\sqrt{h}\right)\left(\tilde{B} - \frac{\gamma}{4}h\right)$ is bounded from above for $h \geq 0$ and $\tilde{B} \in \mathbb{R}$. We therefore obtain from (4.9) the existence of $\tilde{B}$ such that

$$\mathcal{L}_{\bar{\mu}_t^{\otimes N}}\tilde{H}(x_i,v_i) \leq \tilde{B} - \frac{\gamma}{4}\left(H(x_i,v_i)\exp\left(a\sqrt{H(x_i,v_i)}\right)\right) \tag{4.10}$$

$$\frac{d}{dt}\mathbb{E}\tilde{H}\left(\bar{X}_t^i,\bar{V}_t^i\right) \leq \tilde{B} - \frac{\gamma}{4}\mathbb{E}\left(H\left(\bar{X}_t^i,\bar{V}_t^i\right)\exp\left(a\sqrt{H\left(\bar{X}_t^i,\bar{V}_t^i\right)}\right)\right) \tag{4.11}$$

$$\text{and}\quad \frac{d}{dt}\mathbb{E}\tilde{H}\left(\bar{X}_t^i,\bar{V}_t^i\right) \leq \tilde{B} - \frac{\gamma}{4}\mathbb{E}\tilde{H}\left(\bar{X}_t^i,\bar{V}_t^i\right), \tag{4.12}$$

where for this last inequality, we used (4.5). While (4.10) and (4.11) will be useful in ensuring a sufficient restoring force, Equation (4.12) give us a uniform in time bound on $\mathbb{E}\tilde{H}\left(\bar{X}_t^i,\bar{V}_t^i\right)$, provided we have an initial bound.

Now, for the system of particle, we have, using (4.9), $\forall i\in\{1,...,N\},\ \forall x_i,v_i\in\mathbb{R}^d$,

$$\begin{aligned}&\mathcal{L}^{i,N}\tilde{H}\left(x_i,v_i\right)\\&\quad\leq \exp\left(a\sqrt{H\left(x_i,v_i\right)}\right)\left(B+\frac{288a^2}{\gamma}+L_W\left(6+8\lambda\right)\left(\frac{\sum_{j=1}^N|x_j|}{N}\right)^2-\frac{\gamma}{2}H\left(x_i,v_i\right)\right).\end{aligned}$$

Summing over $i\in\{1,..,N\}$, we may calculate

$$\begin{aligned}L_W\left(6+8\lambda\right)&\sum_{j=1}^N\left(\frac{\sum_{j=1}^N|x_j|}{N}\right)^2\sum_{i=1}^N\frac{\exp\left(a\sqrt{H\left(x_i,v_i\right)}\right)}{N}\\&\qquad-\frac{\gamma}{8}\sum_{i=1}^N\frac{H\left(x_i,v_i\right)\exp\left(a\sqrt{H\left(x_i,v_i\right)}\right)}{N}\\&\leq\frac{\gamma}{8}\left(\sum_{i,j=1}^N\frac{H\left(x_i,v_i\right)}{N}\frac{\exp\left(a\sqrt{H\left(x_j,v_j\right)}\right)}{N}-\sum_{i=1}^N\frac{H\left(x_i,v_i\right)\exp\left(a\sqrt{H\left(x_i,v_i\right)}\right)}{N}\right)\\&\leq 0.\end{aligned}\tag{4.13}$$

Here, we used (2.2), the fact that $\forall x,y\geq 0\ \ xe^{a\sqrt{y}}+ye^{a\sqrt{x}}-xe^{a\sqrt{x}}-ye^{a\sqrt{y}}=(e^{a\sqrt{x}}-e^{a\sqrt{y}})(y-x)\leq 0$ and assumed

$$6\frac{L_W}{\lambda}\left(1+\frac{4}{3}\lambda\right)\leq\frac{\gamma}{8}\quad\text{i.e}\quad L_W\leq\frac{\gamma\lambda}{16(3+4\lambda)}.$$

Likewise, there is a constant, which for the sake of clarity we will also denote $\tilde{B}$ (as we may take the maximum of the previous constants), such that we get

$$\begin{aligned}\mathcal{L}^{i,N}\tilde{H}(x_i,v_i)\leq&\tilde{B}+L_W\left(6+8\lambda\right)\left(\frac{\sum_{j=1}^N|x_i|}{N}\right)^2\exp\left(a\sqrt{H\left(x_i,v_i\right)}\right)\\&-\frac{\gamma}{4}H\left(x_i,v_i\right)\exp\left(a\sqrt{H\left(x_i,v_i\right)}\right),\end{aligned}\tag{4.14}$$

$$\mathcal{L}^N\left(\frac{1}{N}\sum_{i=1}^N\tilde{H}(x_i,v_i)\right)\leq\tilde{B}-\frac{\gamma}{4}\left(\frac{1}{N}\sum_{i=1}^N H(x_i,v_i)\exp\left(a\sqrt{H\left(x_i,v_i\right)}\right)\right),\tag{4.15}$$

and

$$\mathcal{L}^N\left(\frac{1}{N}\sum_{i=1}^N\tilde{H}(x_i,v_i)\right)\leq\tilde{B}-\frac{\gamma}{4}\left(\frac{1}{N}\sum_{i=1}^N\tilde{H}(x_i,v_i)\right).\tag{4.16}$$

Once again, (4.14) and (4.15) will be ensure a sufficient restoring force, and (4.16) ensures a uniform in time bound on the expectation of $\tilde{H}(X_t^{i,N},V_t^{i,N})$, since $\mathbb{E}\left(\frac{1}{N}\sum_{j=1}^N\tilde{H}(X_t^{j,N},V_t^{j,N})\right)=\mathbb{E}\left(\tilde{H}(X_t^{i,N},V_t^{i,N})\right)$ by exchangeability of the particles.

More precisely, we obtain from (4.12) and (4.16) the direct corollary

**Lemma 4.2.** *Provided the initial expectations $\mathbb{E}\left(G_0^1\right)$ and $\mathbb{E}\left(\left(G_0^1\right)^2\right)$ are finite, there are constants $\mathcal{C}_{G,1}$ and $\mathcal{C}_{G,2}$, depending on initial conditions, such that for all $t\geq 0$, for all $N\geq 0$, and all $i$*

$$\mathbb{E}\left(G_t^i\right)\leq\mathcal{C}_{G,1}\quad\textit{and}\quad\mathbb{E}\left(\left(G_t^i\right)^2\right)\leq\mathcal{C}_{G,2}.$$

Finally, since $\tilde{H}(x,v)\geq H(x,v)$, Lemma 2.7 still holds for our new semimetric.

### 4.3 New parameters

For the sake of completeness, and since this is similar to Section 2.2, we quickly give some explicit parameters that satisfy the various conditions arising from calculation. T hese parameters are far from optimal, and are just given to show that every constant is explicit. Let $\tilde{B}$ be given by (4.10)-(4.12), and (4.14)-(4.16). Define

$$\alpha = L_U + \frac{\lambda}{4}, \quad R_0 = \sqrt{\frac{160\tilde{B}}{\gamma \min\left(\frac{\lambda}{3}, 3\right)}} \quad \text{and} \quad R_1 = \sqrt{(1+\alpha)^2+\alpha^2} R_0.$$

Recall the definition of $C_{dH,1}$ and $C_{dH,2}$ in (2.8). Denoting

$$C_{f,1} = 8\left(\left(\frac{96}{a^2}\max\left(1, \frac{1}{2\alpha}\right) + \frac{16\sqrt{3}}{a}C_{dH,1}\right)\left(\exp\left(\frac{a^2}{2}\right)-1\right) + 16\sqrt{3}(e-2)C_{dH,2}\right)$$
$$C_{f,2} = 8\left(24\max\left(1,\frac{1}{2\alpha}\right) + 4\sqrt{3}C_{dH,1}a + 8\sqrt{3}C_{dH,2}a^2\right)$$

we set

$$c = \left\{\frac{2\tilde{B}}{5}, \frac{\gamma}{800}, \frac{1}{12}\min\left(2\sqrt{\frac{L_U+L_W}{2\pi\alpha R_1^2}}, \frac{1}{2}\left(1-\frac{L_U+L_W}{\alpha}\right)\right)\right.$$
$$\left.\times \exp\left(-\frac{1}{8}\left(\frac{L_U+L_W}{\alpha} + \alpha + C_{f,1} + C_{f,2}\right)R_1^2\right)\right\},$$

and $\epsilon = \frac{5c}{2\tilde{B}}$. For $s \geq 0$,

$$\phi(s) = \exp\left(-\frac{1}{8}\left(\frac{1}{\alpha}(L_U+L_W) + \alpha + \epsilon C_{f,1} + C_{f,2}\right)s^2\right), \qquad \Phi(s) = \int_0^s \phi(u)\,du$$

$$g(s) = 1 - \frac{c+2\epsilon\tilde{B}}{2}\int_0^s \frac{\Phi(u)}{\phi(u)}du, \qquad f(s) = \int_0^{\min(s,R_1)} \phi(u)\,g(u)\,du.$$

This way we satisfy the following conditions

$$\begin{aligned}
c \leq& \frac{\gamma}{160}\left(1 - \frac{\gamma}{80\epsilon\tilde{B}+\gamma}\right)\\
\alpha >& L_U + L_W\\
\epsilon \leq& 1\\
2c + 4\epsilon\tilde{B} \leq& 2\left(\int_0^{R_1}\frac{\Phi(u)}{\phi(u)}du\right)^{-1}\\
2c + 4\epsilon\tilde{B} \leq& \frac{1}{2}\left(1-\frac{L_U+L_W}{\alpha}\right)\inf_{r\in]0,R_1]}\frac{r\phi(r)}{\Phi(r)}\\
\forall s\geq 0, 0 =& 4\phi'(s) + \left(\frac{1}{\alpha}(L_U+L_W) + \alpha + \epsilon C_{f,1} + C_{f,2}\right)s\phi(s)
\end{aligned}$$

### 4.4 Convergence

The goal of the section is to prove the following result

**Theorem 4.3.** *Let $U \in \mathcal{C}^1\left(\mathbb{R}^d\right)$ satisfy Assumption 1.1 and Assumption 1.3. For all $W \in \mathcal{C}^1\left(\mathbb{R}^d\right)$ satisfying Assumption 1.9 with*

$$L_W \leq \min\left(\frac{\gamma\lambda}{16(3+4\lambda)}, \frac{c}{\mathcal{C}_1}, \frac{\gamma}{64\mathcal{C}_z}, \frac{\gamma a}{256\mathcal{C}_z\epsilon}\right), \tag{4.17}$$

*and for all probability measures $\bar{\nu}_0$ on $\mathbb{R}^{2d}$ such that $\mathbb{E}_{\bar{\nu}_0}\tilde{H}^2(X,V) \leq (\mathcal{C}^0)^2$, for all $N$, $\xi > 0$, and $t \geq 0$,*

$$e^{ct}\mathbb{E}\left(\rho_t\right) \leq \mathbb{E}\left(\rho_0\right) + \xi\left(1+\alpha\right)\mathcal{C}_{G,1}\int_0^t e^{cs}ds + L_W\frac{\mathcal{C}^0\mathcal{C}_{G,2}^{1/2}}{\lambda}\sqrt{\frac{8}{N}}\int_0^t e^{cs}ds.$$

### 4.4.1 Proof of Theorem 1.13 using Theorem 4.3

We first show how Theorem 1.13 is a consequence of Theorem 4.3. Let $\Gamma$ be a coupling of $\nu_0^{\otimes N}$ and $\bar{\nu}_0^{\otimes N}$, such that $\mathbb{E}\rho_0 < \infty$. We consider the coupling previously introduced. For clarity, let us denote

$$A = L_W\frac{\mathcal{C}^0\mathcal{C}_{G,2}^{1/2}}{\lambda}\sqrt{8}, \quad B = (1+\alpha)\mathcal{C}_{G,1},$$

i.e

$$e^{ct}\mathbb{E}\left(\rho_t\right) \leq \mathbb{E}\left(\rho_0\right) + \xi B\int_0^t e^{cs}ds + \frac{A}{\sqrt{N}}\int_0^t e^{cs}ds.$$

Let us consider

$$u(t) = e^{ct}\left(\mathbb{E}\left(\rho_t\right) - \frac{A}{c}\frac{1}{\sqrt{N}} - \xi\frac{B}{c}\right)$$

Then $u(t) \leq u(0)$ i.e

$$\mathbb{E}\left(\rho_t\right) \leq \mathbb{E}\left(\rho_0\right)e^{-ct} + \frac{A}{c}\frac{1}{\sqrt{N}}\left(1-e^{-ct}\right) + \xi\frac{B}{c}\left(1-e^{-ct}\right).$$

We thus obtain the desired result, by taking the limit as $\xi \to 0$ uniformly in time, and by using the exchangeability of the particles to have $\mathbb{E}\left(\rho_t\right) = \mathbb{E}\left(\frac{1}{N}\sum_{i=1}^N \rho_t^i\right) = \mathbb{E}\left(\frac{1}{k}\sum_{i=1}^k \rho_t^i\right)$ for all $k \in \mathbb{N}$.

