Local convergence rates for Wasserstein gradient flows and McKean-Vlasov equations with multiple stationary solutions

Pierre Monmarché[1] and Julien Reygner[2]

[1]LJLL and LCT, Sorbonne Université, Paris, France
[2]CERMICS, Ecole des Ponts, Marne-la-Vallée, France

April 25, 2024

Abstract

Non-linear versions of log-Sobolev inequalities, that link a free energy to its dissipation along the corresponding Wasserstein gradient flow (i.e. corresponds to Polyak-Lojasiewicz inequalities in this context), are known to provide global exponential long-time convergence to the free energy minimizers, and have been shown to hold in various contexts. However they cannot hold when the free energy admits critical points which are not global minimizers, which is for instance the case of the granular media equation in a double-well potential with quadratic attractive interaction at low temperature. This work addresses such cases, extending the general arguments when a log-Sobolev inequality only holds locally and, as an example, establishing such local inequalities for the granular media equation with quadratic interaction either in the one-dimensional symmetric double-well case or in higher dimension in the low temperature regime. The method provides quantitative convergence rates for initial conditions in a Wasserstein ball around the stationary solutions. The same analysis is carried out for the kinetic counterpart of the gradient flow, i.e. the corresponding Vlasov-Fokker-Planck equation. The local exponential convergence to stationary solutions for the mean-field equations, both elliptic and kinetic, is shown to induce for the corresponding particle systems a fast (i.e. uniform in the number or particles) decay of the particle system free energy toward the level of the non-linear limit.

1 Introduction

We are interested in the long-time behavior of the granular media equation

$$\partial_t \rho_t = \nabla \cdot \left(\sigma^2 \nabla \rho_t + \left(\nabla V + \rho_t \star \nabla W \right) \rho_t \right) \tag{1}$$

or of more general McKean-Vlasov semilinear equations (see (9) below). Here, ρ_t is a probability density over $\mathbb{R}^d$, $V \in \mathcal{C}^2(\mathbb{R}^d, \mathbb{R})$, $W \in \mathcal{C}^2(\mathbb{R}^d \times \mathbb{R}^d, \mathbb{R})$ are respectively a confining and interaction potential, $W(x,y) = W(y,x)$ for all $x, y \in \mathbb{R}^d$, $\sigma^2 > 0$ stands for the temperature (or diffusivity) and $\rho \star \nabla W(x) = \int_{\mathbb{R}^d} \nabla_x W(x,y) \rho(y) \mathrm{d}y$. We have typically in mind the double-well case where $d = 1$,

$$V(x) = \frac{x^4}{4} - \frac{x^2}{2}, \qquad W(x,y) = \theta(x-y)^2 \tag{2}$$

for some $\theta \in \mathbb{R}$. When $\theta \leqslant 0$ (repulsive interaction), or when $\theta > 0$ and σ^2/θ is large enough (attractive interaction at high temperature or small interaction), there is a unique stationary solution to (1), which is globally attractive with exponential rate. However for a fixed $\theta > 0$ there is a phase transition at some critical temperature $\sigma_c^2 > 0$ such that, for $\sigma^2 < \sigma_c^2$, (1) with (2) admits three distinct stationary solutions (one symmetric, unstable, and two non-symmetric, stable) [46]. We are interested in this second case, and more precisely on obtaining *local* convergence rates, namely to prove that solutions which start close to a stable stationary solution converge exponentially fast to it. Among other motivations, this is an important question both for the theoretical understanding of metastable interacting particle systems, where the stability property of the non-linear limit drives the short-time behavior and induces the metastable behaviour [22, 12, 3, 20, 28, 5], and for practical questions of optimization in the Wasserstein space, for instance in the mean-field modelling of artificial neural networks [16, 29, 36, 35] (the loss function being convex only in toy models, like one-layer networks).

To illustrate the point, let us state a result obtained with our approach in the specific case of the symmetric double well (2). In the next statement (proven in Section 3.2.1, see Remarks 7 and 12) as in the rest of the work we write $\mathcal{W}_2$ and $\|\cdot\|_{TV}$ respectively the $L2$ Wasserstein distance and total variation norm and $\mathcal{P}_2(\mathbb{R}^d)$ the set of probability measures on $\mathbb{R}^d$ with finite second moment.

Proposition 1. *Consider the granular media* (1) *in the case* (2) *with* $\theta > 0$.

1. *If* $\sigma^2 < \sigma_c^2$, *let* ρ_* *be one of the two non-symmetric stationary solutions. There exist* $\delta, \lambda, C > 0$ *such that for any initial condition* $\rho_0 \in \mathcal{P}_2(\mathbb{R})$ *with* $\mathcal{W}_2(\rho_0, \rho_*) \leqslant \delta$, *the corresponding solution to* (1) *satisfies, for all* $t \geqslant 0$

$$\mathcal{W}_2(\rho_t, \rho_*) + \|\rho_t - \rho_*\|_{TV} \leqslant C e^{-\lambda t} .$$

2. *If* $\sigma^2 = \sigma_c^2$, *denote by* ρ_* *the unique stationary solution. For any* $\rho_0 \in \mathcal{P}_2(\mathbb{R})$, *there exists* $C_0 > 0$ *such that the corresponding solution to* (1) *satisfies, for all* $t \geqslant 0$,

$$\mathcal{W}_2(\rho_t, \rho_*) + \|\rho_t - \rho_*\|_{TV} \leqslant \frac{C_0}{t^{1/3}} .$$

Notice that, in the sub-critical case, there is no assumption on σ^2 other than $\sigma^2 < \sigma_c^2$, namely, the result holds arbitrarily close to the critical temperature (our results also apply in the simpler super-critical regime and provide, as soon as $\sigma^2 > \sigma_c^2$, a global quantitative exponential convergence toward the unique stationary solution, see Remark 8). This is in contrast to the results of [49], also concerned with quantitative local convergence of (1), which require the temperature to be sufficiently small (for a given θ). Moreover, the Wasserstein convergence in [49, Theorem 2.3] only gives a convergence speed of order $e^{-\lambda\sqrt{t}}$ for some $\lambda > 0$, and additionally it requires $\theta > -\inf V''$, which our result does not (under this additional assumption we get that the rate λ is uniform over $\sigma^2 \in (0, \sigma_0^2]$ for any $\sigma_0^2 \in (0, \sigma_c^2)$, as in Proposition 16). The constants δ, λ, C in Proposition 1 can be made explicit (see Remark 9), namely the result is quantitative.

Our method is the following. As recalled in the next section, it is now well-known that (1) can be seen as the gradient flow in the Wasserstein space of some functional. To fix ideas, consider on $\mathbb{R}^d$ such a gradient flow

$$\dot{x}_t = -\nabla f(x_t) .$$

Let x_* be a local minimizer of f. Assuming a Polyak-Lojasiewicz inequality, namely that there exists $\eta > 0$ such that

$$0 \leqslant f(x) - f(x_*) \leqslant \eta |\nabla f(x)|^2 \tag{3}$$

in a neighborhood $\mathcal{A}$ of x_*, then by differentiating $f(x_t) - f(x_*)$ we immediately get that, as long as x_t stays within $\mathcal{A}$,

$$f(x_t) - f(x_*) \leqslant e^{-t/\eta}(f(x_0) - f(x_*)) \tag{4}$$

and

$$\partial_t \sqrt{f(x_t) - f(x_*)} \leqslant -\frac{1}{2\sqrt{\eta}} |\nabla f(x_t)| \,,$$

from which

$$|x_t - x_0| = \left| \int_0^t \nabla f(x_s) \mathrm{d}s \right| \leqslant 2\sqrt{\eta\,(f(x_0) - f(x_*))} \,. \tag{5}$$

Using that f is continuous so that the right hand side can be made arbitrarily small by taking x_0 sufficiently close to x_*, this shows that $\mathcal{A}$ contains a ball centered at x_* such that, starting with an initial condition in this ball, the gradient flows remains in $\mathcal{A}$ and thus the previous inequalities hold for all times $t \geqslant 0$. Assuming furthermore that x_* is the unique critical point of f in $\mathcal{A}$, by the LaSalle invariance principle, x_t converges to x_* and then letting $t \to \infty$ in (5) and applying it with x_0 replaced by x_t gives

$$|x_* - x_t|^2 \leqslant 4\eta\,(f(x_t) - f(x_*)) \leqslant 4\eta e^{-t/\eta}\,(f(x_0) - f(x_*)) \ ,$$

which proves the exponential convergence to x_*.

We show that this argument extends to the infinite dimensional settings of gradient flows over the Wasserstein space (the lack of continuity of f with respect to the Wasserstein distance in this case being circumvented by the regularization properties of (1) which, with the present notations, amounts to say that $x_0 \mapsto f(x_t)$ is continuous at x_* as soon as $t > 0$). The key new ingredient is thus the (local) dissipation inequality (3), which for elliptic McKean-Vlasov equations as (1) corresponds to a (local) non-linear log-Sobolev inequality (LSI). Such inequalities have been investigated in a number of works (see e.g. [11, 22, 26] and references within) but, to our knowledge, only in cases where they are global, corresponding in the finite-dimensional settings above to the case where (3) holds for all $x \in \mathbb{R}^d$ (for instance when f is uniformly strongly convex). Our main contribution is thus to show that the method also works locally and, more importantly, to show that it is indeed possible to establish such local dissipation inequalities in some cases where the global inequality fails. More specifically, the main results of this work are the following:

- In the general framework of gradient flows with respect to $\mathcal{W}_2$, under suitable regularity conditions, the local non-linear LSI that is the analogue of (3) is shown to imply the exponential convergence (in $\mathcal{W}_2$ and relative entropy) towards the local minimizer for all initial conditions in a suitable $\mathcal{W}_2$ ball (this is Theorem 8).

- The same result is established in the kinetic case, i.e. for the Vlasov-Fokker-Planck equation (this is Theorem 20).

- In the particular case of the granular media equation, when the interaction is parametrized by some moments of the measure, a simple criterion for the local non-linear LSI is given in Proposition 9. It is then illustrated in two cases with quadratic interactions, the one-dimensional double-well case (Proposition 13) and the multi-well case in $\mathbb{R}^d$ (Proposition 16).

- Under the same conditions as Theorem 8 (or Theorem 20 in the kinetic case), the free energy of the corresponding system of interacting particles is shown to decay fast below the level of the limit of the non-linear limit, as stated in Proposition 21 (or Proposition 22 in the kinetic case).

This work is organized as follows. The general framework is introduced in Section 2, where the main general result (Theorem 8) is stated and proven. Section 3 addresses the question of establishing local non-Linear LSI, covering the granular media case (with a detailed study of the one-dimensional double-well potential). The Vlasov-Fokker-Planck equation is studied in Section 4, and Section 5 is devoted to interacting particle systems.

To conclude this introduction, let us mention some perspectives of this work. The gradient descent structure underlying our study is quite flexible, as one may restrict the space over which the free energy is minimized (e.g. tensorized distributions, Gaussian distributions, distributions with some fixed marginals…), which is of interest for variational inference [33, 32, 18], and one can modify the metric with respect to which the gradient is taken, which amounts to add some non-constant (possibly non-linear) diffusion matrix, allowing e.g. for slow or fast diffusion processes and congestion effects [23, 9]. We expect most of our analysis to extend to this kind of settings. Non-asymptotic bounds for discrete-time numerical schemes can also be obtained as in [38, 8]. Extension to time-varying temperature (for annealing or as a surrogate to stochastic gradient descent [44]) is straightforward, following e.g. [31, 37] and references within.

Last, let us mention that for McKean-Vlasov equations which do not necessarily write as Wasserstein gradient flows, but under different assumptions than ours, a similar statement to Proposition 1 was recently obtained in [19]. It provides exponential convergence when the initial condition is in a small $\mathcal{W}_1$ neighborhood of a stationary solution. The method is completely different and relies on the differentiation, in the sense of Lions derivatives, of the drift of the underlying nonlinear SDE in the neighborhood of the considered invariant measure, which then provides a criterion for the stability of the invariant measure.

2 Local convergence rates with log-Sobolev inequalities

The main result of this section is Theorem 8, stated and proven in Section 2.6. Before that, the relevant notions, conditions and lemmas are gradually introduced.

2.1 General settings, assumptions and notations

We consider an energy functional $\mathcal{E} : \mathcal{P}_2(\mathbb{R}^d) \to (-\infty, \infty]$ and, for a temperature $\sigma^2 > 0$, the free energy

$$\mathcal{F}(\rho) = \mathcal{E}(\rho) + \sigma^2 \mathcal{H}(\rho)$$

for $\rho \in \mathcal{P}_2(\mathbb{R}^d)$, where $\mathcal{H}(\rho) = \int_{\mathbb{R}^d} \rho \ln \rho$ stands for the entropy (taken as $+\infty$ if ρ is not absolutely continuous; otherwise we also write ρ its density).

Assumption 1 (boundedness from below of the free energy)**.** *The free energy $\mathcal{F}$ is bounded from below.*

Given a functional $\mathcal{G} : \mathcal{P}_2(\mathbb{R}^d) \to (-\infty, \infty]$, a measurable function $\frac{\delta \mathcal{G}}{\delta \mu} : \mathcal{P}_2(\mathbb{R}^d) \times \mathbb{R}^d \to \mathbb{R}$ is called a linear functional derivative of $\mathcal{G}$ if, for all $\mu_1, \mu_0 \in \mathcal{P}_2(\mathbb{R}^d)$ with $\mathcal{G}(\mu_0) + \mathcal{G}(\mu_1) < \infty$,

$$\mathcal{G}(\mu_1) - \mathcal{G}(\mu_0) = \int_0^1 \int_{\mathbb{R}^d} \frac{\delta \mathcal{G}(\mu_t)}{\delta \mu}(x)(\mu_1 - \mu_0)(\mathrm{d}x)\mathrm{d}t \,, \tag{6}$$

where $\mu_t = t\mu_1 + (1-t)\mu_0$.

Assumption 2 (linear functional derivative of the energy)**.** *$\mathcal{E}$ admits a linear functional derivative that we denote $E_\rho(x) = \frac{\delta \mathcal{E}(\rho)}{\delta \mu}(x)$. Moreover,*

(i) the function $(x,\rho) \in \mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d) \mapsto E_\rho(x)$ is continuous, where $\mathcal{P}_2(\mathbb{R}^d)$ is endowed with the $\mathcal{W}_2$ distance;

(ii) for any $\rho \in \mathcal{P}_2(\mathbb{R}^d)$, $E_\rho \in \mathcal{C}^2(\mathbb{R}^d)$;

(iii) for all $\mu_0, \mu_1 \in \mathcal{P}_2(\mathbb{R}^d)$ with $\mathcal{E}(\mu_0) + \mathcal{E}(\mu_1) < \infty$, $x \mapsto \sup_{t\in[0,1]} |E_{\mu_t}(x)|$ is in $L^1(\mu_0) \cap L^1(\mu_1)$.

The last item of Assumption 2 ensures that the integral in the right hand side of (6) is well defined when $\mathcal{G} = \mathcal{E}$, and moreover that

$$\lim_{t\to 0} \frac{\mathcal{E}(\mu_0 + t(\mu_1 - \mu_0)) - \mathcal{E}(\mu_0)}{t} = \int_{\mathbb{R}^d} E_{\mu_0}(x)(\mu_1 - \mu_0)(\mathrm{d}x). \tag{7}$$

We are interested in the McKean-Vlasov equation

$$\partial_t \rho_t = \sigma^2 \Delta \rho_t + \nabla \cdot (\rho_t \nabla E_{\rho_t}) \ , \tag{8}$$

equivalently

$$\partial_t \rho_t = \nabla \cdot \left(\rho_t \nabla \frac{\delta \mathcal{F}(\rho_t)}{\delta \mu} \right) . \tag{9}$$

In particular, the granular media equation (1) corresponds to the free energy

$$\mathcal{F}(\rho) = \sigma^2 \int_{\mathbb{R}^d} \rho \ln \rho + \int_{\mathbb{R}^d} V\rho + \frac{1}{2} \int_{\mathbb{R}^d \times \mathbb{R}^d} W \rho^{\otimes 2} \,, \tag{10}$$

for which $E_\rho(x) = V(x) + \rho \star W(x)$ and Assumptions 1 and 2 are satisfied as soon as, for instance, V is convex outside a compact set and W is lower bounded with at most a quadratic growth at infinity (other conditions can be considered, for instance in repulsive cases W may not be bounded below but Assumption 1 still holds if V grows faster than $-W$ at infinity) Under suitable regularity condition, a straightforward computation gives

$$\partial_t \mathcal{F}(\rho_t) = - \int_{\mathbb{R}^d} \left| \nabla \frac{\delta \mathcal{F}(\rho_t)}{\delta \mu} \right|^2 \mathrm{d}\rho_t \,. \tag{11}$$

In particular, the free energy is decreasing along time, and the stationary solutions ρ_* are the critical points of $\mathcal{F}$, characterized by the fact that $\frac{\delta \mathcal{F}}{\delta \mu}(\rho_*)$ is constant, which, using that $\frac{\delta \mathcal{H}}{\delta \mu}(\rho) = \ln \rho$, is equivalent to the self-consistency equation

$$\rho_* \propto \exp\left(-\frac{1}{\sigma^2} E_{\rho_*} \right) . \tag{12}$$

In cases where $\mathcal{F}$ satisfies a global non-linear log-Sobolev inequality, in the sense that there exists a constant $\overline{\eta} > 0$ such that

$$\forall \rho \in \mathcal{P}_2(\mathbb{R}^d) \,, \qquad \mathcal{F}(\rho) - \inf \mathcal{F} \leqslant \overline{\eta} \int_{\mathbb{R}^d} \left| \nabla \frac{\delta \mathcal{F}(\rho)}{\delta \mu} \right|^2 \mathrm{d}\rho \,, \tag{\textbf{G-NL-LSI}}$$

we immediately get from (11) that the free energy decays exponentially fast toward its infimum. In particular such an inequality is clearly false when $\mathcal{F}$ admits critical points which are not global minimizers, which is precisely the case we are interested in.

Assumption 3 (local equilibrium). *For any $\rho \in \mathcal{P}_2(\mathbb{R}^d)$,*

$$Z_\rho = \int_{\mathbb{R}^d} \exp\left(-\frac{1}{\sigma^2} E_\rho\right) < \infty.$$

For the granular media equation (1) with free energy given by (10), this assumption holds if for instance W is lower bounded and $\ln|x| = o(V(x))$ at infinity. Under Assumption 3, denoting

$$\Gamma(\rho) = Z_\rho^{-1} \exp\left(-\frac{1}{\sigma^2} E_\rho\right) , \tag{13}$$

which we call the local equilibrium as it is the stationary solution of the linear equation

$$\partial \tilde{\rho}_t = \sigma^2 \Delta \tilde{\rho}_t + \nabla \cdot (\tilde{\rho}_t \nabla E_\rho) ,$$

we can also interpret the free energy dissipation in (11) as

$$\int_{\mathbb{R}^d} \left| \nabla \frac{\delta \mathcal{F}}{\delta \mu}(\rho) \right|^2 \mathrm{d}\rho = \sigma^4 \mathcal{I}\left(\rho | \Gamma(\rho)\right) ,$$

where $\mathcal{I}(\nu|\mu)$ stands for the Fisher information of ν with respect to μ,

$$\mathcal{I}(\nu|\mu) = \int_{\mathbb{R}^d} \left| \nabla \ln \frac{\mathrm{d}\nu}{\mathrm{d}\mu} \right|^2 \mathrm{d}\nu .$$

We write

$$\mathcal{K} = \{\rho_* \in \mathcal{P}_2(\mathbb{R}^d),\ \mathcal{F}(\rho_*) < \infty,\ \rho_* = \Gamma(\rho_*)\}$$

the set of critical points of $\mathcal{F}$.

In this work we are interested in cases where the local equilibria satisfy a uniform (classical) log-Sobolev inequality, in the sense that there exist $\eta > 0$ such that

$$\forall \nu, \rho \in \mathcal{P}_2(\mathbb{R}^d) , \qquad \mathcal{H}\left(\nu | \Gamma(\rho)\right) \leqslant \eta \mathcal{I}\left(\nu | \Gamma(\rho)\right) , \tag{\textbf{U-LSI}}$$

where

$$\mathcal{H}(\nu|\mu) = \int_{\mathbb{R}^d} \ln \frac{\mathrm{d}\nu}{\mathrm{d}\mu} \mathrm{d}\nu$$

stands for the relative entropy of ν with respect to μ. Indeed, contrary to (**G-NL-LSI**), there are many tools to establish (**U-LSI**) and it can typically hold even if $\mathcal{F}$ has several critical points. For instance, for the granular media equation (1) in the double well case (2) with attractive interaction (namely $\theta > 0$), we can decompose $V = V_0 + V_1$ where V_0 is strongly convex and V_1 is bounded, so that, for any ρ,

$$V + \rho \star W = V_0 + V_1 + \rho \star W$$

is the sum of a strongly convex potential $V_0 + \rho \star W$ (with a lower bound on the curvature independent from ρ) and a bounded potential V_1 (independent from ρ), so that (**U-LSI**) follows from classical Bakry-Emery and Holley-Stroock results [2]. One motivation of the present work is to understand the difference between the uniform classical log-Sobolev inequality (**U-LSI**) and the non-linear log-Sobolev inequality (**G-NL-LSI**).

The well-posedness for (9) for specific cases is a standard question that we do not address here, we refer the interested reader to [1] for general considerations on Wasserstein gradient flows.

Assumption 4 (well-posedness and regularity of (9))**.** *For all $\rho_0 \in \mathcal{P}_2(\mathbb{R}^d)$,* (9) *has a unique strong solution, continuous in time for $\mathcal{W}_2$, which is the gradient flow of $\mathcal{F}$ in the sense of [1, Definition 11.1.1]; for all $t > 0$, ρ_t has a continuous positive density, $\mathcal{F}(\rho_t)$ and $\mathcal{H}(\rho_t|\Gamma(\rho_t))$ are finite; for almost all times $t > 0$, $\mathcal{I}(\rho_t|\Gamma(\rho_t))$ is finite and* (11) *holds.*

The fact that the free energy, local relative entropy and Fisher information become finite instantaneously will be stated in more quantitative ways along the study. In particular, the results stated in Section 2.5, combined with the next statement from [1], show that Assumption 4 is met in particular in the granular media case (1) in the settings of Proposition 1.