### 4.4.2 Evolution of the coupling semimetric for the particle system

We thus need to start by considering the dynamic of $\rho_t$. Like in Lemma 3.2, we have almost surely for all $t \geq 0$

$$\begin{aligned} d|Q_t^i| = &-e_t^{i,T}\left(\nabla U\left(\bar{X}_t^i\right) - \nabla U(X_t^{i,N})\right)dt - e_t^{i,T}\left(\nabla W * \bar{\mu}_t\left(\bar{X}_t^i\right) - \nabla W * \bar{\mu}_t^N(X_t^{i,N})\right)dt \\ &+ 2\sqrt{2}rc\left(Z_t^i, W_t^i\right)e_t^{i,T}dB_t^{rc,i}. \end{aligned}$$

Hence $e^{ct}f\left(r_t^i\right) = f\left(r_0\right) + \hat{A}_t^i + \hat{M}_t^i$ with

$$\begin{aligned} d\hat{A}_t^i = &\Bigg[cf\left(r_t^i\right) + \alpha f'\left(r_t^i\right)\frac{d|Z_t^i|}{dt} - f'\left(r_t^i\right)e_t^{i^T}\left(\nabla U\left(X_t^i\right) - \nabla U(X_t^{i,N})\right) \\ &- f'\left(r_t^i\right)e_t^{i^T}\left(\nabla W * \mu_t\left(X_t^i\right) - \frac{1}{N}\sum_{j=1}^N \nabla W(X_t^{i,N} - X_t^{j,N})\right) \\ &+ 4f''\left(r_t^i\right)rc^2\left(Z_t^i, W_t^i\right)\Bigg]e^{ct}dt, \\ d\hat{M}_t^i = &e^{ct}2\sqrt{2}f'\left(r_t^i\right)rc\left(Z_t^i, W_t^i\right)e_t^{i^T}dB_t^{rc,i}. \end{aligned}$$

We now consider the evolution of

$$G_t^i = 1 + \epsilon \tilde{H}\left(\bar{X}_t^i, \bar{V}_t^i\right) + \epsilon \tilde{H}(X_t^{i,N}, V_t^{i,N}) + \frac{\epsilon}{N}\sum_{j=1}^N \tilde{H}\left(\bar{X}_t^j, \bar{V}_t^j\right) + \frac{\epsilon}{N}\sum_{j=1}^N \tilde{H}(X_t^{j,N}, V_t^{j,N}).$$

Notice how we have added new terms in $G_t^i$. Those additional quantities will help us in dealing with the non linearity, as will be shown later.

$$\begin{aligned}
dG_t^i = &\ \epsilon\left(\mathcal{L}_{\bar{\mu}_t^{\otimes N}} \tilde{H}\left(\bar{X}_t^i, \bar{V}_t^i\right) + \mathcal{L}^N \tilde{H}(X_t^{i,N}, V_t^{i,N})\right) dt \\
&+ \epsilon\sqrt{2} rc\left(Z_t^i, W_t^i\right)\left(\nabla_v \tilde{H}\left(\bar{X}_t^i, \bar{V}_t^i\right) - \nabla_v \tilde{H}(X_t^{i,N}, V_t^{i,N})\right) \cdot e_t^i {e_t^i}^T dB_t^{rc,i} \\
&+ \epsilon\sqrt{2} rc\left(Z_t^i, W_t^i\right)\left(\nabla_v \tilde{H}\left(\bar{X}_t^i, \bar{V}_t^i\right) + \nabla_v \tilde{H}(X_t^{i,N}, V_t^{i,N})\right) \cdot \left(Id - e_t^i {e_t^i}^T\right) dB_t^{rc,i} \\
&+ \epsilon\sqrt{2} sc\left(Z_t^i, W_t^i\right)\left(\nabla_v \tilde{H}\left(\bar{X}_t^i, \bar{V}_t^i\right) + \nabla_v \tilde{H}(X_t^{i,N}, V_t^{i,N})\right) \cdot dB_t^{sc,i} \\
&+ \frac{\epsilon}{N}\sum_{j=1}^N \left(\mathcal{L}_{\bar{\mu}_t^{\otimes N}} \tilde{H}\left(\bar{X}_t^j, \bar{V}_t^j\right) + \mathcal{L}^N \tilde{H}(X_t^{j,N}, V_t^{j,N})\right) dt \\
&+ \frac{\epsilon\sqrt{2}}{N}\sum_{j=1}^N rc\left(Z_t^j, W_t^j\right)\left(\nabla_v \tilde{H}\left(\bar{X}_t^j, \bar{V}_t^j\right) - \nabla_v \tilde{H}(X_t^{j,N}, V_t^{j,N})\right) \cdot e_t^j {e_t^j}^T dB_t^{rc,j} \\
&+ \frac{\epsilon\sqrt{2}}{N}\sum_{j=1}^N rc\left(Z_t^j, W_t^j\right)\left(\nabla_v \tilde{H}\left(\bar{X}_t^j, \bar{V}_t^j\right) + \nabla_v \tilde{H}(X_t^{j,N}, V_t^{j,N})\right) \cdot \left(Id - e_t^j {e_t^j}^T\right) dB_t^{rc,j} \\
&+ \frac{\epsilon\sqrt{2}}{N}\sum_{j=1}^N sc\left(Z_t^j, W_t^j\right)\left(\nabla_v \tilde{H}\left(\bar{X}_t^j, \bar{V}_t^j\right) + \nabla_v \tilde{H}(X_t^{j,N}, V_t^{j,N})\right) \cdot dB_t^{sc,j}.
\end{aligned}$$

Therefore

$$e^{ct}\rho_t^i = e^{ct} f\left(r_t^i\right) G_t^i = \rho_0 + A_t^i + M_t^i, \tag{4.18}$$

with

$$\begin{aligned}
dA_t^i = &G_t^i d\hat{A}_t^i + \epsilon e^{ct} f\left(r_t^i\right)\left(\mathcal{L}_{\bar{\mu}_t^{\otimes N}} \tilde{H}\left(\bar{X}_t^i, \bar{V}_t^i\right) + \mathcal{L}^N \tilde{H}(X_t^{i,N}, V_t^{i,N})\right. \\
&\left. + \frac{1}{N}\sum_{j=1}^N \mathcal{L}_{\bar{\mu}_t^{\otimes N}} \tilde{H}\left(\bar{X}_t^j, \bar{V}_t^j\right) + \frac{1}{N}\mathcal{L}^N \sum_{j=1}^N \tilde{H}(X_t^{j,N}, V_t^{j,N})\right) dt \\
&+ 4\epsilon\left(1 + \frac{1}{N}\right) e^{ct} f'\left(r_t^i\right) rc^2\left(Z_t^i, W_t^i\right)\left(\nabla_v \tilde{H}\left(\bar{X}_t^i, \bar{V}_t^i\right) - \nabla_v \tilde{H}(X_t^{i,N}, V_t^{i,N})\right) \cdot e_t^i dt
\end{aligned}$$

and $M_t^i$ is a continuous local martingale. Let us deal with this last line. For the sake of conciseness, from now on we denote for all $i$

$$\bar{H}_i := H\left(\bar{X}_t^i, \bar{V}_t^i\right), \quad \text{and} \quad H_i^N := H(X_t^{i,N}, V_t^{i,N})$$

We have

$$
\begin{aligned}
&|\nabla_v \tilde{H}\left(\bar{X}_t^i, \bar{V}_t^i\right) - \nabla_v \tilde{H}(X_t^{i,N}, V_t^{i,N})| \\
&\quad = \left|\nabla_v \bar{H}_i \exp\left(a\sqrt{\bar{H}_i}\right) - \nabla_v H_i^N \exp\left(a\sqrt{H_i^N}\right)\right| \\
&\quad \leq \left|12\bar{X}_t^i + 24\bar{V}_t^i - 12X_t^{i,N} - 24V_t^{i,N}\right| \left(\exp\left(a\sqrt{\bar{H}_i}\right) + \exp\left(a\sqrt{H_i^N}\right)\right) \\
&\qquad + a\left|12\bar{X}_t^i + 24\bar{V}_t^i\right| \left|\sqrt{\bar{H}_i} - \sqrt{H_i^N}\right| \left(\exp\left(a\sqrt{\bar{H}_i}\right) + \exp\left(a\sqrt{H_i^N}\right)\right) \\
&\quad \leq 24\max\left(1, \frac{1}{2\alpha}\right) r_t^i \left(\exp\left(a\sqrt{\bar{H}_i}\right) + \exp\left(a\sqrt{H_i^N}\right)\right) \\
&\qquad + 4a\sqrt{3}\left|\bar{H}_i - H_i^N\right| \left(\exp\left(a\sqrt{\bar{H}_i}\right) + \exp\left(a\sqrt{H_i^N}\right)\right).
\end{aligned}
$$

Now, using Lemma 2.8, we get

$$
\begin{aligned}
&|\nabla_v \tilde{H}\left(\bar{X}_t^i, \bar{V}_t^i\right) - \nabla_v \tilde{H}(X_t^{i,N}, V_t^{i,N})| \\
&\quad \leq \left(24\max\left(1, \frac{1}{2\alpha}\right) + 4\sqrt{3}C_{dH,1}a\right) r_t^i \left(\exp\left(a\sqrt{\bar{H}_i}\right) + \exp\left(a\sqrt{H_i^N}\right)\right) \\
&\qquad + 4\sqrt{3}C_{dH,2} a r_t^i \left(\sqrt{\bar{H}_i} + \sqrt{H_i^N}\right) \left(\exp\left(a\sqrt{\bar{H}_i}\right) + \exp\left(a\sqrt{H_i^N}\right)\right) \\
&\quad \leq \left(24\max\left(1, \frac{1}{2\alpha}\right) + 4\sqrt{3}C_{dH,1}a\right) r_t^i \left(\exp\left(a\sqrt{\bar{H}_i}\right) + \exp\left(a\sqrt{H_i^N}\right)\right) \\
&\qquad + 8\sqrt{3}C_{dH,2} a r_t^i \left(\sqrt{\bar{H}_i}\exp\left(a\sqrt{\bar{H}_i}\right) + \sqrt{H_i^N}\exp\left(a\sqrt{H_i^N}\right)\right).
\end{aligned}
$$

Hence why, using (4.5) and (4.6), we get

$$
\begin{aligned}
&|\nabla_v \tilde{H}\left(\bar{X}_t^i, \bar{V}_t^i\right) - \nabla_v \tilde{H}(X_t^{i,N}, V_t^{i,N})| \\
&\quad \leq \left(24\max\left(1, \frac{1}{2\alpha}\right) + 4\sqrt{3}C_{dH,1}a\right) r_t^i \left(\frac{4}{a^2}\left(e^{\frac{a^2}{2}} - 1\right) + \tilde{H}\left(\bar{X}_t^i, \bar{V}_t^i\right) + \tilde{H}(X_t^{i,N}, V_t^{i,N})\right) \\
&\qquad + 8\sqrt{3}C_{dH,2} a^2 r_t^i \left(\frac{2}{a^2}(e-2) + \tilde{H}\left(\bar{X}_t^i, \bar{V}_t^i\right) + \tilde{H}(X_t^{i,N}, V_t^{i,N})\right),
\end{aligned}
$$

and thus

$$
\begin{aligned}
&4\epsilon\left(1 + \frac{1}{N}\right) e^{ct} f'\left(r_t^i\right) rc^2\left(Z_t^i, W_t^i\right) \left(\nabla_v \tilde{H}\left(\bar{X}_t^i, \bar{V}_t^i\right) - \nabla_v \tilde{H}(X_t^{i,N}, V_t^{i,N})\right) \cdot e_t^i dt \\
&\quad \leq 8\epsilon r_t^i f'\left(r_t^i\right) e^{ct} rc^2\left(Z_t^i, W_t^i\right) \\
&\qquad\qquad \times \left(\left(\frac{96}{a^2}\max\left(1, \frac{1}{2\alpha}\right) + \frac{16\sqrt{3}}{a}C_{dH,1}\right)\left(e^{\frac{a^2}{2}} - 1\right) + 16\sqrt{3}(e-2)C_{dH,2}\right) \\
&\qquad + 8 r_t^i f'\left(r_t^i\right) e^{ct} rc^2\left(Z_t^i, W_t^i\right) \left(24\max\left(1, \frac{1}{2\alpha}\right) + 4\sqrt{3}C_{dH,1}a + 8\sqrt{3}C_{dH,2}a^2\right) \\
&\qquad\qquad \times \left(\epsilon\tilde{H}\left(\bar{X}_t^i, \bar{V}_t^i\right) + \epsilon\tilde{H}(X_t^{i,N}, V_t^{i,N})\right) \\
&\quad \leq \left(\epsilon C_{f,1} + C_{f,2}\right) r_t^i f'\left(r_t^i\right) rc^2\left(Z_t^i, W_t^i\right) G_t^i.
\end{aligned}
$$

Then we use

$$\begin{aligned}
&\left|\nabla W * \bar{\mu}_t\left(\bar{X}_t^i\right)-\frac{1}{N} \sum_{j=1}^N \nabla W(X_t^{i,N}-X_t^{j,N})\right| \\
&\quad \leq\left|\nabla W * \bar{\mu}_t\left(\bar{X}_t^i\right)-\frac{1}{N} \sum_{j=1}^N \nabla W\left(\bar{X}_t^i-\bar{X}_t^j\right)\right| \\
&\qquad +\left|\frac{1}{N} \sum_{j=1}^N \nabla W\left(\bar{X}_t^i-\bar{X}_t^j\right)-\frac{1}{N} \sum_{j=1}^N \nabla W(X_t^{i,N}-X_t^{j,N})\right|, \\
&\quad \leq\left|\nabla W * \bar{\mu}_t\left(\bar{X}_t^i\right)-\frac{1}{N} \sum_{j=1}^N \nabla W\left(\bar{X}_t^i-\bar{X}_t^j\right)\right| \\
&\qquad +\frac{1}{N} \sum_{j=1}^N\left|\nabla W\left(\bar{X}_t^i-\bar{X}_t^j\right)-\nabla W(X_t^{i,N}-X_t^{j,N})\right|, \\
&\quad \leq\left|\nabla W * \bar{\mu}_t\left(\bar{X}_t^i\right)-\frac{1}{N} \sum_{j=1}^N \nabla W\left(\bar{X}_t^i-\bar{X}_t^j\right)\right|+\frac{L_W}{N} \sum_{j=1}^N\left(\left|\bar{X}_t^i-X_t^{i,N}\right|+\left|\bar{X}_t^j-X_t^{j,N}\right|\right).
\end{aligned}$$

Thus

$$\begin{aligned}
&\left|\nabla W * \bar{\mu}_t\left(\bar{X}_t^i\right)-\frac{1}{N} \sum_{j=1}^N \nabla W(X_t^{i,N}-X_t^{j,N})\right| \\
&\qquad\qquad \leq\left|\nabla W * \bar{\mu}_t\left(\bar{X}_t^i\right)-\frac{1}{N} \sum_{j=1}^N \nabla W\left(\bar{X}_t^i-\bar{X}_t^j\right)\right|+L_W|Z_t^i|+L_W \frac{\sum_{j=1}^N|Z_t^j|}{N}.
\end{aligned}$$

And finally we use (4.10), (4.14) and (4.15) to have

$$\begin{aligned}
&\mathcal{L}_{\bar{\mu}_t^{\otimes N}} \tilde{H}\left(\bar{X}_t^i, \bar{V}_t^i\right)+\mathcal{L}^N \tilde{H}(X_t^{i,N}, V_t^{i,N}) \\
&\qquad +\frac{1}{N} \sum_{j=1}^N \mathcal{L}_{\bar{\mu}_t^{\otimes N}} \tilde{H}\left(\bar{X}_t^j, \bar{V}_t^j\right)+\frac{1}{N} \mathcal{L}^N \sum_{j=1}^N \tilde{H}(X_t^{j,N}, V_t^{j,N}) \\
&\quad \leq 4 \tilde{B}+L_W(6+8 \lambda)\left(\frac{\sum_{j=1}^N|X_t^{j,N}|}{N}\right)^2 \exp \left(a \sqrt{H_i^N}\right) \\
&\qquad -\frac{\gamma}{4} \bar{H}_i \exp \left(a \sqrt{\bar{H}_i}\right)-\frac{\gamma}{4} H_i^N \exp \left(a \sqrt{H_i^N}\right) \\
&\qquad -\frac{\gamma}{4 N} \sum_{j=1}^N\left(\bar{H}_j \exp \left(a \sqrt{\bar{H}_j}\right)+H_j^N \exp \left(a \sqrt{H_j^N}\right)\right).
\end{aligned}$$

We thus obtain

$$dA_t^i \leq e^{ct} K_t^i dt \tag{4.19}$$

with

$$K_t^i = f'\left(r_t^i\right)G_t^i\left(\alpha\frac{d|Z_t^i|}{dt} + \left(L_U + L_W\right)|Z_t^i| + \left(\epsilon C_{f,1} + C_{f,2}\right)r_t^i rc^2\left(Z_t^i, W_t^i\right)\right) + 2cf\left(r_t^i\right)G_t^i \tag{4.20}$$

$$+ 4f''\left(r_t^i\right)G_t^i rc^2\left(Z_t^i, W_t^i\right) + \left|\nabla W * \bar\mu_t\left(\bar X_t^i\right) - \frac{1}{N}\sum_{j=1}^N \nabla W\left(\bar X_t^i - \bar X_t^j\right)\right| f'\left(r_t^i\right)G_t^i \tag{4.21}$$

$$+ \epsilon f\left(r_t^i\right)\Bigg(4\tilde B - \frac{\gamma}{16}\tilde H\left(\bar X_t^i, \bar V_t^i\right) - \frac{\gamma}{16}\tilde H(X_t^{i,N}, V_t^{i,N}) - \frac{\gamma}{16N}\sum_{j=1}^N \tilde H\left(\bar X_t^j, \bar V_t^j\right) - \frac{\gamma}{16N}\sum_{j=1}^N \tilde H(X_t^{j,N}, V_t^{j,N})\Bigg) \tag{4.22}$$

$$+ L_W\frac{\sum_{j=1}^N |Z_t^j|}{N} f'\left(r_t^i\right)G_t^i - cf\left(r_t^i\right)G_t^i - \epsilon f\left(r_t^i\right)\Bigg[\frac{\gamma}{16}\bar H_i \exp\left(a\sqrt{\bar H_i}\right) + \frac{\gamma}{16}H_i^N \exp\left(a\sqrt{H_i^N}\right) + \frac{\gamma}{16N}\sum_{j=1}^N \bar H_j \exp\left(a\sqrt{\bar H_j}\right) + \frac{\gamma}{16N}\sum_{j=1}^N H_j^N \exp\left(a\sqrt{H_j^N}\right)\Bigg] \tag{4.23}$$

$$+ \epsilon L_W\left(6 + 8\lambda\right)f\left(r_t^i\right)\left(\frac{\sum_{j=1}^N |X_t^{j,N}|}{N}\right)^2 \exp\left(a\sqrt{H_i^N}\right) - \frac{\gamma\epsilon}{8}f\left(r_t^i\right)\left(H_i^N \exp\left(a\sqrt{H_i^N}\right) + \frac{1}{N}\sum_{j=1}^N H_j^N \exp\left(a\sqrt{H_j^N}\right)\right). \tag{4.24}$$

This formulation of $K_t^i$ might seem cumbersome (and to some degree it is...) but we have actually grouped the various terms based on how we will have them compensate one another. Thus,

- lines (4.20) and (4.21) will be managed thanks to the construction of the function $f$ like before, with a special care given to the last term of line (4.21), on which we will use a law of large number,
- line (4.22) will come into play when considering the "last region of space" introduced previously,
- line (4.23) will, under some conditions on $L_W$, be nonpositive when summing up all $\left(K_t^j\right)_j$,
- and finally, line (4.24) will be nonpositive thanks to Lemma 2.1, provided $L_W$ is sufficiently small.

This highlights two important ideas in the construction of the function $\rho$: we both added in $G_t^i$ the empirical mean of $H(X_t^{i,N}, V_t^{i,N}) + H\left(\bar X_t^i, \bar V_t^i\right)$ and constructed a Lyapunov function with a greater restoring force. This is what allows us to tackle the non linearity appearing in (4.23) and (4.24) respectively in the terms $\frac{\sum_{j=1}^N |Z_t^j|}{N}G_t^i$ and $\left(\frac{\sum_{j=1}^N |X_t^{j,N}|}{N}\right)^2 \exp\left(a\sqrt{H_i^N}\right)$.

### 4.4.3 Some calculations

Like previously, we now have to show contraction in all three regions of space. Recall $f'\left(r_t^i\right) \leq 1$. The same calculations as before will be used, we only detail here the differences.

- First, since $\frac{L_W}{\lambda}\left(6+8\lambda\right) \leq \frac{\gamma}{8}$, by using Lemma 2.1 and since

$$H_j^N \exp\left(a\sqrt{H_i^N}\right) \leq H_i^N \exp\left(a\sqrt{H_i^N}\right) + H_j^N \exp\left(a\sqrt{H_j^N}\right)$$

we obtain

$$\epsilon L_W\left(6+8\lambda\right) f\left(r_t^i\right)\left(\frac{\sum_{j=1}^N |X_t^{j,N}|}{N}\right)^2 \exp\left(a\sqrt{H_i^N}\right) \\ - \frac{\gamma\epsilon}{8N} f\left(r_t^i\right)\left(N H_i^N \exp\left(a\sqrt{H_i^N}\right) + \sum_{j=1}^N H_j^N \exp\left(a\sqrt{H_j^N}\right)\right) \leq 0.$$

This takes care of (4.24).