Proposition 2. *In the case* (10) *with $\nabla^2 V \geq \lambda I_d$ for some $\lambda \in \mathbb{R}$, $W(x,y) = w(x-y)$ with an even and convex function $w : \mathbb{R}^d \to [0,\infty)$ which has the doubling property $w(x+y) \leq c_w(1 + w(x) + w(y))$, the following holds. For any initial condition $\rho_0 \in \mathcal{P}_2(\mathbb{R}^d)$, there is a unique distributional solution to* (1)*, which is the gradient flow of $\mathcal{F}$ and possesses the following properties:*

(i) for any $t > 0$, ρ_t has a density with respect to the Lebesgue measure on $\mathbb{R}^d$;

(ii) $\rho_t \to \rho_0$ when $t \to 0$ in $\mathcal{P}_2(\mathbb{R}^d)$;

(iii) $\rho_\cdot \in L^1_{\mathrm{loc}}((0,\infty), W^{1,1}_{\mathrm{loc}}(\mathbb{R}^d))$;

(iv) for any $0 < s < t < \infty$,

$$\mathcal{F}(\rho_s) = \mathcal{F}(\rho_t) + \sigma^4 \int_s^t \mathcal{I}(\rho_r|\Gamma(\rho_r))\mathrm{d}r < \infty.$$

Proof. With the present conditions on V and W the assumptions of [1, Example 11.2.7] are satisfied, so that [1, Theorem 11.2.8] applies, which gives all the points of Proposition 2. □

From now on, Assumptions 1, 2, 3 and 4 are systematically enforced.

2.2 Two known cases with global convergence

To further motivate our study, let us now highlight that it is known that (**U-LSI**) is in fact sufficient to conclude (concerning the long-time behavior of the equation) in two cases: when $\mathcal{E}$ is functional-convex, or when the interaction is sufficiently small (two cases which, of course, do not allow for multiple stationary solutions).

We start with functional-convexity (to be distinguished from displacement-convexity), which by definition means that

$$\mathcal{E}(t\mu_0 + (1-t)\mu_1) \leqslant t\mathcal{E}(\mu_0) + (1-t)\mathcal{E}(\mu_1) \qquad t \in [0,1]\,, \tag{14}$$

for all $\mu_0, \mu_1 \in \mathcal{P}_2(\mathbb{R}^d)$. As stated in e.g. [16], assuming furthermore that $\mathcal{F}$ admits a minimizer $\rho_* \in \mathcal{K}$ then the latter is unique (thanks to the strict convexity of the entropy), and for all $\rho \in \mathcal{P}_2(\mathbb{R}^d)$, the following entropy sandwich inequalities hold:

$$\sigma^2 \mathcal{H}(\rho|\rho_*) \leqslant \mathcal{F}(\rho) - \mathcal{F}(\rho_*) \leqslant \sigma^2 \mathcal{H}\left(\rho|\Gamma(\rho)\right)\,, \tag{15}$$

see also Lemma 3 below. In particular, the second inequality, together with (**U-LSI**), implies (**G-NL-LSI**) with $\overline{\eta} = \eta/\sigma^2$, hence the exponential decay of $\mathcal{F}(\rho_t)$ to its global minimum which, by the first inequality of (15), gives the exponential convergence of $\mathcal{H}(\rho_t|\rho_*)$ to zero (which in turns implies the exponential convergence of ρ_t to ρ_* in total variation and, because ρ_* satisfies a classical LSI thanks to (**U-LSI**) and thus a Talagrand inequality [41], in Wasserstein 2 distance).

The convexity condition (14) is known to hold in various settings, in particular for mean-field models of one-layer neuron networks [29, 16]. In the granular media case (10), assume that the interaction potential is of the form

$$W(x,y) = W_0(x) + W_0(y) + 2\sum_{k\in\mathbb{N}} r_k(x)r_k(y) - 2\sum_{k\in\mathbb{N}} a_k(x)a_k(y) \tag{16}$$

for some functions W_0, a_k, r_k (as in e.g. the quadratic case (2); the letters a and r refers to *attractive* and *repulsive* by analogy with the quadratic case). Then we see that, for $\mathcal{E}(\rho) = \int(V + \frac{1}{2}\rho \star W)\rho$ and $t \in [0,1]$,

$$\begin{aligned}\mathcal{E}(t\mu_0 + (1-t)\mu_1) - t\mathcal{E}(\mu_0) - (1-t)\mathcal{E}(\mu_1)& \\ = t(1-t)\sum_{k\in\mathbb{N}}\left[\left(\int_{\mathbb{R}^d} a_k(\mu_0-\mu_1)\right)^2 - \left(\int_{\mathbb{R}^d} r_k(\mu_0-\mu_1)\right)^2\right]&,\end{aligned} \tag{17}$$

so that (14) holds for all μ_0, μ_1, t if and only if $a_k = 0$ for all $k \in \mathbb{N}$, which corresponds to repulsive interaction (and for quadratic potentials (2) to the case where $\theta \leqslant 0$, i.e. W is concave). See also [15, Section 3.1] for further examples satisfying (14).

In the general non-functional-convex cases, assuming that $\rho_* \in \mathcal{K}$ is a global minimizer of $\mathcal{F}$, notice that the first inequality in (15) cannot hold if there is another global minimizer $\rho' \in \mathcal{K}$ (since we would have $\mathcal{H}(\rho'|\rho_*) > 0$ while $\mathcal{F}(\rho') = \mathcal{F}(\rho_*)$) and the second inequality (which is the one we used to get the global non-linear LSI (**G-NL-LSI**) from the uniform classical LSI (**U-LSI**)) cannot hold as soon as there is a critical point of $\mathcal{F}$ which is not a global minimizer (since $\mathcal{H}(\rho|\Gamma(\rho)) = 0$ when $\rho \in \mathcal{K}$).

Second, consider the non-convex but small Lipschitz interaction settings. We assume that $\mu \mapsto \nabla E_\mu$ is uniformly Lipschitz, namely there exists $L > 0$ such that

$$\forall \rho, \rho' \in \mathcal{P}_2(\mathbb{R}^d)\,, \qquad \|\nabla E_\rho - \nabla E_{\rho'}\|_\infty \leqslant L\mathcal{W}_2(\rho,\rho')\,. \tag{\textbf{Lip}}$$

We still assume (**U-LSI**), which by [41] implies the uniform T_2 Talagrand inequality

$$\forall \nu, \rho \in \mathcal{P}_2(\mathbb{R}^d)\,, \qquad \mathcal{W}_2^2\left(\nu, \Gamma(\rho)\right) \leqslant 4\eta\mathcal{H}\left(\nu|\Gamma(\rho)\right)\,. \tag{18}$$

Let $\rho_* \in \mathcal{K}$. A classical computation shows that

$$\partial_t \mathcal{H}(\rho_t|\rho_*) = -\sigma^2\mathcal{I}(\rho_t|\rho_*) + \int_{\mathbb{R}^d}(\nabla E_{\rho_*} - \nabla E_{\rho_t})\cdot\nabla\ln\left(\frac{\rho_t}{\rho_*}\right)\mathrm{d}\rho_t\,.$$

Using the Cauchy-Schwarz, log-Sobolev and Talagrand inequalities,

$$\begin{aligned}\partial_t\mathcal{H}(\rho_t|\rho_*) &\leqslant -\sigma^2\mathcal{I}(\rho_t|\rho_*) + \|\nabla E_{\rho_*} - \nabla E_{\rho_t}\|_\infty\sqrt{\mathcal{I}(\rho_t|\rho_*)} \\ &\leqslant -\frac{\sigma^2}{2}\mathcal{I}(\rho_t|\rho_*) + \frac{L^2}{2\sigma^2}\mathcal{W}_2^2(\rho_t,\rho_*) \\ &\leqslant \left(-\frac{\sigma^2}{2\eta} + \frac{2L^2\eta}{\sigma^2}\right)\mathcal{H}(\rho_t|\rho_*)\,,\end{aligned}$$

In particular, as soon as $L\eta < 2\sigma^2$, we get an exponential convergence in relative entropy (hence total variation and $\mathcal{W}_2$) of ρ_t to ρ_*, and in particular uniqueness of the critical point ρ_*.

2.3 Lower-bounded curvature

To revisit and generalize the two previous cases, let us consider the case where there exists a cost functional $\mathcal{C} : \mathcal{P}_2(\mathbb{R}^d) \times \mathcal{P}_2(\mathbb{R}^d) \to (-\infty, \infty]$ with $\mathcal{C}(\mu_0, \mu_1) = \mathcal{C}(\mu_1, \mu_0)$ and such that

$$\forall \mu_0, \mu_1 \in \mathcal{P}_2(\mathbb{R}^d), t \in [0,1], \\ \mathcal{E}(t\mu_0 + (1-t)\mu_1) \leqslant t\mathcal{E}(\mu_0) + (1-t)\mathcal{E}(\mu_1) + t(1-t)\mathcal{C}(\mu_0, \mu_1)\,. \qquad (\mathcal{C}\textbf{-curv-}\mathcal{E})$$

Functional-convexity corresponds to $\mathcal{C} = 0$. If $\mathcal{C} = \lambda \mathcal{W}_2^2$ for some $\lambda \in \mathbb{R}$, namely if

$$\forall \mu_0, \mu_1 \in \mathcal{P}_2(\mathbb{R}^d), t \in [0,1], \\ \mathcal{E}(t\mu_0 + (1-t)\mu_1) \leqslant t\mathcal{E}(\mu_0) + (1-t)\mathcal{E}(\mu_1) + \lambda t(1-t)\mathcal{W}_2^2(\mu_0, \mu_1)\,, \qquad (\mathcal{W}_2^2\textbf{-curv-}\mathcal{E})$$

it can be interpreted in terms of Otto calculus by saying that the Hessian of $\mathcal{E}$ is lower-bounded by $-\lambda$. In the case (16) a bound ($\mathcal{C}$**-curv-**$\mathcal{E}$) follows from (17), and if a_k is L_k-Lipschitz for all $k \in \mathbb{N}$ with $\lambda = \sum_{k\in\mathbb{N}} L_k^2 < \infty$, ($\mathcal{W}_2^2$**-curv-**$\mathcal{E}$) holds (in fact, a stronger inequality holds, with $\mathcal{W}_2$ replaced by $\mathcal{W}_1$, which could be of interest in some cases; in this work we use $\mathcal{W}_2$ everywhere for simplicity).