- We have, since $f'\left(r_t^i\right) \leq 1$: $\frac{1}{N}\sum_{i=1}^N \frac{\sum_{j=1}^N |Z_t^j|}{N} f'\left(r_t^i\right) G_t^i \leq \frac{\sum_{i,j=1}^N |Z_t^j| G_t^i}{N^2}$. Then, using Lemma 2.7

$$\begin{aligned}
&\frac{1}{N^2}\sum_{i,j=1}^N |Z_t^i| G_t^j \\
=&\frac{1}{N}\sum_{i=1}^N |Z_t^i| + \frac{2\epsilon}{N^2}\sum_{i,j=1}^N |Z_t^i| \tilde{H}\left(\bar{X}_t^j, \bar{V}_t^j\right) + \frac{2\epsilon}{N^2}\sum_{i,j=1}^N |Z_t^i| \tilde{H}(X_t^{j,N}, V_t^{j,N}) \\
\leq&\frac{\mathcal{C}_1}{N}\sum_{i=1}^N \rho_t^i + \frac{2\mathcal{C}_z\epsilon}{N^2}\sum_{i,j=1}^N f(r_t^i)\left(\tilde{H}\left(\bar{X}_t^j, \bar{V}_t^j\right) + \tilde{H}(X_t^{j,N}, V_t^{j,N})\right) \\
&+ \frac{4\mathcal{C}_z\epsilon^2}{aN^2}\sum_{i,j=1}^N f(r_t^i)\left(\sqrt{\bar{H}_i} + \sqrt{H_i^N}\right)\left(\sqrt{\bar{H}_j}\exp\left(a\sqrt{\bar{H}_j}\right) + \sqrt{H_j^N}\exp\left(a\sqrt{H_j^N}\right)\right).
\end{aligned}$$

First, using (4.5)

$$\frac{2\mathcal{C}_z\epsilon}{N^2}\sum_{i,j=1}^N f(r_t^i)\left(\tilde{H}\left(\bar{X}_t^j, \bar{V}_t^j\right) + \tilde{H}(X_t^{j,N}, V_t^{j,N})\right) \\ \leq \frac{2\mathcal{C}_z\epsilon}{N^2}\sum_{i,j=1}^N f(r_t^i)\left(\bar{H}_j \exp\left(a\sqrt{\bar{H}_j}\right) + H_j^N \exp\left(a\sqrt{H_j^N}\right)\right).$$

Since

$$\left(\sqrt{\bar{H}_i} + \sqrt{H_i^N}\right)\left(\sqrt{\bar{H}_j}\exp\left(a\sqrt{\bar{H}_j}\right) + \sqrt{H_j^N}\exp\left(a\sqrt{H_j^N}\right)\right) \\ \leq 2\Big[\bar{H}_i \exp\left(a\sqrt{\bar{H}_i}\right) + \bar{H}_j \exp\left(a\sqrt{\bar{H}_j}\right) \\ + H_i^N \exp\left(a\sqrt{H_i^N}\right) + H_j^N \exp\left(a\sqrt{H_j^N}\right)\Big],$$

we have

$$\begin{aligned}
&\frac{4\mathcal{C}_z\epsilon^2}{aN^2}\sum_{i,j=1}^{N} f(r_t^i)\left(\sqrt{\bar{H}_i}+\sqrt{H_i^N}\right)\left(\sqrt{\bar{H}_j}\exp\left(a\sqrt{\bar{H}_j}\right)+\sqrt{H_j^N}\exp\left(a\sqrt{H_j^N}\right)\right)\\
&\qquad\le \frac{8\mathcal{C}_z\epsilon^2}{aN^2}\sum_{i,j=1}^{N} f(r_t^i)\Big[\bar{H}_i\exp\left(a\sqrt{\bar{H}_i}\right)+\bar{H}_j\exp\left(a\sqrt{\bar{H}_j}\right)\\
&\qquad\qquad\qquad\qquad + H_i^N\exp\left(a\sqrt{H_i^N}\right)+H_j^N\exp\left(a\sqrt{H_j^N}\right)\Big].
\end{aligned}$$

This way, since $2\mathcal{C}_z L_W \le \frac{\gamma}{32}$, $L_W\epsilon\frac{8\mathcal{C}_z}{a}\le\frac{\gamma}{32}$, and $L_W\mathcal{C}_1\le c$, we get

$$\begin{aligned}
&\frac{1}{N}\sum_{i=1}^{N}\Bigg(L_W\frac{\sum_{j=1}^{N}|Z_t^j|}{N}f'\left(r_t^i\right)G_t^i - cf\left(r_t^i\right)G_t^i\\
&\qquad -\epsilon f\left(r_t^i\right)\Bigg(\frac{\gamma}{16}\bar{H}_i\exp\left(a\sqrt{\bar{H}_i}\right)+\frac{\gamma}{16}H_i^N\exp\left(a\sqrt{H_i^N}\right)\\
&\qquad\qquad +\frac{\gamma}{16N}\sum_{j=1}^{N}\bar{H}_j\exp\left(a\sqrt{\bar{H}_j}\right)+\frac{\gamma}{16N}\sum_{j=1}^{N}H_j^N\exp\left(a\sqrt{H_j^N}\right)\Bigg)\Bigg)\le 0
\end{aligned}$$

- Using Cauchy-Schwarz inequality

$$\begin{aligned}
&\mathbb{E}\left(G_t^i\left|\nabla W*\bar{\mu}_t\left(\bar{X}_t^i\right)-\frac{1}{N}\sum_{j=1}^{N}\nabla W\left(\bar{X}_t^i-\bar{X}_t^j\right)\right|\right)\\
&\quad\le \mathbb{E}\left({G_t^i}^2\right)^{1/2}\mathbb{E}\left(\left|\nabla W*\bar{\mu}_t\left(\bar{X}_t^i\right)-\frac{1}{N}\sum_{j=1}^{N}\nabla W\left(\bar{X}_t^i-\bar{X}_t^j\right)\right|^2\right)^{1/2},\\
&\quad\le \mathbb{E}\left({G_t^i}^2\right)^{1/2}\mathbb{E}\left(\mathbb{E}\left(\left|\nabla W*\bar{\mu}_t\left(\bar{X}_t^i\right)-\frac{1}{N}\sum_{j=1}^{N}\nabla W\left(\bar{X}_t^i-\bar{X}_t^j\right)\right|^2\middle|\bar{X}_t^i\right)\right)^{1/2}.
\end{aligned}$$

Moreover, we notice that given $\bar{X}_t^i$, the random variables $\bar{X}_t^j$ for $j\neq i$ are i.i.d with law $\bar{\mu}_t$. Hence

$$\begin{aligned}
&\mathbb{E}\left(\left|\nabla W*\bar{\mu}_t\left(\bar{X}_t^i\right)-\frac{1}{N-1}\sum_{j=1}^{N}\nabla W\left(\bar{X}_t^i-\bar{X}_t^j\right)\right|^2\middle|\bar{X}_t^i\right)\\
&\qquad\qquad = \frac{1}{N-1}\mathrm{Var}_{\bar{\mu}_t}\left(\nabla W\left(\bar{X}_t^i-\cdot\right)\right)\le\frac{4L_W^2}{N-1}\mathbb{E}_{\bar{\mu}_t}\left(|\cdot|^2\right),
\end{aligned}$$

so

$$\begin{aligned}
&\mathbb{E}\left(\left|\nabla W*\bar{\mu}_t\left(\bar{X}_t^i\right)-\frac{1}{N}\sum_{j=1}^{N}\nabla W\left(\bar{X}_t^i-\bar{X}_t^j\right)\right|^2\right)\\
&\qquad\le\mathbb{E}\left(\left|\nabla W*\bar{\mu}_t\left(\bar{X}_t^i\right)-\frac{1}{N-1}\sum_{j=1}^{N}\nabla W\left(\bar{X}_t^i-\bar{X}_t^j\right)\right|^2\right)
\end{aligned}$$

$$
\begin{aligned}
&+\mathbb{E}\left(\left|\frac{1}{N}\sum_{j=1}^{N}\nabla W\left(\bar{X}_t^i-\bar{X}_t^j\right)-\frac{1}{N-1}\sum_{j=1}^{N}\nabla W\left(\bar{X}_t^i-\bar{X}_t^j\right)\right|^2\right),\\
&\leq\frac{4L_W^2}{N-1}\mathbb{E}_{\bar{\mu}_t}\left(|\cdot|^2\right)+\left(\frac{1}{N-1}-\frac{1}{N}\right)^2N\sum_{j=1}^{N}L_W^2\mathbb{E}\left(|\bar{X}_t^i-\bar{X}_t^j|^2\right),\\
&\leq 4L_W^2\left(\frac{1}{N-1}+\frac{1}{(N-1)^2}\right)\mathbb{E}_{\bar{\mu}_t}\left(|\cdot|^2\right).
\end{aligned}
$$

We may then use $\mathbb{E}_{\bar{\mu}_t}\left(|\cdot|^2\right)\leq\frac{\mathcal{C}^0}{\lambda}$.

Thus, by the same exact construction as before, we can obtain the existence of a function $f$ and a constant $c>0$ such that in all regions of space, for $L_W$ sufficiently small,

$$
\mathbb{E}\left(\frac{1}{N}\sum_i K_t^i\right)\leq\xi\left(1+\alpha\right)\mathcal{C}_{G,1}+L_W\frac{\mathcal{C}^0\mathcal{C}_{G,2}^{1/2}}{\lambda}\left(\frac{4}{N-1}+\frac{4}{(N-1)^2}\right)^{1/2}.
$$

By taking the expectation in the dynamic of $\rho_t$ given by (4.18) and (4.19) at stopping times $\tau_n$ increasingly converging to $t$, we prove Theorem 4.3 by using Fatou's lemma for $n\to\infty$.

## 5 $\nabla U$ locally Lipschitz continuous

As previously mentioned, the new Lyapunov function $\tilde{H}$ given in the previous section allows for a greater restoring force, recall (4.9). Let us now see how using this function allows for a perturbation of the global Lipschitz Assumption.

In this section we replace Assumption 1.3 with Assumption 1.4. We assume, for $\nu_0^1$ and $\nu_0^2$ the initial conditions,

$$
\forall i\in\{1,2\},\ \mathbb{E}_{\nu_0^i}\left(\left(\int_0^{H(X,V)}e^{a\sqrt{u}}du\right)^2\right)\leq(\mathcal{C}^0)^2 \tag{5.1}
$$

We show how the proof can be modified to still obtain contraction. As explained in Assumption 1.4, the coefficient $L_\psi$ will be considered sufficiently small with respect to the parameters of the problem. For now, let us simply assume $L_\psi$ is smaller than an *a priori* bound, for instance $L_\psi\leq 1$. Some conditions on $L_\psi$ will appear in the calculations below and we will deal with these later.

Like previously, we consider

$$
G_t=1+\epsilon\tilde{H}\left(X_t,V_t\right)+\epsilon\tilde{H}(\tilde{X}_t,\tilde{V}_t).
$$

Hence following the same method as previously we obtain

$$
\begin{aligned}
K_t\leq&G_t\Big(cf\left(r_t\right)+\alpha f'\left(r_t\right)\frac{d|Z_t|}{dt}+(L_U+L_W)f'\left(r_t\right)|Z_t|+L_Wf'\left(r_t\right)\mathbb{E}\left(|Z_t|\right) && (5.2)\\
&\quad+4f''\left(r_t\right)rc^2\left(Z_t,W_t\right)\Big)+\frac{1}{2}\left(\epsilon\mathcal{C}_{f,1}+\mathcal{C}_{f,2}\right)r_tf'\left(r_t\right)rc\left(Z_t,W_t\right)^2 && (5.3)\\
&+\epsilon\left(2\tilde{B}-\frac{\gamma}{8}\left(\tilde{H}\left(X_t,V_t\right)+\tilde{H}(\tilde{X}_t,\tilde{V}_t)\right)\right)f\left(r_t\right) && (5.4)\\
&+\left(\psi\left(X_t\right)+\psi(\tilde{X}_t)\right)|Z_t|f'\left(r_t\right)G_t\\
&\quad-\frac{\gamma\epsilon}{8}\left(H\left(X_t,V_t\right)\exp\left(a\sqrt{H\left(X_t,V_t\right)}\right)+H(\tilde{X}_t,\tilde{V}_t)\exp\left(a\sqrt{H(\tilde{X}_t,\tilde{V}_t)}\right)\right) && (5.5)
\end{aligned}
$$

$$+\epsilon\frac{L_W}{\lambda}\left(6+8\lambda\right)$$
$$\times\left(\exp\left(a\sqrt{H\left(X_t,V_t\right)}\right)\mathbb{E}H\left(X_t,V_t\right)+\exp\left(a\sqrt{H(\tilde{X}_t,\tilde{V}_t)}\right)\mathbb{E}H(\tilde{X}_t,\tilde{V}_t)\right)f\left(r_t\right). \tag{5.6}$$

We describe briefly how the terms will compensate each other before writing the calculations that are different.

- Like previously, lines (5.2) and (5.3) will be dealt with through the choice of function $f$, with the non linearity appearing at the end of (5.2) giving us a remaining expectation (cf bullet **1** below),
- line (5.4) will intervene like before in the last region of space (where we use that for all $x$ in $\mathbb{R}$, $\tilde{H}\geq H$ to come back to calculations we've made in Section 3.2.3, cf bullet **2** below) and in the first two region of space to compensate line (5.5) (cf bullet **3** below),
- and line (5.6) will give us a remaining expectation (cf bullet **4** below).