Example 1. *Consider (in dimension one for simplicity) the Gaussian kernel interaction potential*

$$W(x,y) = e^{-(x-y)^2} = \sum_{k\in\mathbb{N}} (-1)^k e^{-x^2} \frac{x^k y^k}{k!} e^{-y^2}\,,$$

which appears for instance in the Adaptive Biasing Potential method [6] and more generally in regularized approximations of processes which are influenced by the local density of particles (e.g. [34]). It is of the form (16) *with $a_k(x) = e^{-x^2} x^k/\sqrt{k!}$ for odd $k \in \mathbb{N}$ and $r_k(x) = e^{-x^2} x^k/\sqrt{k!}$ for even $k \in \mathbb{N}$. An elementary analysis and the Stirling formula shows that $\|a_k'\|_\infty = \mathcal{O}(k/2^{k/2})$, which means that* ($\mathcal{W}_2^2$**-curv-**$\mathcal{E}$) *holds.*

Example 2. *Considering the settings of [17], let $\mathbb{H}$ be a Hilbert space with norm $\|\cdot\|_{\mathbb{H}}$, $\varphi : \mathbb{R}^d \to \mathbb{H}$, $R : \mathbb{H} \to \mathbb{R}$ and $V_0 : \mathbb{R}^d \to \mathbb{R}$. Consider*

$$\mathcal{E}(\mu) = \int_{\mathbb{R}^d} V_0(x)\mu(\mathrm{d}x) + R\left(\int_{\mathbb{R}^d} \varphi(x)\mu(\mathrm{d}x)\right)\,,$$

where the integral of an $\mathbb{H}$-valued function is understood in Bochner's sense, and we take the convention that $\mathcal{E}(\mu) = \infty$ if $\int_{\mathbb{R}^d} \|\varphi(x)\|_{\mathbb{H}}\mu(\mathrm{d}x) = \infty$. As discussed in [17], this encompasses many cases of interest in optimization, machine learning and stastistics in high dimension. Assume that the curvature of R is lower-bounded in the sense that there exists $\theta > 0$ such that $R + \theta\|\cdot\|_{\mathbb{H}}^2$ is convex. This implies that ($\mathcal{C}$**-curv-**$\mathcal{E}$) *holds with*

$$\mathcal{C}(\mu_0, \mu_1) = \theta \left\| \int_{\mathbb{R}^d} \varphi \mathrm{d}\mu_0 - \int_{\mathbb{R}^d} \varphi \mathrm{d}\mu_1 \right\|_{\mathbb{H}}^2\,.$$

This generalizes the quadratic finite-dimensional case (17)*. Assuming furthemore that φ is Lipschitz continuous (namely $\|\varphi(x) - \varphi(y)\|_{\mathbb{H}} \leqslant \ell|x-y|$ for all $x, y \in \mathbb{R}^d$), we get* ($\mathcal{W}_2^2$**-curv-**$\mathcal{E}$) *with $\lambda = \theta\ell^2$.*

The following generalizes the entropy sandwich inequality (15) when $\mathcal{C} \neq 0$.

Lemma 3. *Under* ($\mathcal{C}$**-curv-**$\mathcal{E}$)*, for all $\mu_0, \mu_1 \in \mathcal{P}_2(\mathbb{R}^d)$ and $\rho_* \in \mathcal{K}$,*

$$\mathcal{F}(\mu_0) \leqslant \mathcal{F}(\mu_1) + \sigma^2 \mathcal{H}\left(\mu_0 | \Gamma(\mu_0)\right) + \mathcal{C}(\mu_0, \mu_1)\,, \qquad (19)$$

and

$$\mathcal{F}(\rho_*) + \sigma^2 \mathcal{H}\left(\mu_0 | \rho_*\right) \leqslant \mathcal{F}(\mu_0) + \mathcal{C}(\mu_0, \rho_*)\,. \qquad (20)$$

Proof. To prove each inequality, we assume that the right-hand side is finite, otherwise the result is trivial.

Dividing by t and sending t to zero in ($\mathcal{C}$**-curv-**$\mathcal{E}$) yields (using (7)),

$$\int_{\mathbb{R}^d} E_{\mu_1}(\mu_0 - \mu_1) + \mathcal{E}(\mu_1) \leqslant \mathcal{E}(\mu_0) + \mathcal{C}(\mu_0, \mu_1)\,,$$

and thus, interverting the roles of μ_1 and μ_0,

$$\int_{\mathbb{R}^d} E_{\mu_1}(\mu_0 - \mu_1) - \mathcal{C}(\mu_0, \mu_1) \leqslant \mathcal{E}(\mu_0) - \mathcal{E}(\mu_1) \leqslant \int_{\mathbb{R}^d} E_{\mu_0}(\mu_0 - \mu_1) + \mathcal{C}(\mu_0, \mu_1)\,.$$

Then

$$\begin{aligned}
\mathcal{F}(\mu_0) - \mathcal{F}(\mu_1) &\leqslant \int_{\mathbb{R}^d} E_{\mu_0}(\mu_0 - \mu_1) + \mathcal{C}(\mu_0, \mu_1) + \sigma^2\mathcal{H}(\mu_0) - \sigma^2\mathcal{H}(\mu_1) \\
&= \sigma^2\mathcal{H}\left(\mu_0|\Gamma(\mu_0)\right) - \sigma^2\mathcal{H}\left(\mu_1|\Gamma(\mu_0)\right) + \mathcal{C}(\mu_0, \mu_1) \\
&\leqslant \sigma^2\mathcal{H}\left(\mu_0|\Gamma(\mu_0)\right) + \mathcal{C}(\mu_0, \mu_1)\,,
\end{aligned}$$

and

$$\begin{aligned}
\mathcal{F}(\mu_0) - \mathcal{F}(\mu_1) &\geqslant \int_{\mathbb{R}^d} E_{\mu_1}(\mu_0 - \mu_1) - \mathcal{C}(\mu_0, \mu_1) + \sigma^2\mathcal{H}(\mu_0) - \sigma^2\mathcal{H}(\mu_1) \\
&= \sigma^2\mathcal{H}\left(\mu_0|\Gamma(\mu_1)\right) - \sigma^2\mathcal{H}\left(\mu_1|\Gamma(\mu_1)\right) - \mathcal{C}(\mu_0, \mu_1)\,.
\end{aligned}$$

The second term of the right hand side vanishes if $\mu_1 \in \mathcal{K}$, which concludes. □

This result will prove useful in our cases of interest, namely when $\mathcal{F}$ has several critical points. Nevertheless, for now, in the spirit of the previous section, let us discuss how (**U-LSI**) may already give a global convergence in some cases under ($\mathcal{C}$**-curv-**$\mathcal{E}$).

Indeed, assuming a Talagrand type inequality

$$\mathcal{C}(\mu_0, \rho_*) \leqslant \eta'\mathcal{H}(\mu_0|\rho_*)$$

for some $\eta' > 0$ (which, if $\mathcal{C} \leqslant \lambda\mathcal{W}_2^2$ for some $\lambda > 0$, is implied by (**U-LSI**) with $\eta' = 4\eta\lambda$), using (20) and then (19),

$$\begin{aligned}
\mathcal{C}(\mu_0, \rho_*) \leqslant \eta'\mathcal{H}(\mu_0|\rho_*) &\leqslant \frac{\eta'}{\sigma^2}\left(\mathcal{F}(\mu_0) - \mathcal{F}(\rho_*) + \mathcal{C}(\mu_0, \rho_*)\right) \\
&\leqslant \eta'\mathcal{H}\left(\mu_0|\Gamma(\mu_0)\right) + \frac{2\eta'}{\sigma^2}\mathcal{C}(\mu_0, \rho_*)\,.
\end{aligned}$$

Hence, under the condition $2\eta' < \sigma^2$ (i.e., again, high temperature or small interaction), we obtain a non-linear transport inequality

$$\mathcal{C}(\mu_0, \rho_*) \leqslant \frac{\eta'}{1 - 2\eta'/\sigma^2}\mathcal{H}\left(\mu_0|\Gamma(\mu_0)\right)\,. \tag{21}$$

Together with the LSI for ρ_* and (19), this proves (**G-NL-LSI**), since

$$\mathcal{F}(\mu_0) - \mathcal{F}(\rho_*) \leqslant \left(\sigma^2 + \frac{\eta'}{1 - 2\eta'/\sigma^2}\right)\mathcal{H}\left(\mu_0|\Gamma(\mu_0)\right) \leqslant \eta\left(\sigma^2 + \frac{\eta'}{1 - 2\eta'/\sigma^2}\right)\mathcal{I}\left(\mu_0|\Gamma(\mu_0)\right)\,.$$

Notice that, conversely, the non-linear transport inequality (21) with $\mathcal{C} = \mathcal{W}_2^2$ is shown to be a consequence of (**G-NL-LSI**) in [22] (see also Lemma 4 below).

2.4 Local non-linear log-Sobolev inequalities

Consider a non-empty set $\mathcal{A} \subset \mathcal{P}_2(\mathbb{R}^d)$ with $\inf_{\mathcal{A}} \mathcal{F} < \infty$. Instead of the global condition (**G-NL-LSI**), the key ingredient in our approach is to obtain (local) non-linear LSI of the form

$$\forall \rho \in \mathcal{A}\,, \qquad \mathcal{F}(\rho) - \inf_{\mu \in \mathcal{A}} \mathcal{F}(\mu) \leqslant \overline{\eta}\sigma^4 \mathcal{I}\left(\rho|\Gamma(\rho)\right) \tag{NL-LSI}$$

for some $\overline{\eta} > 0$. The fact that it is possible to get this in some subsets $\mathcal{A}$ even in cases where $\mathcal{F}$ has several critical points is the topic of Section 3. For now we assume that this holds in some set $\mathcal{A}$ and we discuss its consequences.

First, since the left hand side of (**NL-LSI**) is non-negative and the right hand side vanishes when $\rho \in \mathcal{K}$, we get that

$$\rho_* \in \mathcal{A} \cap \mathcal{K} \qquad \Rightarrow \qquad \mathcal{F}(\rho_*) = \inf_{\mu \in \mathcal{A}} \mathcal{F}(\mu)\,. \tag{22}$$

Second, the results which were known to hold under (**G-NL-LSI**) still hold as long as the solution of (9) stays within $\mathcal{A}$, as we state in the next lemma. For conciseness we write

$$\mathcal{F}_{\mathcal{A}}(\rho) = \mathcal{F}(\rho) - \inf_{\mu \in \mathcal{A}} \mathcal{F}(\mu)\,.$$

For $\rho_0 \in \mathcal{P}_2(\mathbb{R}^d)$, we write

$$T_{\mathcal{A}}(\rho_0) = \inf\{t \geqslant 0,\ \rho_t \notin \mathcal{A}\}\,,$$

with $(\rho_t)_{t\geqslant 0}$ the solution to (9). The next lemma is the infinite-dimensional version of (4) and (5) and follows the same proof.

Lemma 4. *Assume that* (**NL-LSI**) *holds on $\mathcal{A}$. Then, for all $\rho_0 \in \mathcal{A}$ and $t \leqslant T_{\mathcal{A}}(\rho_0)$,*

$$\mathcal{F}_{\mathcal{A}}(\rho_t) \leqslant e^{-t/\overline{\eta}} \mathcal{F}_{\mathcal{A}}(\rho_0) \tag{23}$$

and

$$\mathcal{W}_2^2(\rho_t, \rho_0) \leqslant 4\overline{\eta} \mathcal{F}_{\mathcal{A}}(\rho_0)\,. \tag{24}$$

Proof. The first inequality is straightforward from (11) and (**NL-LSI**). The proof of the second inequality is the same as the proof of [22, Theorem 3.2] (although a factor 2 seems to disappear wrongly in the latter when using the Benamou-Brenier formula, which explains the difference with our result ; notice that our factor 4 is consistent with the classical linear case [41], or the finite-dimensional settings (5)). Since (9) can be written as the continuity equation

$$\partial_t \rho_t = \nabla \cdot (\rho_t v_t) \qquad v_t := \nabla \frac{\delta \mathcal{F}(\rho_t)}{\delta \mu}\,,$$

by the Benamou-Brenier formulation of the Wasserstein distance [7], for any $0 \le s < t$,

$$\mathcal{W}_2^2(\rho_s, \rho_t) \le (t-s) \int_s^t \int_{\mathbb{R}^d} |v_r|^2 \rho_r \mathrm{d}r = \sigma^4 (t-s) \int_s^t \mathcal{I}(\rho_r|\Gamma(\rho_r))\mathrm{d}r.$$

Thus,

$$\frac{\mathcal{W}_2(\rho_s, \rho_t)}{t-s} \le \sigma^2 \sqrt{\frac{1}{t-s} \int_s^t \mathcal{I}(\rho_r|\Gamma(\rho_r))\mathrm{d}r},$$

so by Lebesgue's differentiation theorem, the metric derivative of $(\rho_t)_{t\geq 0}$ defined in [1, Theorem 1.1.2] satisfies

$$|\rho_t'| \le \sigma^2 \sqrt{\mathcal{I}(\rho_t|\Gamma(\rho_t))},$$

$\mathrm{d}t$-almost everywhere. Since, by [1, Theorem 1.1.2], we have on the other hand

$$\mathcal{W}_2(\rho_s,\rho_t) \leq \int_s^t |\rho_r'|\mathrm{d}r,$$

we get

$$\mathcal{W}_2(\rho_0,\rho_t) \leqslant \sigma^2 \int_0^t \sqrt{\mathcal{I}\left(\rho_s|\Gamma(\rho_s)\right)}\mathrm{d}s\,.$$

Besides, for $t \leqslant T_{\mathcal{A}}(\rho_0)$,

$$\partial_t \mathcal{F}_{\mathcal{A}}(\rho_t) = -\sigma^4 \mathcal{I}\left(\rho_t|\Gamma(\rho_t)\right) \leqslant -\sqrt{\sigma^4 \mathcal{I}\left(\rho_t|\Gamma(\rho_t)\right)\mathcal{F}_{\mathcal{A}}(\rho_t)/\eta}\,,$$

from which

$$\sqrt{\mathcal{F}_{\mathcal{A}}(\rho_t)} - \sqrt{\mathcal{F}_{\mathcal{A}}(\rho_0)} \leqslant -\frac{\sigma^2}{2\sqrt{\eta}}\int_0^t \sqrt{\mathcal{I}\left(\rho_s|\Gamma(\rho_s)\right)}\mathrm{d}s \leqslant -\frac{1}{2\sqrt{\eta}}\mathcal{W}_2(\rho_t,\rho_0)\,.$$

Rearranging the terms and using that $\sqrt{\mathcal{F}_{\mathcal{A}}(\rho_t)} \geqslant 0$ concludes. □

As a consequence, when (**NL-LSI**) holds in a set $\mathcal{A} \subset \mathcal{P}_2(\mathbb{R}^d)$ that contains a unique critical point $\rho_* \in \mathcal{K}$, in order to get the long-time convergence to ρ_* at exponential speed starting from an initial condition $\rho_0 \in \mathcal{A}$, it mainly remains to show that $T_{\mathcal{A}}(\rho_0) = +\infty$. In specific cases, this can be shown by various means, for instance in dimension 1 by monotonicity arguments as in [48], or in general using that the free energy decreases as in [47], etc. In the following, we focus on a self-contained argument which establishes the exponential convergence towards ρ_* under the natural condition that $\mathcal{W}_2(\rho_*,\rho_0)$ is sufficiently small (in particular, without assuming that $\mathcal{F}(\rho_0) < \infty$).

2.5 Auxiliary regularization results

Let us recall two results. The first is proven in [15, Lemma 4.9].

Lemma 5. *Under* (**Lip**) *and* (**U-LSI**)*, for all $\rho \in \mathcal{P}_2(\mathbb{R}^d)$ and $\rho_* \in \mathcal{K}$,*

$$\mathcal{H}\left(\rho|\Gamma(\rho)\right) \leqslant \left(1 + \eta L + \frac{L^2\eta^2}{2}\right)\mathcal{H}(\rho|\rho_*)\,.$$

The second result is from [50, Corollary 4.3].

Proposition 6. *Assume that there exists $\kappa_0,\kappa_1,\kappa_2 > 0$ such that for all $\nu,\mu \in \mathcal{P}_2(\mathbb{R}^d)$ and all $x,y \in \mathbb{R}^d$,*

$$|\nabla E_\mu(0)|^2 + \sigma^2 \leqslant \kappa_0\left(1 + \int_{\mathbb{R}^d}|x|^2\mu(\mathrm{d}x)\right)$$

and

$$-2\left(\nabla E_\mu(x) - \nabla E_\nu(y)\right)\cdot(x-y) \leqslant \kappa_1|x-y|^2 + \kappa_2|x-y|\mathcal{W}_2(\nu,\mu)\,. \tag{25}$$

Then, given two solutions $(\rho_t)_{t\geqslant 0}$ and $(\rho_t')_{t\geqslant 0}$ of (9)*, for all $t > 0$,*

$$\mathcal{H}\left(\rho_t|\rho_t'\right) \leqslant s_t \mathcal{W}_2^2(\rho_0,\rho_0')\,,$$

where

$$s_t = \frac{1}{\sigma^2}\left(\frac{\kappa_1}{1-e^{-\kappa_1 t}} + \frac{1}{2}t\kappa_2^2 e^{2(\kappa_1+\kappa_2)t}\right)\,.$$

Notice that, under the uniform Lipschitz condition (**Lip**), since we can bound $|\nabla E_\mu(0)| \leqslant |\nabla E_{\delta_0}(0)| + L\mathcal{W}_2(\mu,\delta_0)$ for any $\mu \in \mathcal{P}_2(\mathbb{R}^d)$, the assumptions of Proposition 6 are fulfilled if additionally there exists $\kappa_1 > 0$ such that the one-sided Lipschitz condition

$$\forall \mu \in \mathcal{P}_2(\mathbb{R}^d), x, y \in \mathbb{R}^d, \qquad 2\left(\nabla E_\mu(x) - \nabla E_\mu(y)\right) \cdot (x-y) \geqslant -\kappa_1 |x-y|^2 \qquad \textbf{(o-s-Lip)}$$

holds, and then we can take $\kappa_2 = 2L$ to have (25).

Corollary 7. *Assume* ($\mathcal{W}_2^2$**-curv-**$\mathcal{E}$)*,* (**Lip**)*,* (**o-s-Lip**) *and* (**U-LSI**) *for some* $\lambda, L, \kappa_1, \eta > 0$*. Then, for all* $(\rho_t)_{t\geqslant 0}$ *solution to* (9)*, all* $\rho_* \in \mathcal{K}$ *and all* $t > 0$*,*

$$\mathcal{F}(\rho_t) - \mathcal{F}(\rho_*) \leqslant q_t \mathcal{W}_2^2(\rho_0, \rho_*)$$

with

$$q_t = \left(1 + \eta L + \frac{4\eta\lambda}{\sigma^2} + \frac{L^2\eta^2}{2}\right)\left(\frac{\kappa_1}{1 - e^{-\kappa_1 t}} + \frac{1}{2} t L^2 e^{(2\kappa_1 + 4L)t}\right). \tag{26}$$

Notice that q_t is of order $1/t$ as $t \to 0$. See [1, Theorem 4.0.4] for a similar result.

Proof. Using (19) in Lemma 3 with $\mathcal{C} = \lambda \mathcal{W}_2^2$ and then Lemma 5 and the Talagrand inequality (18) for $\rho_* = \Gamma(\rho_*)$ implied by (**U-LSI**),

$$\begin{aligned} \mathcal{F}(\rho_t) - \mathcal{F}(\rho_*) &\leqslant \sigma^2 \mathcal{H}\left(\rho_t | \Gamma(\rho_t)\right) + \lambda \mathcal{W}_2^2(\rho_t, \rho_*) \\ &\leqslant \sigma^2 \left(1 + \eta L + \frac{4\eta\lambda}{\sigma^2} + \frac{L^2\eta^2}{2}\right) \mathcal{H}(\rho_t | \rho_*). \end{aligned}$$

Since $t \mapsto \rho_*$ solves (9), conclusion follows from Proposition 6. □

2.6 Conclusion

For $\mu \in \mathcal{P}_2(\mathbb{R}^d)$ and $r > 0$, write $\mathcal{B}_{\mathcal{W}_2}(\mu, r) = \{\nu \in \mathcal{P}_2(\mathbb{R}^d), \mathcal{W}_2(\nu, \mu) \leqslant r\}$.

Assumption 5. *The conditions* ($\mathcal{W}_2^2$**-curv-**$\mathcal{E}$)*,* (**Lip**)*,* (**o-s-Lip**) *and* (**U-LSI**) *hold for some* $\lambda, L, \kappa_1, \eta > 0$*, and* (**NL-LSI**) *holds on some non-empty set* $\mathcal{A} \subset \mathcal{P}_2(\mathbb{R}^d)$ *for some* $\overline{\eta} > 0$*. Furthermore there exist* $\rho_* \in \mathcal{A} \cap \mathcal{K}$ *and* $\delta > 0$ *such that* $\mathcal{B}_{\mathcal{W}_2}(\rho_*, \delta) \subset \mathcal{A}$*.*

Theorem 8. *Under Assumption 5, set*

$$\delta' = \frac{\delta}{2(2\sqrt{\overline{\eta}q_1} + e^{(\kappa_1/2 + L)})},$$

where q_1 *is given by* (26) *(with* $t = 1$*). Then, for all* $\rho_0 \in \mathcal{B}_{\mathcal{W}_2}(\rho_*, \delta')$*,* $T_{\mathcal{A}}(\rho_0) = \infty$ *and in particular*

$$\mathcal{F}_{\mathcal{A}}(\rho_t) \leqslant e^{-t/\overline{\eta}} \mathcal{F}_{\mathcal{A}}(\rho_0)$$

for all $t \geqslant 0$*. Furthermore, if we assume additionally that* $\mathcal{B}_{\mathcal{W}_2}(\rho_*, \delta) \cap \mathcal{K} = \{\rho_*\}$*, then:*

- *The following non-linear Talagrand inequality holds:*

$$\mathcal{W}_2^2(\rho_0, \rho_*) \leqslant 4\overline{\eta} \mathcal{F}_{\mathcal{A}}(\rho_0). \tag{27}$$

- *Setting* $C = e^{1/\overline{\eta}} \max(4\overline{\eta} q_1, e^{\kappa_1 + 2L})$*, for all* $t \geqslant 0$*,*

$$\mathcal{W}_2^2(\rho_t, \rho_*) \leqslant C e^{-t/\overline{\eta}} \mathcal{W}_2^2(\rho_0, \rho_*) \tag{28}$$

$$\mathcal{F}_{\mathcal{A}}(\rho_t) \leqslant e^{-t/\overline{\eta}} \min\left(\mathcal{F}_{\mathcal{A}}(\rho_0)\ ,\ q_{\min(t,1)} e^{1/\overline{\eta}} \mathcal{W}_2^2(\rho_0, \rho_*)\right) \tag{29}$$

$$\sigma^2 \mathcal{H}(\rho_t | \rho_*) \leqslant \mathcal{F}_{\mathcal{A}}(\rho_t) + \lambda \mathcal{W}_2^2(\rho_t, \rho_*). \tag{30}$$

The last inequality on the relative entropy is just (20) in Lemma 3, we simply recall it as it gives here the exponential convergence of $\mathcal{H}(\rho_t|\rho_*)$ (hence of $\|\rho_t - \rho_*\|_{TV}$) to zero.