Notice how we use the Lyapunov function to compensate $\psi$ appearing when considering $\nabla U$ only locally Lipschitz continuous.

- **1**. We can find a constant $\mathcal{C}_{1,e}$ such that for all $x,v,\tilde{x},\tilde{v}\in\mathbb{R}^d$,

$$|x-\tilde{x}|+|v-\tilde{v}|\leq\mathcal{C}_{1,e}\rho(x,v,\tilde{x},\tilde{v}),$$

and thus

$$\mathbb{E}\left(\mathbb{E}\left(|Z_t|\right)G_tf'\left(r_t\right)\right)\leq\mathcal{C}_{1,e}\mathbb{E}\left(\rho_t\right)\mathbb{E}\left(G_t\right).$$

- **2**. In the last region of space, we use the fact that

$$K_t\mathbb{1}_{R_3}\leq\left(\left(c-\frac{\gamma}{8}\right)G_t+2\epsilon\tilde{B}+\frac{\gamma}{8}\right)f(r_t) \tag{5.7}$$

We deal with (5.7) exactly like in Section 3.2.3.

- **3**. Let us deal with the only locally Lipschitz continuous aspect. In the first two regions of space we use $f'(r_t)|Z_t|\leq f'(r_t)r_t/\alpha\leq f(r_t)/\alpha$ and the upper bound in (4.6).

$$\begin{aligned}
&G_t\left(\psi\left(X_t\right)+\psi(\tilde{X}_t)\right)f'\left(r_t\right)|Z_t|\\
&\quad-\epsilon\frac{\gamma}{8}\left(H\left(X_t,V_t\right)\exp\left(a\sqrt{H\left(X_t,V_t\right)}\right)+H(\tilde{X}_t,\tilde{V}_t)\exp\left(a\sqrt{H(\tilde{X}_t,\tilde{V}_t)}\right)\right)f\left(r_t\right)\\
&\leq\left(\psi\left(X_t\right)+\psi(\tilde{X}_t)\right)f'\left(r_t\right)|Z_t|\\
&\quad+\frac{2\epsilon}{a\alpha}f\left(r_t\right)\left(\psi\left(X_t\right)+\psi(\tilde{X}_t)\right)\\
&\qquad\times\left(\sqrt{H\left(X_t,V_t\right)}\exp\left(a\sqrt{H\left(X_t,V_t\right)}\right)+\sqrt{H(\tilde{X}_t,\tilde{V}_t)}\exp\left(a\sqrt{H(\tilde{X}_t,\tilde{V}_t)}\right)\right)\\
&\quad-\epsilon\frac{\gamma}{8}\left(H\left(X_t,V_t\right)\exp\left(a\sqrt{H\left(X_t,V_t\right)}\right)+H(\tilde{X}_t,\tilde{V}_t)\exp\left(a\sqrt{H(\tilde{X}_t,\tilde{V}_t)}\right)\right)f\left(r_t\right),
\end{aligned}$$

On one hand, since $\psi(x)\leq L_\psi\sqrt{H(x,v)}$, we have

$$\psi(x)+\psi(\tilde{x})\leq L_\psi\sqrt{H(x,v)}+L_\psi\sqrt{H(\tilde{x},\tilde{v})}\leq L_\psi\left(\frac{H(x,v)+H(\tilde{x},\tilde{v})}{2}+1\right),$$

and thus

$$\begin{aligned}&(\psi(x)+\psi(\tilde{x}))\, f'(r_t)|Z_t|\\&\quad\le \frac{L_\psi}{2}\left(H(x,v)\exp\left(a\sqrt{H(x,v)}\right)+H(\tilde{x},\tilde{v})\exp\left(a\sqrt{H(\tilde{x},\tilde{v})}\right)\right)\frac{f(r_t)}{\alpha}\\&\qquad+L_\psi f'(r_t)|Z_t|\end{aligned}$$

On the other hand

$$\begin{aligned}&(\psi(x)+\psi(\tilde{x}))\left(\sqrt{H(x,v)}\exp\left(a\sqrt{H(x,v)}\right)+\sqrt{H(\tilde{x},\tilde{v})}\exp\left(a\sqrt{H(\tilde{x},\tilde{v})}\right)\right)\\&\quad\le L_\psi\left(\sqrt{H(x,v)}+\sqrt{H(\tilde{x},\tilde{v})}\right)\\&\qquad\times\left(\sqrt{H(x,v)}\exp\left(a\sqrt{H(x,v)}\right)+\sqrt{H(\tilde{x},\tilde{v})}\exp\left(a\sqrt{H(\tilde{x},\tilde{v})}\right)\right)\\&\quad\le 2L_\psi\left(H(x,v)\exp\left(a\sqrt{H(x,v)}\right)+H(\tilde{x},\tilde{v})\exp\left(a\sqrt{H(\tilde{x},\tilde{v})}\right)\right).\end{aligned}$$

This way, assuming in a first step that

$$\frac{L_\psi}{2\alpha}\le\epsilon\frac{\gamma}{16}\quad\text{and}\quad\frac{4L_\psi\epsilon}{a\alpha}\le\epsilon\frac{\gamma}{16},\tag{5.8}$$

we get, since in the third region of space $f'(r_t)=0$,

$$\begin{aligned}&G_t\left(\psi(X_t)+\psi(\tilde{X}_t)\right)f'(r_t)|Z_t|\\&\qquad-\epsilon\frac{\gamma}{8}\left(H(X_t,V_t)\exp\left(a\sqrt{H(X_t,V_t)}\right)+H(\tilde{X}_t,\tilde{V}_t)\exp\left(a\sqrt{H(\tilde{X}_t,\tilde{V}_t)}\right)\right)f(r_t)\\&\quad\le L_\psi|Z_t|f'(r_t)G_t.\end{aligned}$$

At this stage, lines (5.2), (5.3) and (5.5) (without the non linearity dealt with in bullet **1**), can be bounded by the quantity

$$\begin{aligned}\tilde{K}_t=&G_t\Big(cf(r_t)+\alpha f'(r_t)\frac{d|Z_t|}{dt}+(L_U+1+\frac{\lambda}{8})f'(r_t)|Z_t|+4f''(r_t)rc^2(Z_t,W_t)\Big)\\&+\frac{1}{2}(\epsilon\mathcal{C}_{f,1}+\mathcal{C}_{f,2})r_tf'(r_t)rc(Z_t,W_t)^2,\end{aligned}$$

where we used the *a priori* bounds $0\le L_\psi\le 1$ and $0\le L_W<\frac{\lambda}{8}$. The righthand side is then dealt with through the choice of the concave function $f$ like previously.

- **4**. Likewise, we can bound

$$\begin{aligned}&\mathbb{E}\left(\epsilon\left(\exp\left(a\sqrt{H(X_t,V_t)}\right)\mathbb{E}H(X_t,V_t)+\exp\left(a\sqrt{H(\tilde{X}_t,\tilde{V}_t)}\right)\mathbb{E}H(\tilde{X}_t,\tilde{V}_t)\right)f(r_t)\right)\\&\quad\le\mathcal{C}_{H,\tilde{H}}\mathbb{E}(\rho_t)+\mathcal{C}^0_{H,\tilde{H}}\mathbb{E}(\rho_t)e^{-\gamma t},\end{aligned}$$

with $\mathcal{C}_{H,\tilde{H}}$ a constant independent of initial conditions and $\mathcal{C}^0_{H,\tilde{H}}$ another constant, possibly depending on initial conditions. Here, we used (2.6) and (4.5).

We can thus construct a function $f$ and constants $c$ and $\epsilon$, through the same calculations as before, such that there are $C$ and $C^0$ constants (resp. independent and dependent on initial conditions) such that

$$\forall t,\quad e^{ct}\mathbb{E}(\rho_t)\le\mathbb{E}(\rho_0)+\xi(1+\alpha)\mathbb{E}(G_t)e^{ct}+L_WC\int_0^t e^{cs}\mathbb{E}(\rho_s)\,ds$$

$$+ L_W C^0 \int_0^t e^{(c-\gamma\lambda)s} \mathbb{E}(\rho_s)\, ds.$$

Since $\mathbb{E}G_t$ is bounded uniformly in time, we may now conclude using Gronwall's lemma.

We have used in the proof the assumption (5.8) on $L_\psi$. Let us explain why it can be enforced. Here, the parameter $\epsilon$ is independent of $L_\psi$ (as above we have bounded it using the *a priori* bounds $0 \leq L_\psi \leq 1$) and is similar to the expression of $\epsilon$ given in Section 4.3. Using the fact that $\alpha > L_U$, we assume

$$L_\psi \leq c_\psi(L_U, \lambda, \tilde{A}, d, a) := \min\left(\frac{L_U \gamma a}{64}, \frac{L_U \gamma \epsilon}{8}, 1\right) \tag{5.9}$$

with $\gamma = \frac{\lambda}{2(\lambda+1)}$.

# A Various results

## A.1 Proof of lemma 1.2

The property only depends on the distance to the the origin, not the direction. We therefore only need to prove it in dimension 1, making sure the constant $\tilde{A}$ is independent of the direction. There is $x_0 > 0$ such that $\frac{\lambda}{2}x_0^2 = 2A$. Therefore, for $x \geq 0$, using (1.4):

$$U'(x_0+x)(x_0+x) \geq 2\lambda U(x_0+x) + \frac{\lambda}{2}(x_0+x)^2 - 2A = 2\lambda U(x_0+x) + \frac{\lambda}{2}x^2 + \lambda x x_0.$$

Then, for $x \geq 0$:

$$\begin{aligned}
U(x_0+x) - U(x_0) &= \int_0^1 U'(x_0+tx)\, x dt = \int_0^1 U'(x_0+tx)(x_0+tx)\frac{x}{x_0+tx} dt \\
&\geq \frac{x}{x_0+x}\int_0^1 2\lambda U(x_0+tx) + \frac{\lambda}{2}t^2x^2 + \lambda t x x_0 dt \\
&\geq \frac{x}{x_0+x}\left(\frac{\lambda}{6}x^2 + \frac{\lambda}{2}x x_0\right) \qquad \text{since } U \geq 0 \\
&= \frac{\lambda}{6}\frac{x^3}{x_0+x} + \frac{\lambda}{2}\frac{x^2 x_0}{x_0+x}.
\end{aligned}$$

We thus have for all $x \geq x_0$:

$$\begin{aligned}
U(x) - U(x_0) &\geq \frac{\lambda}{6}\frac{(x-x_0)^3}{x} + \frac{\lambda}{2}\frac{(x-x_0)^2 x_0}{x} \\
&= \frac{\lambda}{6}x^2 - \frac{\lambda}{2}x x_0 + \frac{\lambda}{2}x_0^2 - \frac{\lambda}{6}\frac{x_0^3}{x} + \frac{\lambda}{2}x x_0 - \lambda x_0^2 + \frac{\lambda}{2}\frac{x_0^3}{x} \\
&= \frac{\lambda}{6}x^2 - \frac{\lambda}{2}x_0^2 + \frac{\lambda}{3}\frac{x_0^3}{x}.
\end{aligned}$$

However, $-\frac{\lambda}{2}x_0^2 + \frac{\lambda}{3}\frac{x_0^3}{x} \geq -\frac{\lambda}{2}x_0^2 = -2A$ for $x \geq x_0$. We therefore have the desired result for $x \geq x_0$. The same reasoning gives us the result for $x \leq -x_0$.

Hence, if $|x| \geq |x_0| = \sqrt{\frac{4A}{\lambda}}$, $U(x) - U(x_0) \geq \frac{\lambda}{6}x^2 - 2A$. We then use the fact that $U(x)$ is continuous on the sphere of center 0 and radius $\sqrt{\frac{4A}{\lambda}}$, hence bounded on this set, to give a lower bound on $U(x_0)$ independent of the direction. Finally, for $|x| \in [-x_0, x_0]$, the function $x \mapsto U(x) - \frac{\lambda}{6}x^2$ is continuous, therefore bounded.