Remark 1. *Write $\mathcal{A}_\infty(\rho_*) = \{\rho_0 \in \mathcal{A} : T_A(\rho_0) = \infty$ and $\mathcal{W}_2(\rho_t, \rho_*) \to 0$ as $t \to \infty\}$. Then it is clear by following the proof of Theorem 8 that its conclusion (namely* (27), (28) *and* (29)*) hold for any initial condition $\rho_0 \in \mathcal{A}_\infty(\rho_*)$. Moreover, for any $\rho_0 \in \mathcal{P}_2(\mathbb{R}^d)$ such that $\mathcal{W}_2(\rho_t, \rho_*) \to 0$ as $t \to \infty$, there exists a $t_* > 0$ such that $\rho_{t_*} \in \mathcal{B}_{\mathcal{W}_2}(\rho_*, \delta) \subset \mathcal{A}_\infty(\rho_*)$, from which there exists $C_0 > 0$ such that for all $t \geqslant 1$,*

$$\mathcal{W}_2^2(\rho_t, \rho_*) + \mathcal{H}(\rho_t|\rho_*) + \mathcal{F}(\rho_t) - \mathcal{F}(\rho_*) \leqslant C_0 e^{-t/\overline{\eta}} . \tag{31}$$

Proof. Given a solution $(\rho_t)_{t\geqslant 0}$ of (8), we consider the time-inhomogeneous Markov diffusion process

$$\mathrm{d}X_t = -\nabla E_{\rho_t}(X_t)\mathrm{d}t + \sqrt{2}\sigma \mathrm{d}B_t$$

where B is a Brownian motion and X_0 is distributed according to ρ_0, so that $X_t \sim \rho_t$ for all $t \geqslant 0$. Existence of strong solutions for this SDE is justified by the fact $(t,x) \mapsto \nabla E_{\rho_t}$ is continuous in time (by continuity of $t \mapsto \rho_t$ and (**Lip**)) and $\mathcal{C}^1$ in x, and explosion in finite time is prevented by (**o-s-Lip**). Considering a synchronous coupling (X,Y) (i.e. using the same Brownian motion for two processes) of such diffusions, the first one being associated with an arbitrary solution ρ, the second being associated to the stationary solution $t \mapsto \rho_*$, we get that

$$\partial_t |X_t - Y_t|^2 = -2(X_t - Y_t) \cdot (\nabla E_{\rho_t}(X_t) - \nabla E_{\rho_*}(Y_t)) \leqslant (\kappa_1 + L)|X_t - Y_t|^2 + L\mathcal{W}_2^2(\rho_t, \rho_*) ,$$

where we used (**Lip**) and (**o-s-Lip**). Taking the expectation, using the Gronwall Lemma, that $\mathcal{W}_2^2(\rho_t, \rho_*) \leqslant \mathbb{E}(|X_t - Y_t|^2)$ and taking the infimum over the coupling of the initial conditions, we obtain

$$\mathcal{W}_2(\rho_t, \rho_*) \leqslant e^{(\kappa_1/2+L)t} \mathcal{W}_2(\rho_0, \rho_*) . \tag{32}$$

In view of the definition of δ', this implies that for $\rho_0 \in \mathcal{B}_{\mathcal{W}_2}(\rho_*, \delta')$, $T_{\mathcal{A}}(\rho_0) \geqslant 1$ and $\rho_1 \in \mathcal{A}$. On the other hand, using Corollary 7 and the decay of the free energy along time, for all $t \geqslant 1$,

$$\mathcal{F}(\rho_t) - \mathcal{F}(\rho_*) \leqslant q_1 (\delta')^2 .$$

Hence, for all $t \in [1, T_{\mathcal{A}}(\rho_0)]$, using (24),

$$\mathcal{W}_2(\rho_t, \rho_*) \leqslant \mathcal{W}_2(\rho_t, \rho_1) + \mathcal{W}_2(\rho_1, \rho_*) \leqslant \delta' \left(2\sqrt{\overline{\eta} q_1} + e^{(\kappa_1/2+L)}\right) = \frac{\delta}{2} .$$

By contradiction, since $t \mapsto \mathcal{W}_2(\rho_t, \rho_*)$ is continuous, we get that $T_{\mathcal{A}}(\rho_0) = \infty$.

Now, assume that $\mathcal{B}_{\mathcal{W}_2}(\rho_*, \delta) \cap \mathcal{K} = \{\rho_*\}$. Since Equation (9) admits $\mathcal{F}$ as a strict Lyapunov function, the LaSalle invariance principle established in [10] shows that the $\mathcal{W}_2$ distance between ρ_t and $\mathcal{K}$ vanishes as $t \to \infty$. Since the trajectory stays within $\mathcal{B}_{\mathcal{W}_2}(\rho_*, \delta/2)$ and thus at a distance at least $\delta/2$ of $\mathcal{K} \setminus \{\rho_*\}$, this implies the convergence towards ρ_*. We can then let $t \to \infty$ in (24) to get (27), that we use then to bound, for $t \geqslant 1$,

$$\mathcal{W}_2^2(\rho_t, \rho_*) \leqslant 4\overline{\eta}\mathcal{F}_{\mathcal{A}}(\rho_t) \leqslant 4\overline{\eta} e^{-(t-1)/\overline{\eta}} \mathcal{F}_{\mathcal{A}}(\rho_1) \leqslant 4\overline{\eta} e^{-(t-1)/\overline{\eta}} q_1 \mathcal{W}_2^2(\rho_t, \rho_*) .$$

For $t \leqslant 1$, we simply use (32). The bound on $\mathcal{F}_{\mathcal{A}}(\rho_t)$ is straightforward from (23) and Corollary 7. □

3 Local non-linear LSI in a parametric case

In this section we focus on the settings of Example 2, where we recall that

$$\mathcal{E}(\mu) = \int_{\mathbb{R}^d} V_0(x)\mu(\mathrm{d}x) + R\left(\int_{\mathbb{R}^d} \varphi(x)\mu(\mathrm{d}x)\right). \tag{33}$$

We assume that $V_0 \in \mathcal{C}^2(\mathbb{R}^d, \mathbb{R})$, $R \in \mathcal{C}^1(\mathbb{H}, \mathbb{R})$ and $\varphi \in \mathcal{C}^2(\mathbb{R}^d, \mathbb{H})$. In this case, denoting by $\nabla_{\mathbb{H}}$ and $\langle \cdot, \cdot \rangle_{\mathbb{H}}$ the gradient and scalar product over $\mathbb{H}$,

$$E_\rho(x) = V_0(x) + \langle \nabla_{\mathbb{H}} R(\varphi(\rho)), \varphi(x) \rangle_{\mathbb{H}},$$

with $\varphi(\rho) := \int_{\mathbb{R}^d} \varphi\rho \in \mathbb{H}$, and Assumption 2 is satisfied if, for any $x \in \mathbb{R}^d$, $\|\varphi(x)\|_{\mathbb{H}} \leq C(1 + |x|^2)$.

The general idea that we develop in this section is that, since the non-linearity is somehow parametrized by $\varphi(\rho)$, information for the Wasserstein gradient flow can be deduced from standard analysis of fixed point or optimization on Hilbert spaces. General considerations are gathered in Section 3.1, which are then applied to the granular media case (10) in Section 3.2.

Remark 2. *If R is convex then $\mathcal{E}$ is functional convex. As discussed in Section 2.2, thanks to Lemma 3, in that case* (**U-LSI**) *implies* (**G-NL-LSI**) *and then global convergence at constant rate towards a unique stationary solution. Hence, this is not the case we are interested in.*

3.1 General results

3.1.1 The associated fixed-point problem

In the case (33), we get that $\Gamma(\rho) = \rho_{\varphi(\rho)}$ where, for $\psi \in \mathbb{H}$,

$$\rho_\psi(x) \propto \exp\left(-\frac{1}{\sigma^2}\left[V_0(x) + \langle \nabla_{\mathbb{H}} R(\psi), \varphi(x) \rangle_{\mathbb{H}}\right]\right). \tag{34}$$

Let us assume that $\mathcal{F}(\rho_\psi) < \infty$ for all $\psi \in \mathbb{H}$. This implies in particular that

$$f(\psi) = \varphi(\rho_\psi) = \int_{\mathbb{R}^d} \varphi(x)\rho_\psi(x)\mathrm{d}x \tag{35}$$

is well-defined, and we see that $\varphi(\Gamma(\rho)) = f(\varphi(\rho))$. In particular, any $\rho_* \in \mathcal{K}$ satisfies $\rho_* = \rho_{\varphi(\rho_*)}$. Thus, denoting by $\mathcal{K}' = \{\psi \in \mathbb{H},\ f(\psi) = \psi\}$ the set of fixed-points of f,

$$\mathcal{K} = \{\rho_\psi,\ \psi \in \mathcal{K}'\}.$$

One may hope that the stability properties of $\rho_* = \rho_{\psi_*} \in \mathcal{K}$ as a stationary solution of the (continuous-time) Wasserstein gradient descent is related to the stability of ψ_* as a fixed-point of the (discrete-time) dynamical system $\psi \mapsto f(\psi)$ on $\mathbb{H}$, which is a classical question. In fact, we can find examples where ρ_{ψ_*} is stable although ψ_* is not (see the granular media equation with repulsive interaction in Figure 3 below). However, on the contrary, the next result shows that (**NL-LSI**) holds in a neighborhood ρ_{ψ_*} (which is thus stable thanks to Theorem 8) if ψ_* is a geometrically attracting fixed-point.

Proposition 9. *Let $\psi_* \in \mathcal{K}'$, $\mathcal{A}' \subset \mathbb{H}$ and $\alpha \in [0, 1)$ be such that for all $\psi \in \mathcal{A}'$,*

$$\|f(\psi) - \psi_*\|_{\mathbb{H}} \leqslant \alpha\|\psi - \psi_*\|_{\mathbb{H}}. \tag{36}$$

Then, for any $\psi \in \mathcal{A}'$,

$$\|\psi - \psi_*\|_{\mathbb{H}} \leqslant \frac{1}{1-\alpha}\|f(\psi) - \psi\|_{\mathbb{H}} \,. \tag{37}$$

Assuming moreover that φ is ℓ-Lipschitz continuous, then, for all $\rho \in \mathcal{P}_2(\mathbb{R}^d)$ such that $\varphi(\rho) \in \mathcal{A}'$,

$$\|\varphi(\rho) - \psi_*\|_{\mathbb{H}} \leqslant \frac{\ell}{1-\alpha}\mathcal{W}_2\left(\rho, \Gamma(\rho)\right) \,.$$

Finally, assuming furthermore (**U-LSI**) *and that $R + \theta\|\cdot\|_{\mathbb{H}}^2$ is convex for some $\theta > 0$, then, for all $\rho \in \mathcal{P}_2(\mathbb{R}^d)$ such that $\varphi(\rho) \in \mathcal{A}'$,*

$$\mathcal{F}(\rho) - \mathcal{F}(\rho_{\psi_*}) \leqslant \eta\left(\sigma^2 + \frac{4\eta\theta\ell^2}{(1-\alpha)^2}\right)\mathcal{I}\left(\rho|\Gamma(\rho)\right) \,,$$

i.e. (**NL-LSI**) *holds over $\mathcal{A} = \{\rho \in \mathcal{P}_2,\ \varphi(\rho) \in \mathcal{A}'\}$.*

Proof. The first inequality (37) simply follows from the triangular inequality

$$\|\psi - \psi_*\|_{\mathbb{H}} \leqslant \|\psi - f(\psi)\|_{\mathbb{H}} + \|f(\psi) - \psi_*\|_{\mathbb{H}} \leqslant \|\psi - f(\psi)\|_{\mathbb{H}} + \alpha\|\psi - \psi_*\|_{\mathbb{H}} \,.$$

Applying this to $\psi = \varphi(\rho)$ reads

$$\|\varphi(\rho) - \psi_*\|_{\mathbb{H}} \leqslant \frac{1}{1-\alpha}\|\varphi(\rho) - \varphi\left(\Gamma(\rho)\right)\|_{\mathbb{H}} \leqslant \frac{\ell}{1-\alpha}\mathcal{W}_2\left(\rho, \Gamma(\rho)\right) \,.$$

Finally, applying Lemma 3 (with $\mathcal{C}(\mu_0, \mu_1) = \theta\|\varphi(\mu_0) - \varphi(\mu_1)\|_{\mathbb{H}}^2$ as in Example 2),

$$\begin{aligned}
\mathcal{F}(\rho) - \mathcal{F}(\rho_{\psi_*}) &\leqslant \sigma^2\mathcal{H}\left(\rho|\Gamma(\rho)\right) + \theta\|\varphi(\rho) - \psi_*\|_{\mathbb{H}}^2 \\
&\leqslant \sigma^2\mathcal{H}\left(\rho|\Gamma(\rho)\right) + \frac{\theta\ell^2}{(1-\alpha)^2}\mathcal{W}_2^2\left(\rho, \Gamma(\rho)\right) \\
&\leqslant \left(\sigma^2 + \frac{4\eta\theta\ell^2}{(1-\alpha)^2}\right)\mathcal{H}\left(\rho|\Gamma(\rho)\right)
\end{aligned}$$

thanks to Talagrand inequality (18). Conclusion follows from (**U-LSI**). □

Remark 3. *The constant $\overline{\eta}$ in* (**NL-LSI**) *obtained when applying Proposition 9, and thus the quantitative convergence estimates obtained when applying Theorem 8, are explicit in terms of the constants in the assumptions. In particular there is no additionnal dependency in the dimension.*

We have thus identified the following general conditions:

Assumption 6. *There exist $\theta, \ell, \eta > 0$ such that φ is ℓ-Lipschitz continuous, $R + \theta\|\cdot\|_{\mathbb{H}}^2$ is convex and* (**U-LSI**) *holds. Moreover, $\mathcal{F}(\rho_\psi) < \infty$ for any $\psi \in \mathbb{H}$.*

Remark 4. *As discussed in Example 2, these conditions imply* ($\mathcal{W}_2^2$**-curv-**$\mathcal{E}$) *with $\lambda = \theta\ell^2$. Moreover, since*

$$\nabla E_\rho(x) = \nabla V_0(x) + \langle\nabla_{\mathbb{H}} R\left(\varphi(\rho)\right), \nabla\varphi(x)\rangle_{\mathbb{H}} \,,$$

assuming that $\nabla_{\mathbb{H}} R$ is L'-Lipschitz continuous, we get that, for any $x \in \mathbb{R}^d$,

$$|\nabla E_{\mu_0}(x) - \nabla E_{\mu_1}(x)| \leqslant \ell L'\|\varphi(\mu_0) - \varphi(\mu_1)\|_{\mathbb{H}} \leqslant \ell^2 L'\mathcal{W}_2(\mu_0, \mu_1) \,,$$

which is (**Lip**).

Under the general conditions of Assumption 6, given a specific problem, what remains to be done is to establish the contraction (36). A simple way to get it locally at some fixed point ψ_* is to check that $|\nabla_{\mathbb{H}} f(\psi_*)|$, the operator norm of the Jacobian matrix of f, is strictly less than 1 at ψ_*. We end up with the following.

Corollary 10. *Under Assumption 6, let $\psi_* \in \mathcal{K}'$. Assume that f is differentiable at ψ_* with $|\nabla_{\mathbb{H}} f(\psi_*)| < 1$. Then there exist $\delta, \overline{\eta} > 0$ such that* (**NL-LSI**) *holds on $\mathcal{A} = \mathcal{B}_{\mathcal{W}_2}(\rho_{\psi_*}, \delta)$.*

Proof. Using that ψ_* is a fixed point of f and a Taylor expansion,

$$\|f(\psi) - \psi_*\|_{\mathbb{H}} = \|\nabla_{\mathbb{H}} f(\psi_*)(\psi - \psi_*)\|_{\mathbb{H}} + o(\|\psi - \psi_*\|_{\mathbb{H}}) \leqslant \alpha \|\psi - \psi_*\|_{\mathbb{H}}$$

for some $\alpha < 1$ uniformly over some neighborhood $\mathcal{N}$ of ψ_*. The function φ being Lipschitz continuous, $\rho \mapsto \varphi(\rho)$ is Lipschitz continuous for the $\mathcal{W}_2$ distance, which means that $\varphi(\rho)$ lies in $\mathcal{N}$ for all $\rho \in \mathcal{B}_{\mathcal{W}_2}(\rho_{\psi_*}, \delta)$ for δ small enough. Conclusion follows from Proposition 9. □

3.1.2 Decomposition of the free energy

The difference between (**U-LSI**) and (**G-NL-LSI**) obviously lies in the difference between $\mathcal{H}(\rho|\Gamma(\rho))$ and $\mathcal{F}(\rho)$. In the case (33), we can decompose

$$\mathcal{F}(\rho) = \sigma^2 \mathcal{H}\left(\rho|\Gamma(\rho)\right) + g\left(\varphi(\rho)\right)$$

where

$$g(\psi) = R\left(\psi\right) - \langle \nabla_{\mathbb{H}} R\left(\psi\right), \psi\rangle_{\mathbb{H}} - \sigma^2 \ln \int_{\mathbb{R}^d} e^{-\frac{1}{\sigma^2}[V_0(x) + \langle \nabla_{\mathbb{H}} R(\psi), \varphi(x)\rangle_{\mathbb{H}}]} \mathrm{d}x\,. \tag{38}$$

This second part of the free energy only depends on ρ through the parameter $\varphi(\rho)$.

Remark 5. *If R is concave, then $M(\psi_1, \psi_2) := R(\psi_2) - R(\psi_1) + \langle \nabla_{\mathbb{H}} R(\psi_2), \psi_1 - \psi_2\rangle_{\mathbb{H}} \geqslant 0$ (this is the so-called Bregman divergence associated with $-R$). Assuming furthermore that $V_0(x) = V(x) - R(\varphi(x))$ with V such that e^{-V/σ^2} is integrable, and that $M(\psi_1, \psi_2) \to +\infty$ when $\psi_2 \to \infty$, we get by dominated convergence that*

$$e^{-\frac{1}{\sigma^2} g(\psi)} = \int_{\mathbb{R}^d} e^{-\frac{1}{\sigma^2}[V(x) + M(\varphi(x), \psi)]} \mathrm{d}x \underset{\|\psi\|_{\mathbb{H}} \to +\infty}{\longrightarrow} 0\,.$$

As a consequence, g is lower bounded and goes to infinity at infinity. Moreover, it is continuous, and thus in particular when $\mathbb{H}$ has finite dimension then g reaches its minimum. Notice that, in the quadratic case where $R(\psi) = -\theta\|\psi\|_{\mathbb{H}}^2$, $M(\psi_1, \psi_2) = \theta\|\psi_1 - \psi_2\|_{\mathbb{H}}^2$.

Assuming that $R \in \mathcal{C}^2(\mathbb{H}, \mathbb{R})$,

$$\nabla_{\mathbb{H}} g(\psi) = \nabla_{\mathbb{H}}^2 R(\psi)\left(f(\psi) - \psi\right)\,, \tag{39}$$

where we used that

$$\frac{\int_{\mathbb{R}^d} \nabla_{\mathbb{H}}^2 R(\psi) \varphi e^{-\frac{1}{\sigma^2}[V_0(x) + \langle \nabla_{\mathbb{H}} R(\psi), \varphi(x)\rangle_{\mathbb{H}}]} \mathrm{d}x}{\int_{\mathbb{R}^d} e^{-\frac{1}{\sigma^2}[V_0(x) + \langle \nabla_{\mathbb{H}} R(\psi), \varphi(x)\rangle_{\mathbb{H}}]} \mathrm{d}x} = \nabla_{\mathbb{H}}^2 R(\psi) \varphi\left(\rho_\psi\right) = \nabla_{\mathbb{H}}^2 R(\psi) f(\psi)\,.$$

As a consequence, if ψ_* is a critical point of g such that $\nabla_{\mathbb{H}}^2 R(\psi_*)$ is non-singular (which is in particular the case for all critical points of g if R is strictly convex or concave, which covers many cases of interest) then necessarily $\psi_* \in \mathcal{K}'$. In that case, if moreover ψ_* is a local minimizer of g, then ρ_{ψ_*} is a minimizer of both parts of the free energy, $\rho \mapsto \mathcal{H}(\rho|\Gamma(\rho))$ and $\rho \mapsto g\left(\varphi(\rho)\right)$ (at least locally in the latter). It is thus a good candidate to be a stable stationary solution for the Wasserstein gradient descent for the free energy. This is indeed the case:

Proposition 11. *Assume* ($\mathcal{W}_2^2$**-curv-**$\mathcal{E}$), (**Lip**), (**o-s-Lip**) *and* (**U-LSI**) *for some* $\lambda, L, \kappa_1, \eta > 0$ *and* φ *is* ℓ*-Lipschitz continuous. Let* $\psi_* \in \mathcal{K}'$ *be a proper isolated local minimizer of* g*, in the sense that for all* $\varepsilon > 0$ *small enough,* $\zeta(\varepsilon) := \inf\{g(\psi) - g(\psi_*),\ \psi \in \mathbb{H},\ \|\psi - \psi_*\|_{\mathbb{H}} = \varepsilon\} > 0$ *and* $\mathcal{K}' \cap \mathcal{B}(\psi_*, \varepsilon) = \{\psi_*\}$*. Then there exists* $\delta > 0$ *such that for all initial conditions* $\rho_0 \in \mathcal{B}_{\mathcal{W}_2}(\rho_{\psi_*}, \delta)$*, the flow* (9) *converges in long-time to* ρ_{ψ_*}*.*

Proof. Let $\varepsilon > 0$ be small enough so that $\zeta(\varepsilon) > 0$ and $\mathcal{K}' \cap \mathcal{B}(\psi_*, \varepsilon) = \{\psi_*\}$. Notice that, for $\psi, \tilde{\psi} \in \mathcal{K}'$, $\|\psi - \tilde{\psi}\|_{\mathbb{H}} = \|\varphi(\rho_\psi) - \varphi(\rho_{\tilde{\psi}})\|_{\mathbb{H}} \leq \ell \mathcal{W}_2(\rho_\psi, \rho_{\tilde{\psi}})$, so that the second condition implies that $\mathcal{K} \cap \mathcal{B}_{\mathcal{W}_2}(\rho_{\psi_*}, \varepsilon/\ell) = \{\rho_{\psi_*}\}$, and thus ρ_{ψ_*} is an isolated critical point of $\mathcal{F}$. For any $\rho \in \mathcal{P}_2$ with $\|\varphi(\rho) - \psi_*\|_{\mathbb{H}} = \varepsilon$,

$$\mathcal{F}(\rho) \geqslant g\left(\varphi(\rho)\right) \geqslant g(\psi_*) + \zeta(\varepsilon) = \mathcal{F}(\rho_{\psi_*}) + \zeta(\varepsilon)\,.$$

As a consequence, given a solution $(\rho_t)_{t\geqslant 0}$ of (1), if there is a time t_0 such that $\|\varphi(\rho_{t_0}) - \psi_*\|_{\mathbb{H}} < \varepsilon$ and $\mathcal{F}(\rho_{t_0}) - \mathcal{F}(\rho_{\psi_*}) < \zeta(\varepsilon)$ then the monotonicity of the free energy along the flow together with the continuity of $t \mapsto \|\varphi(\rho_t) - \psi_*\|_{\mathbb{H}}$ implies that $\|\varphi(\rho_t) - \psi_*\|_{\mathbb{H}} < \varepsilon$ for all $t \geqslant t_0$. Conclusion would thus follow from LaSalle invariance principle [10] using that ρ_{ψ_*} is isolated.