### A.2 Proof of lemma 1.10

We have

$$\nabla W * \mu(x) - \nabla W * \nu(\tilde{x}) = \nabla W * \mu(x) - \nabla W * \mu(\tilde{x}) + \nabla W * \mu(\tilde{x}) - \nabla W * \nu(\tilde{x})$$

Let $(X, \tilde{X})$ be a coupling of $\mu$ and $\nu$. Then

$$\begin{aligned}|\nabla W * \mu_t(x) - \nabla W * \tilde{\mu}_t(\tilde{x})| =& \left|\mathbb{E}\left(\nabla W(x-X) - \nabla W(\tilde{x}-\tilde{X})\right)\right| \\ \leq& L_W \mathbb{E}\left(|x - X - \tilde{x} + \tilde{X}|\right) \\ \leq& L_W \mathbb{E}\left(|x-\tilde{x}| + |X-\tilde{X}|\right)\end{aligned}$$

This being true for all coupling, we obtain the desired result.

### A.3 Proof of Lemma 2.1

**Remark A.1.** With $\gamma$ given by (2.1), we have $\gamma \leq \frac{1}{2}$.

We have

$$\begin{aligned}\mathcal{L}_\mu H(x,v) =& v \cdot \nabla_x H(x,v) - v \cdot \nabla_v H(x,v) - \nabla U(x) \cdot \nabla_v H(x,v) \\ &- \nabla W * \mu(x) \cdot \nabla_v H(x,v) + \Delta_v H(x,v) \\ =& v \cdot (24\nabla U(x) + 12(1-\gamma)x + 2\lambda x + 12v) - v \cdot (12x + 24v) \\ &- \nabla U(x) \cdot (12x + 24v) - \nabla W * \mu(x) \cdot (12x + 24v) + 24d \\ =& 24d - 12\nabla U(x) \cdot x + x \cdot v(12(1-\gamma) + 2\lambda - 12) \\ &- \nabla W * \mu(x) \cdot (12x + 24v) - 12|v|^2,\end{aligned}$$

with

$$\begin{aligned}-\gamma H(x,v) =& -24\gamma U(x) - 6\gamma(1-\gamma)|x|^2 - \gamma\lambda|x|^2 - 12\gamma x \cdot v - 12\gamma|v|^2 \\ -12\nabla U(x) \cdot x \leq& -24\lambda U(x) - 6\lambda|x|^2 + 24A,\end{aligned}$$

and

$$\begin{aligned}&-\nabla W * \mu(x) \cdot (12x + 24v) \\ &\quad \leq (L_W|x| + L_W \mathbb{E}_\mu(|\cdot|))(12|x| + 24|v|) \\ &\quad \leq 12L_W|x|^2 + 24L_W|x||v| \\ &\qquad + L_W \mathbb{E}_\mu(|\cdot|)\left(6\frac{|x|^2}{a_x \mathbb{E}_\mu(|\cdot|)} + 6a_x \mathbb{E}_\mu(|\cdot|) + 12\frac{|v|^2}{a_v \mathbb{E}_\mu(|\cdot|)} + 12a_v \mathbb{E}_\mu(|\cdot|)\right),\end{aligned}$$

where this last inequality holds for any $a_x, a_v > 0$. Therefore

$$\begin{aligned}\mathcal{L}_\mu H(x,v) \leq& 24A + 24d + 6L_W \mathbb{E}_\mu(|\cdot|)^2(a_x + 2a_v) - \gamma H(x,v) + 24\gamma U(x) + 6\gamma(1-\gamma)|x|^2 \\ &+ \gamma\lambda|x|^2 + 12\gamma x \cdot v + 12\gamma|v|^2 - 24\lambda U(x) - 6\lambda|x|^2 + 12L_W|x|^2 + 24L_W|x||v| \\ &+ \frac{6L_W}{a_x}|x|^2 + \frac{12L_W}{a_v}|v|^2 + x \cdot v(12(1-\gamma) + 2\lambda - 12) - 12|v|^2,\end{aligned}$$

and then

$$\begin{aligned}\mathcal{L}_\mu H(x,v) \leq& 24A + 24d + 6L_W \mathbb{E}_\mu(|\cdot|)^2(a_x + 2a_v) - \gamma H(x,v) + 24U(x)(\gamma - \lambda) \\ &+ |x||v|(|12\gamma + 12(1-\gamma) + 2\lambda - 12| + 24L_W)\end{aligned}$$

$$+ |x|^2 \left( 6\gamma (1-\gamma) + \gamma\lambda - 6\lambda + 12L_W + \frac{6L_W}{a_x} \right)$$
$$+ |v|^2 \left( 12\gamma - 12 + \frac{12L_W}{a_v} \right).$$

We now use $|x||v| \le \frac{\lambda}{3}|x|^2 + \frac{3}{4\lambda}|v|^2$, and $|12\gamma\lambda + 12(1-\gamma\lambda) + 2\lambda - 12| = 2\lambda$.

We have $(\gamma - \lambda) < 0$. Hence $24U(x)(\gamma-\lambda) \le 4\lambda(\gamma-\lambda)|x|^2 - 24(\gamma-\lambda)\tilde{A}$ using Lemma 1.2. Then

$$\mathcal{L}_\mu H(x,v) \le 24A - 24(\gamma-\lambda)\tilde{A} + 24d + 6L_W \mathbb{E}_\mu(|\cdot|)^2 (a_x + 2a_v) - \gamma H(x,v)$$
$$+ |x|^2 \left( 4\lambda(\gamma-\lambda) + 6\gamma(1-\gamma) + \gamma\lambda - 6\lambda + \frac{6L_W}{a_x} + 12L_W + \frac{2\lambda^2}{3} + \frac{24L_W\lambda}{3} \right)$$
$$+ |v|^2 \left( 12\gamma - 12 + \frac{12L_W}{a_v} + \frac{3}{2} + \frac{3}{4\lambda}24L_W \right).$$

We now consider each term individually.

**Coefficient of $|x|^2$.** We have, using $0 < \gamma < 1$ and $L_W \le \frac{\lambda}{8}$

$$4\lambda(\gamma-\lambda) + 6\gamma(1-\gamma) + \gamma\lambda - 6\lambda + \frac{2\lambda^2}{3} + \frac{24L_W\lambda}{3} + 12L_W$$
$$\le \gamma(5\lambda+6) - \left( 4\lambda^2 + 6\lambda - \frac{2\lambda^2}{3} - \lambda^2 - \frac{3\lambda}{2} \right).$$

Therefore, it is sufficient that

$$\gamma \le \lambda \frac{\frac{7}{3}\lambda + \frac{9}{2}}{5\lambda+6}.$$

We check this holds for $\gamma = \frac{\lambda}{2\lambda+2}$. Then

$$4\lambda(\gamma-\lambda) + 6\gamma(1-\gamma) + \gamma\lambda - 6\lambda + \frac{2\lambda^2}{3} + \frac{24L_W\lambda}{3} + 12L_W + \frac{6L_W}{a_x}$$
$$\le (5\lambda+6)\left( \gamma - \frac{\frac{7}{3}\lambda^2 + \frac{9}{2}\lambda}{5\lambda+6} \right) + \frac{3\lambda}{4a_x}.$$

We therefore choose

$$\frac{3\lambda}{4a_x} \le -(5\lambda+6)\left( \gamma - \frac{\frac{7}{3}\lambda + \frac{9}{2}\lambda}{5\lambda+6} \right) = \frac{7}{3}\lambda^2 + \frac{9}{2}\lambda - \frac{5\lambda^2+6\lambda}{2\lambda+2}.$$

It is, for that, sufficient to take

$$\frac{3\lambda}{4a_x} = \frac{3}{4}\lambda, \quad \text{i.e} \quad a_x = 1.$$

Furthermore

$$6\lambda(\gamma-\lambda) + 6\gamma(1-\gamma) + \gamma\lambda - 6\lambda + \frac{2\lambda^2}{3} + \frac{24L_W\lambda}{3} + 12L_W + \frac{6L_W}{a_x}$$
$$\le \frac{5\lambda^2+6\lambda}{2\lambda+2} - \frac{7}{3}\lambda^2 - \frac{9}{2}\lambda + \frac{3}{4}\lambda = \frac{5\lambda^2+6\lambda}{2\lambda+2} - \frac{7}{3}\lambda^2 - \frac{15}{4}\lambda.$$

We then observe

$$6\lambda(\gamma-\lambda) + 6\gamma(1-\gamma) + \gamma\lambda - 6\lambda + \frac{6L_W}{a_x} + \frac{2\lambda^2}{3} + \frac{24L_W\lambda}{3} + 12L_W \le -\lambda^2 - \frac{3}{4}\lambda.$$

And finally, for all $\lambda > 0$ and for all $x$

$$|x|^2\left(6\lambda\left(\gamma-\lambda\right)+6\gamma\left(1-\gamma\right)+\gamma\lambda-6\lambda+\frac{6L_W}{a_x}+\frac{2\lambda^2}{3}+\frac{48L_W\lambda}{3}+12L_W\right)$$
$$\leq -\lambda^2|x|^2-\frac{3}{4}\lambda|x|^2$$

**Coefficient of $|v|^2$.** We have, using $0<\gamma\leq\frac{1}{2}$ and $L_W\leq\lambda/8$

$$12\gamma-12+\frac{3}{2}+\frac{3}{4\lambda}24L_W\leq -6+\frac{3}{2}+\frac{18}{\lambda}\cdot\frac{\lambda}{8}=-6+\frac{3}{2}+\frac{9}{4}=-\frac{9}{4}.$$

We then choose

$$\frac{12\lambda}{8a_v}=\frac{9}{4},\quad \text{i.e}\quad a_v=\frac{2}{3}\lambda.$$

Therefore

$$\forall\lambda>0,\ \forall v,\qquad |v|^2\left(12\gamma-12+\frac{3}{2}+\frac{3}{4\lambda}24L_W+\frac{12L_W}{a_v}\right)\leq 0.$$

We thus obtain

$$\mathcal{L}_\mu H\left(x,v\right)$$
$$\leq 24\left(A-\left(\gamma-\lambda\right)\tilde{A}+d\right)+6L_W\mathbb{E}_\mu\left(|\cdot|\right)^2\left(1+\frac{4}{3}\lambda\right)-\lambda^2|x|^2-\frac{3}{4}\lambda|x|^2-\gamma H\left(x,v\right),$$

i.e

$$\mathcal{L}_\mu H\left(x,v\right)\leq 24\left(A-\left(\gamma-\lambda\right)\tilde{A}+d\right)+\mathbb{E}_\mu\left(|\cdot|\right)^2\left(\frac{3}{4}\lambda+\lambda^2\right)-\lambda^2|x|^2-\frac{3}{4}\lambda|x|^2-\gamma H\left(x,v\right). \tag{A.1}$$

### A.4 Proof of Lemma 2.3

Using $1-\gamma\geq\frac{1}{2}$, we get

$$H\left(x,v\right)\geq 24U(x)+\left(3+\lambda\right)|x|^2+12\left|v+\frac{x}{2}\right|^2-3|x|^2,$$

which is (2.2). We then have

$$H\left(x,v\right)\geq\min\left(\frac{2}{3}\lambda,6\right)\left(|v|^2+|x+v|^2\right).$$

Thus

$$\begin{aligned}
r(x,\tilde{x},v,\tilde{v})^2&\leq\left(\left(1+\alpha\right)|x-\tilde{x}+v-\tilde{v}|+\alpha|v-\tilde{v}|\right)^2\\
&\leq 2\left(1+\alpha\right)^2|x-\tilde{x}+v-\tilde{v}|^2+2\alpha^2|v-\tilde{v}|^2\\
&\leq 4\left(\left(1+\alpha\right)^2+\alpha^2\right)\left(|x+v|^2+|v|^2+|\tilde{x}+\tilde{v}|^2+|\tilde{v}|^2\right).
\end{aligned}$$

Therefore we obtain the final point.

### A.5 Proof of control of L1 and L2 Wasserstein distances

We prove Lemma 2.7. Using the definition of $R_1$ and (2.1), and since $B\geq d\geq 1$ and $\gamma\leq\frac{1}{2}$, we have $R_1\geq 1$.

- First for the L1-Wasserstein distance

$$|x-x'|+|v-v'| \leq |v-v'+x-x'|+2|x-x'| \leq \max\left(\frac{2}{\alpha},1\right) r\left((x,v),(x',v')\right).$$

If $r\left((x,v),(x',v')\right) \leq 1 \leq R_1$

$$r\left((x,v),(x',v')\right) \leq \frac{f(r)}{f'_-(R_1)} \leq \frac{\rho\left((x,v),(x',v')\right)}{\phi(R_1)\,g(R_1)}.$$

If $r\left((x,v),(x',v')\right) \geq 1$, we have shown (2.7)

$$r\left((x,v),(x',v')\right) \leq r^2\left((x,v),(x',v')\right) \leq 4\frac{(1+\alpha)^2+\alpha^2}{\min\left(\frac{2}{3}\lambda,6\right)}\left(H(x,v)+H(x',v')\right).$$

Thus

$$\begin{aligned}
r\left((x,v),(x',v')\right) &\leq \frac{4}{\epsilon}\frac{(1+\alpha)^2+\alpha^2}{\min\left(\frac{2}{3}\lambda,6\right)}\left(\epsilon H(x,v)+\epsilon H(x',v')\right)\\
&\leq \frac{4}{\epsilon}\frac{(1+\alpha)^2+\alpha^2}{\min\left(\frac{2}{3}\lambda,6\right)}\frac{\rho\left((x,v),(x',v')\right)}{f(r)}\\
&\leq \frac{4}{\epsilon}\frac{(1+\alpha)^2+\alpha^2}{\min\left(\frac{2}{3}\lambda,6\right)}\frac{\rho\left((x,v),(x',v')\right)}{f(1)}.
\end{aligned}$$

Therefore

$$\begin{aligned}
&|x-x'|+|v-v'|\\
&\quad\leq \max\left(\frac{2}{\alpha},1\right)\max\left(\frac{4\left((1+\alpha)^2+\alpha^2\right)}{\epsilon\min\left(\frac{2}{3}\lambda,6\right)f(1)},\frac{1}{\phi(R_1)\,g(R_1)}\right)\rho\left((x,v),(x',v')\right).
\end{aligned}$$

- Then for the L2-Wasserstein distance

$$|v-v'|^2 = |v-v'+x-x'-(x-x')|^2 \leq 2|v-v'+x-x'|^2+2|x-x'|^2.$$

Hence

$$|x-x'|^2+|v-v'|^2 \leq 3\left(|v-v'+x-x'|^2+|x-x'|^2\right).$$

But

$$\begin{aligned}
r^2\left((x,v),(x',v')\right) &= \left(\alpha|x-x'|+|x-x'+v-v'|\right)^2\\
&\geq \alpha^2|x-x'|^2+|x-x'+v-v'|^2\\
&\geq \left(1+\alpha^2\right)\left(|x-x'|^2+|x-x'+v-v'|^2\right)\\
&\geq \frac{1+\alpha^2}{3}\left(|x-x'|^2+|v-v'|^2\right).
\end{aligned}$$

If $r\left((x,v),(x',v')\right) \leq 1 \leq R_1$

$$r^2\left((x,v),(x',v')\right) \leq r\left((x,v),(x',v')\right) \leq \frac{f(r)}{f'_-(R_1)} \leq \frac{\rho\left((x,v),(x',v')\right)}{\phi(R_1)\,g(R_1)}.$$

If $r\left((x,v),(x',v')\right) \geq 1$, we have shown (2.7)

$$r^2\left((x,v),(x',v')\right) \leq 4\frac{(1+\alpha)^2+\alpha^2}{\min\left(\frac{2}{3}\lambda,6\right)}\left(H(x,v)+H(x',v')\right).$$

Thus

$$\begin{aligned} r\left(\left(x,v\right),\left(x',v'\right)\right) &\leq \frac{4}{\epsilon}\frac{\left(1+\alpha\right)^2+\alpha^2}{\min\left(\frac{2}{3}\lambda,6\right)}\left(\epsilon H\left(x,v\right)+\epsilon H\left(x',v'\right)\right) \\ &\leq \frac{4}{\epsilon}\frac{\left(1+\alpha\right)^2+\alpha^2}{\min\left(\frac{2}{3}\lambda,6\right)}\frac{\rho\left(\left(x,v\right),\left(x',v'\right)\right)}{f\left(r\right)} \\ &\leq \frac{4}{\epsilon}\frac{\left(1+\alpha\right)^2+\alpha^2}{\min\left(\frac{2}{3}\lambda,6\right)}\frac{\rho\left(\left(x,v\right),\left(x',v'\right)\right)}{f\left(1\right)}. \end{aligned}$$

Therefore

$$\begin{aligned} &|x-x'|^2+|v-v'|^2 \\ &\qquad \leq \frac{3}{1+\alpha^2}\max\left(\frac{4\left(\left(1+\alpha\right)^2+\alpha^2\right)}{\epsilon\min\left(\frac{2}{3}\lambda,6\right)f\left(1\right)},\frac{1}{\phi\left(R_1\right)g\left(R_1\right)}\right)\rho\left(\left(x,v\right),\left(x',v'\right)\right). \end{aligned}$$

### A.6 Proof of Lemma 2.8

We have

$$\begin{aligned} &H(x,v)-H(\tilde{x},\tilde{v}) \\ &\quad =24\left(U(x)-U(\tilde{x})\right)+\left(6(1-\gamma)+\lambda\right)\left(|x|^2-|\tilde{x}|^2\right)+12\left(x\cdot v-\tilde{x}\cdot\tilde{v}\right)+12\left(|v|^2-|\tilde{v}|^2\right) \\ &\quad =24\left(U(x)-U(\tilde{x})\right)+\left(6(1-\gamma)+\lambda-3\right)\left(|x|^2-|\tilde{x}|^2\right)+12\left(\left|v+\frac{x}{2}\right|^2-\left|\tilde{v}+\frac{\tilde{x}}{2}\right|^2\right). \end{aligned}$$

We first have

$$\left||x|^2-|\tilde{x}|^2\right|\leq|x-\tilde{x}|\left(|x|+|\tilde{x}|\right)\leq\frac{r(x,v,\tilde{x},\tilde{v})}{\alpha\sqrt{\lambda}}\left(\sqrt{H(x,v)}+\sqrt{H(\tilde{x},\tilde{v})}\right).$$

Then

$$\begin{aligned} \left|\left|v+\frac{x}{2}\right|^2-\left|\tilde{v}+\frac{\tilde{x}}{2}\right|^2\right| &\leq\left|v+\frac{x}{2}-\tilde{v}-\frac{\tilde{x}}{2}\right|\left(\left|v+\frac{x}{2}\right|+\left|\tilde{v}+\frac{\tilde{x}}{2}\right|\right) \\ &\leq\frac{1}{\sqrt{12}}|v-\tilde{v}+\frac{1}{2}(x-\tilde{x})|\left(\sqrt{H(x,v)}+\sqrt{H(\tilde{x},\tilde{v})}\right) \\ &\leq\frac{1}{2\sqrt{3}}\max\left(1,\frac{1}{2\alpha}\right)r(x,v,\tilde{x},\tilde{v})\left(\sqrt{H(x,v)}+\sqrt{H(\tilde{x},\tilde{v})}\right). \end{aligned}$$

And finally

$$\begin{aligned} |U(x)-U(\tilde{x})| &=\left|\int_0^1\nabla U\left(\tilde{x}+t(x-\tilde{x})\right)\cdot(x-\tilde{x})dt\right| \\ &\leq\sup_{t\in[0,1]}\left|\nabla U\left(\tilde{x}+t(x-\tilde{x})\right)\right||x-\tilde{x}| \\ &\leq\left(\nabla U(0)+L_U(|x|+|\tilde{x}|)\right)|x-\tilde{x}| \\ &\leq\left(\nabla U(0)+\frac{L_U}{\sqrt{\lambda}}\left(\sqrt{H(x,v)}+\sqrt{H(\tilde{x},\tilde{v})}\right)\right)\frac{r(x,v,\tilde{x},\tilde{v})}{\alpha}. \end{aligned}$$

These three inequalities yield the desired result.

## B Proof of Lemma 2.5

We first rewrite the various conditions on the parameters.

- Since for all $u \geq 0$, $0 < \phi(u) \leq 1$, we have $0 < \Phi(s) = \int_0^s \phi(u)\,du \leq s$, i.e $s/\Phi(s) \geq 1$. Therefore

$$\inf_{r\in]0,R_1]} \frac{r\phi(r)}{\Phi(r)} \geq \inf_{r\in]0,R_1]} \phi(r) = \phi(R_1).$$

  It is thus sufficient for (2.11) that

$$c + 2\epsilon B \leq \frac{1}{2}\left(1 - \frac{1}{\alpha}(L_U + L_W)\right)\phi(R_1).$$

- We have

$$\phi(r) \leq \exp\left(-\frac{L_U + L_W}{8\alpha} r^2\right).$$

  So

$$\Phi(r) \leq \int_0^\infty \exp\left(-\frac{L_U + L_W}{8\alpha} s^2\right) ds = \sqrt{\frac{2\pi\alpha}{L_U + L_W}}.$$

  Then

$$\int_0^{R_1} \frac{\Phi(r)}{\phi(r)} dr \leq \sqrt{\frac{2\pi\alpha}{L_U + L_W}} R_1 \frac{1}{\phi(R_1)}.$$

  It is thus sufficient for (2.12) that

$$c + 2\epsilon B \leq 2\sqrt{\frac{L_U + L_W}{2\pi\alpha}} \frac{\phi(R_1)}{R_1}.$$

At this point, we have now proven that under Assumption 1.1, Assumption 1.3 and Assumption 1.9, for the parameters to satisfy Lemma 2.5 it is sufficient for them to satisfy

$$\alpha > L_U + L_W, \tag{B.1}$$

$$c \leq \frac{\gamma}{6}\left(1 - \frac{\frac{5}{6}\gamma}{2\epsilon B + \frac{5}{6}\gamma}\right), \tag{B.2}$$

$$c + 2\epsilon B \leq \frac{1}{2}\left(1 - \frac{1}{\alpha}(L_U + L_W)\right)\phi(R_1), \tag{B.3}$$

$$c + 2\epsilon B \leq 2\sqrt{\frac{L_U + L_W}{2\pi\alpha}} \frac{\phi(R_1)}{R_1}, \tag{B.4}$$

with, again

$$B = 24\left(A + (\lambda - \gamma)\tilde{A} + d\right), \quad R_1 = \sqrt{(1+\alpha)^2 + \alpha^2}\sqrt{\frac{24}{5\gamma \min\left(3, \frac{1}{3}\lambda\right)}B}.$$

Let us show that there are positive parameters $\epsilon$, $\alpha$, $L_W$ and $c$ satisfying those conditions.

For inequality (B.1) it is sufficient, as $L_W < \frac{\lambda}{8}$, to consider

$$\alpha = L_U + \frac{\lambda}{4},$$

while inequality (B.2) first invites us to consider $2\epsilon B$ of a comparable order to $c$

$$2\epsilon B = \delta c.$$

We have thus switched parameter $\epsilon$ for $\delta$. First we translate (B.2) into our new parameter:

$$\begin{aligned} c \leq \frac{\gamma}{6}\left(1 - \frac{\frac{5}{6}\gamma}{2\epsilon B + \frac{5}{6}\gamma}\right) &\iff c \leq \frac{\gamma}{6}\frac{\delta c}{\delta c + \frac{5}{6}\gamma} \\ &\iff 1 \leq \frac{\gamma}{6}\frac{\delta}{\delta c + \frac{5}{6}\gamma} \text{ (since } c \geq 0) \\ &\iff c \leq \frac{\gamma}{6}\frac{\delta - 5}{\delta}. \end{aligned}$$

The appearance of $\phi(R_1)$ in (B.3) and (B.4) suggests we should try to minimize it. Let us assume, for simplicity, that $\epsilon \leq 1$, which is equivalent to having $c \leq \frac{2B}{\delta}$. We then have

$$\begin{aligned} \phi(r) &= \exp\left(-\frac{1}{8}\left(\frac{1}{\alpha}(L_U + L_W) + \alpha + 96\epsilon\max\left(\frac{1}{2\alpha}, 1\right)\right)r^2\right) \\ &\geq \exp\left(-\frac{1}{8}\left(\frac{L_U + L_W}{\alpha} + \alpha + 96\max\left(\frac{1}{2\alpha}, 1\right)\right)r^2\right) \text{ on } [0, R_1]. \end{aligned} \tag{B.5}$$

Now, using (B.5), we have for (B.3) and (B.4) that it is sufficient that

$$c \leq \frac{1}{2(\delta+1)}\left(1 - \frac{1}{\alpha}(L_U + L_W)\right)\exp\left(-\frac{1}{8}\left(\frac{L_U + L_W}{\alpha} + \alpha + 96\max\left(\frac{1}{2\alpha}, 1\right)\right)R_1^2\right),$$

and

$$c \leq \frac{2}{\delta+1}\sqrt{\frac{L_U + L_W}{2\pi\alpha}}\exp\left(-\frac{1}{8}\left(\frac{L_U + L_W}{\alpha} + \alpha + 96\max\left(\frac{1}{2\alpha}, 1\right)\right)R_1^2\right).$$

We could now optimize parameter $\delta$, but for the sake of conciseness, we choose $\delta = 6$.

Recall $0 \leq L_W < \frac{\lambda}{8}$. This way, both in (2.8) and (3.14), $c$ and $\mathcal{C}_1$ can be bounded independently of $L_W$. Hence why $L_W$, in (3.14) and (4.17), can be chosen last, and every other quantities can be chosen independently of $L_W$.

## References


[1] Dallas Albritton, Scott Armstrong, Jean-Christophe Mourrat, and Matthew Novack, *Variational methods for the kinetic Fokker-Planck equation*, 2019, arXiv:1902.04037.

[2] Dominique Bakry, Patrick Cattiaux, and Arnaud Guillin, *Rate of convergence for ergodic continuous Markov processes: Lyapunov versus Poincaré*, J. Funct. Anal. **254** (2008), no. 3, 727–759. MR2381160

[3] Fabrice Baudoin, *Bakry-Émery meet Villani*, J. Funct. Anal. **273** (2017), no. 7, 2275–2291. MR3677826

[4] Fabrice Baudoin, Maria Gordina, and David P. Herzog, *Gamma calculus beyond Villani and explicit convergence estimates for Langevin dynamics with singular potentials*, Arch. Ration. Mech. Anal. **241** (2021), no. 2, 765–804. MR4275746

[5] Dario Benedetto, Emanuele Caglioti, José A. Carrillo, and Mario Pulvirenti, *A non-Maxwellian steady distribution for one-dimensional granular media*, J. Statist. Phys. **91** (1998), no. 5-6, 979–990. MR1637274

[6] F. Bolley, *Separability and completeness for the Wasserstein distance*, Séminaire de probabilités XLI, Lecture Notes in Math., vol. 1934, Springer, Berlin, 2008, pp. 371–377. MR2483740

[7] François Bolley, Ivan Gentil, and Arnaud Guillin, *Uniform convergence to equilibrium for granular media*, Arch. Ration. Mech. Anal. **208** (2013), no. 2, 429–445. MR3035983

[8] François Bolley, Arnaud Guillin, and Florent Malrieu, *Trend to equilibrium and particle approximation for a weakly selfconsistent Vlasov-Fokker-Planck equation*, M2AN Math. Model. Numer. Anal. **44** (2010), no. 5, 867–884. MR2731396

[9] F. Bouchut and J. Dolbeault, *On long time asymptotics of the Vlasov-Fokker-Planck equation and of the Vlasov-Poisson-Fokker-Planck system with Coulombic and Newtonian potentials*, Differential Integral Equations **8** (1995), no. 3, 487–514. MR1306570

[10] Yu Cao, Jianfeng Lu, and Lihan Wang, *On explicit $l^2$-convergence rate estimate for underdamped langevin dynamics*, 2019, arXiv:1908.04746.

[11] José A. Carrillo, Matias G. Delgadino, and Grigorios A. Pavliotis, *A $\lambda$-convexity based proof for the propagation of chaos for weakly interacting stochastic particles*, J. Funct. Anal. **279** (2020), no. 10, 108734, 30. MR4137700

[12] 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** (2003), no. 3, 971–1018. MR2053570

[13] Patrick Cattiaux, Arnaud Guillin, and Florent Malrieu, *Probabilistic approach for granular media equations in the non-uniformly convex case*, Probab. Theory Related Fields **140** (2008), no. 1-2, 19–40. MR2357669

[14] Patrick Cattiaux, Arnaud Guillin, Pierre Monmarché, and Chaoen Zhang, *Entropic multipliers method for Langevin diffusion and weighted log Sobolev inequalities*, J. Funct. Anal. **277** (2019), no. 11, 108288, 24. MR4013832

[15] Louis-Pierre Chaintron and Antoine Diez, *Propagation of chaos: a review of models, methods and applications. II. Applications*, 2021, arXiv:2106.14812.

[16] Louis-Pierre Chaintron and Antoine Diez, *Propagation of chaos: a review of models, methods and applications. I. Models and methods*, 2022, arXiv:2203.00446.

[17] Xiang Cheng, Niladri S. Chatterji, Yasin Abbasi-Yadkori, Peter L. Bartlett, and Michael I. Jordan, *Sharp convergence rates for Langevin dynamics in the nonconvex setting*, May 2018, p. arXiv:1805.01648.

[18] Xiang Cheng, Niladri S. Chatterji, Peter L. Bartlett, and Michael I. Jordan, *Underdamped langevin mcmc: A non-asymptotic analysis*, Proceedings of the 31st Conference On Learning Theory (Sébastien Bubeck, Vianney Perchet, and Philippe Rigollet, eds.), Proceedings of Machine Learning Research, vol. 75, PMLR, 06–09 Jul 2018, pp. 300–323.

[19] Jean Dolbeault, Clément Mouhot, and Christian Schmeiser, *Hypocoercivity for linear kinetic equations conserving mass*, Trans. Amer. Math. Soc. **367** (2015), no. 6, 3807–3828. MR3324910

[20] Manh Hong Duong and Julian Tugaut, *The Vlasov-Fokker-Planck equation in non-convex landscapes: convergence to equilibrium*, Electron. Commun. Probab. **23** (2018), Paper No. 19, 10. MR3779816

[21] Alain Durmus, Andreas Eberle, Arnaud Guillin, and Raphael Zimmer, *An elementary approach to uniform in time propagation of chaos*, Proc. Amer. Math. Soc. **148** (2020), no. 12, 5387–5398. MR4163850

[22] Andreas Eberle, *Reflection couplings and contraction rates for diffusions*, Probab. Theory Related Fields **166** (2016), no. 3-4, 851–886. MR3568041

[23] Andreas Eberle, Arnaud Guillin, and Raphael Zimmer, *Couplings and quantitative contraction rates for Langevin dynamics*, Ann. Probab. **47** (2019), no. 4, 1982–2010. MR3980913

[24] Andreas Eberle, Arnaud Guillin, and Raphael Zimmer, *Quantitative Harris-type theorems for diffusions and McKean-Vlasov processes*, Trans. Amer. Math. Soc. **371** (2019), no. 10, 7135–7173. MR3939573

[25] Tadahisa Funaki, *A certain class of diffusion processes associated with nonlinear parabolic equations*, Z. Wahrsch. Verw. Gebiete **67** (1984), no. 3, 331–348. MR762085

[26] Arnaud Guillin, Wei Liu, Liming Wu, and Chaoen Zhang, *The kinetic Fokker-Planck equation with mean field interaction*, J. Math. Pures Appl. (9) **150** (2021), 1–23. MR4248461

[27] Arnaud Guillin, Wei Liu, Liming Wu, and Chaoen Zhang, *Uniform Poincaré and logarithmic Sobolev inequalities for mean field particle systems*, The Annals of Applied Probability **32** (2022), no. 3, 1590–1614.

[28] Arnaud Guillin and Pierre Monmarché, *Optimal linear drift for the speed of convergence of an hypoelliptic diffusion*, Electron. Commun. Probab. **21** (2016), Paper No. 74, 14. MR3568348

[29] 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** (2021), no. 2, Paper No. 15, 20. MR4333408

[30] Martin Hairer and Jonathan C. Mattingly, *Spectral gaps in Wasserstein distances and the 2D stochastic Navier-Stokes equations*, Ann. Probab. **36** (2008), no. 6, 2050–2091. MR2478676

[31] Samuel Herrmann and Julian Tugaut, *Non-uniqueness of stationary measures for self-stabilizing processes*, Stochastic Process. Appl. **120** (2010), no. 7, 1215–1246. MR2639745

[32] Mark Kac, *Foundations of kinetic theory*, Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability, 1954–1955, vol. III, University of California Press, Berkeley-Los Angeles, Calif., 1956, pp. 171–197. MR0084985

[33] Anna Kazeykina, Zhenjie Ren, Xiaolu Tan, and Junjian Yang, *Ergodicity of the underdamped mean-field Langevin dynamics*, arXiv e-prints (2020), arXiv:2007.14660.

[34] Don S. Lemons and Anthony Gythiel, *Paul Langevin's 1908 paper "On the Theory of Brownian Motion" ["Sur la théorie du mouvement brownien," C. R. Acad. Sci. (Paris) 146, 530-533 (1908)]*, American Journal of Physics **65** (1997), no. 11, 1079–1081.

[35] Yulong Lu and Jonathan C. Mattingly, *Geometric ergodicity of Langevin dynamics with Coulomb interactions*, Nonlinearity **33** (2020), no. 2, 675–699. MR4048775

[36] Henry P. McKean, Jr., *A class of Markov processes associated with nonlinear parabolic equations*, Proc. Nat. Acad. Sci. U.S.A. **56** (1966), 1907–1911. MR221595

[37] Sylvie Méléard, *Asymptotic behaviour of some interacting particle systems; McKean-Vlasov and Boltzmann models*, Probabilistic models for nonlinear partial differential equations (Montecatini Terme, 1995), Lecture Notes in Math., vol. 1627, Springer, Berlin, 1996, pp. 42–95. MR1431299

[38] Pierre Monmarché, *Long-time behaviour and propagation of chaos for mean field kinetic particles*, Stochastic Process. Appl. **127** (2017), no. 6, 1721–1737. MR3646428

[39] Pierre Monmarché, *Generalized $\Gamma$ calculus and application to interacting particles on a graph*, Potential Anal. **50** (2019), no. 3, 439–466. MR3925589

[40] Grigorios A. Pavliotis, *Stochastic processes and applications*, Texts in Applied Mathematics, vol. 60, Springer, New York, 2014, Diffusion processes, the Fokker-Planck and Langevin equations. MR3288096

[41] Daniel Revuz and Marc Yor, *Continuous martingales and Brownian motion*, third ed., Grundlehren der mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], vol. 293, Springer-Verlag, Berlin, 1999. MR1725357

[42] Samir Salem, *A gradient flow approach of propagation of chaos*, Discrete Contin. Dyn. Syst. **40** (2020), no. 10, 5729–5754. MR4128323

[43] Katharina Schuh, *Global contractivity for langevin dynamics with distribution-dependent forces and uniform in time propagation of chaos*, 2022, arXiv:2206.03082.

[44] Alain-Sol Sznitman, *Topics in propagation of chaos*, École d'Été de Probabilités de Saint-Flour XIX—1989, Lecture Notes in Math., vol. 1464, Springer, Berlin, 1991, pp. 165–251. MR1108185

[45] Denis Talay, *Stochastic Hamiltonian systems: exponential convergence to the invariant measure, and discretization by the implicit Euler scheme*, vol. 8, 2002, Inhomogeneous random systems (Cergy-Pontoise, 2001), pp. 163–198. MR1924934

[46] Cédric Villani, *Hypocoercivity*, Mem. Amer. Math. Soc. **202** (2009), no. 950, iv+141. MR2562709

[47] Anatoly A. Vlasov, *The vibrational properties of an electron gas*, Soviet Physics Uspekhi **10** (1968), no. 6, 721–733.