Hence, let us prove that the previous conditions are met for initial conditions that are sufficiently close to ρ_{ψ_*}. We follow arguments similar to the proof of Theorem 8. Indeed, for all $t \geqslant 0$, using (32) ,

$$\|\varphi(\rho_t) - \psi_*\|_{\mathbb{H}} \leq \ell e^{(\kappa_1/2+L)t} \mathcal{W}_2\left(\rho_0, \rho_{\psi_*}\right)\,.$$

By taking $\delta < \varepsilon e^{-\kappa_1/2-L}/\ell$, we ensure that $\|\varphi(\rho_{t_0}) - \psi_*\|_{\mathbb{H}} < \varepsilon$ at time $t_0 = 1$ for all $\rho_0 \in \mathcal{B}_{\mathcal{W}_2}(\rho_{\psi_*}, \delta)$. On the other hand, thanks to Corollary 7, by taking δ small enough, we can also ensure that $\mathcal{F}(\rho_{t_0}) - \mathcal{F}(\rho_{\psi_*}) < \zeta(\varepsilon)$ with such initial conditions. □

Remark 6. *The argument that initial conditions that start with a free energy sufficiently close to* $\mathcal{F}(\rho_*)$ *will stay within a ball centered at* ρ_* *has been used e.g. in [47, 4]. Here we combine it with the small-time regularization result of Corollary 7 to get a result only in terms of a* $\mathcal{W}_2$ *ball.*

In fact, we can now reinterpret Proposition 9 in light of Proposition 11. Indeed, differentiating again (39) (assuming suitable regularity) at some $\psi_* \in \mathcal{K}'$ yields

$$\nabla^2_{\mathbb{H}} g(\psi_*) = \nabla^2_{\mathbb{H}} R(\psi_*)\left(\nabla_{\mathbb{H}} f(\psi_*) - I\right) \tag{40}$$

(where we used that $f(\psi_*) = \psi_*$). For brevity, denote $\nabla^2 R_* = \nabla^2_{\mathbb{H}} R(\psi_*)$. For any $u \in \mathbb{H}$,

$$\langle u, \nabla^2_{\mathbb{H}} g(\psi_*) u\rangle_{\mathbb{H}} = \langle u, \nabla^2 R_* \nabla_{\mathbb{H}} f(\psi_*) u\rangle_{\mathbb{H}} - \langle u, \nabla^2 R_* u\rangle_{\mathbb{H}}\,.$$

Assuming that $\nabla^2 R_*$ is negative definite, we can consider the norm $\|u\|_* = \sqrt{-\langle u, \nabla^2 R_* u\rangle_{\mathbb{H}}}$ and the associated operator norm $\|\nabla_{\mathbb{H}} f(\psi_*)\|_*$. Then,

$$\langle u, \nabla^2_{\mathbb{H}} g(\psi_*) u\rangle_{\mathbb{H}} \geqslant \left(1 - \|\nabla_{\mathbb{H}} f(\psi_*)\|_*\right) \|u\|_*^2\,.$$

We get that ψ_* is a non-degenerate local minimizer of g if $\|\nabla_{\mathbb{H}} f\|_* < 1$. On the other hand, if $\|\nabla_{\mathbb{H}} f\|_* < 1$ and $(\nabla^2 R_*)^{-1}$ is bounded, then we can assume that $\|\cdot\|_* = \|\cdot\|_{\mathbb{H}}$ (which amounts to the linear change of coordinates given by $(\nabla^2 R_*)^{-1/2}$) and apply Proposition 9 or, alternatively, we can use the contraction (36) expressed with norm $\|\cdot\|_*$ to get (37) with the same norm and then with the initial $\|\cdot\|_{\mathbb{H}}$ by the equivalence of the norms and then proceed with Proposition 9). At least, this suggests that in some cases, in order to apply Proposition 9 at some point ψ_*, it may be more natural to work with the norm $\|\cdot\|_*$ associated to $\nabla^2 R_*$ (when the latter is negative definite).

Besides, we compute that

$$\nabla_{\mathbb{H}} f(\psi) = -\frac{1}{\sigma^2}\left[\int_{\mathbb{R}^d} \varphi\left(\nabla^2_{\mathbb{H}} R(\psi)\varphi\right)^T \rho_\psi - f(\psi)\left(\nabla^2_{\mathbb{H}} R(\psi) f(\psi)\right)^T\right],$$

from which, for $u_1, u_2 \in \mathbb{H}$, writing $w_i = \nabla^2 R_* u_i$ for $i = 1, 2$,

$$\begin{aligned}
\langle u_1, \nabla^2 R_* \nabla_{\mathbb{H}} f(\psi_*) u_2 \rangle_{\mathbb{H}} &= -\frac{1}{\sigma^2}\left\langle w_1, \left[\int_{\mathbb{R}^d} \varphi\varphi^T \rho_{\psi_*} - f(\psi_*)\left(f(\psi_*)\right)^T\right] w_2 \right\rangle_{\mathbb{H}} \\
&= -\frac{1}{\sigma^2}\langle w_1, \mathrm{covar}_{\rho_{\psi_*}}\left(\varphi(X)\right) w_2 \rangle_{\mathbb{H}}\,. \qquad (41)
\end{aligned}$$

In particular, we see that $\nabla_{\mathbb{H}} f(\psi_*)$ is symmetric and nonnegative for the scalar product associated to $\|\cdot\|_*$, from which

$$\|\nabla_{\mathbb{H}} f(\psi_*)\|_* = \sup_{u\neq 0} \frac{\langle u, \nabla^2 R_* \nabla_{\mathbb{H}} f(\psi_*) u\rangle_{\mathbb{H}}}{\langle u, \nabla^2 R_* u\rangle_{\mathbb{H}}}\,. \qquad (42)$$

Furthermore, if we work with suitable coordinates in order to enforce that $\nabla^2 R_* = -I$, we get that, for $\psi_* \in \mathcal{K}'$,

$$\nabla^2_{\mathbb{H}} g(\psi_*) = I - \frac{1}{\sigma^2}\mathrm{covar}_{\rho_{\psi_*}}(\varphi(X))\,.$$

3.1.3 Degenerate minima

There is no convergence rate in Proposition 11, and we cannot expect one in general if we do not assume that ψ_* is a non-degenerate local minimizer of g. Assuming again that $\nabla^2_{\mathbb{H}} R(\psi_*)$ is non-singular, in view of (40), if $\nabla^2_{\mathbb{H}} g(\psi_*)$ is singular then $|\nabla_{\mathbb{H}} f(\psi_*)| \geqslant 1$ (since $\nabla_{\mathbb{H}} f(\psi_*)u = u$ for $u \in \ker\nabla^2_{\mathbb{H}} g(\psi_*)$, and thus this does not depend on the norm we use). In other words, degenerate minima of g fall in the limit case where we cannot apply Proposition 9 anymore. Then we can weaken the contraction condition (36) to allow for high-order stable fixed points. The proof of the following is the same as the one of Proposition 9 (hence omitted).

Proposition 12. *Let $\psi_* \in \mathcal{K}'$, $\mathcal{A}' \subset \mathbb{H}$ and $\beta > 0, \nu \in (0,1)$ be such that for all $\psi \in \mathcal{A}'$,*

$$\|\psi - \psi_*\|_{\mathbb{H}} \leqslant \beta \|f(\psi) - \psi\|^\nu_{\mathbb{H}}\,. \qquad (43)$$

Assuming furthermore (**U-LSI**)*, that $R + \theta\|\cdot\|^2_{\mathbb{H}}$ is convex for some $\theta > 0$ and that φ is ℓ-Lipschitz continuous, then, for all $\rho \in \mathcal{P}_2(\mathbb{R}^d)$ such that $\varphi(\rho) \in \mathcal{A}'$,*

$$\mathcal{F}(\rho) - \mathcal{F}(\rho_{\psi_*}) \leqslant \eta\sigma^2 \mathcal{I}\left(\rho|\Gamma(\rho)\right) + \theta\beta^2(4\eta^2\ell^2)^\nu \mathcal{I}^\nu\left(\rho|\Gamma(\rho)\right)\,. \qquad (44)$$

Under the settings of Proposition 11, assume that (43) holds in a neighborhood of ψ_*. Thanks to Proposition 11, there exists $\delta > 0$ such that for all initial conditions $\rho_0 \in \mathcal{B}_{\mathcal{W}_2}(\rho_{\psi_*}, \delta)$, the gradient flow converges to ρ_* and the inequality (44) holds for ρ_t for all time $t \geqslant 0$. Combined with Corollary 7, this yields algebraic convergence rates of the form

$$\mathcal{F}(\rho_t) - \mathcal{F}(\rho_{\psi_*}) \leqslant \frac{C}{\left(\mathcal{W}_2^{2r}(\rho_0, \rho_{\psi_*}) + t\right)^{1/r}}\,,$$

with $r = 1/\nu - 1 > 0$ and some constant $C > 0$, for all $t \geqslant 1$. From this algebraic decay of the free energy, we can get convergence of ρ_t to ρ_{ψ_*} since (20) in Lemma 3 gives

$$\mathcal{H}\left(\rho_t|\rho_{\psi_*}\right) \leqslant \mathcal{F}(\rho_t) - \mathcal{F}(\rho_{\psi_*}) + \theta\|\varphi(\rho_t) - \psi_*\|^2_{\mathbb{H}}\,.$$

When following the proof of Proposition 9 to get Proposition 12, we see that we obtain as an intermediary inequality that

$$\|\varphi(\rho_t) - \psi_*\|_{\mathbb{H}}^2 \leqslant \beta^2 (4\eta\ell^2)^\nu \mathcal{H}^\nu\left(\rho_t|\Gamma(\rho_t)\right)$$

as long as $\varphi(\rho_t) \in \mathcal{A}'$ (hence for all times here since $\rho_0 \in \mathcal{B}_{\mathcal{W}_2}(\rho_{\psi_*}, \delta)$). Moreover, since ψ_* is a local minimizer of g, by taking δ sufficiently small we get that $g(\varphi(\rho_t)) \geqslant g(\psi_*)$ for all $t \geqslant 0$, in other words

$$\mathcal{H}(\rho_t|\Gamma(\rho_t)) \leqslant \mathcal{F}(\rho_t) - \mathcal{F}(\rho_{\psi_*}) .$$

Gathering all these bounds give

$$\mathcal{H}\left(\rho_t|\rho_{\psi_*}\right) \leqslant \frac{C_0}{t^{\nu^2/(1-\nu)}} \tag{45}$$

for some $C_0 > 0$ for all $t \geqslant 0$.

3.2 Application to granular media equation

In this section we focus on the granular media equation (1) on $\mathbb{R}^d$, which corresponds to the free energy (10) (in fact we will consider a slightly more general case in Section 3.2.2). For clarity, for now, let us focus on the quadratic interaction case where $W(x, y) = \theta|x - y|^2$ for some $\theta > 0$ (recall from (17) that, for $\theta \leqslant 0$, $\mathcal{E}$ is convex and thus (**G-NL-LSI**) follows from (**U-LSI**) thanks to Lemma 3). This is a particular case of (33) with $\mathcal{H} = \mathbb{R}^d$,

$$V_0(x) = V(x) + \theta|x|^2 , \qquad \varphi(x) = x , \qquad R(\psi) = -\theta|\psi|^2 .$$

Assumption 6 is satisfied since $R + \theta|\cdot|^2$ is convex and φ is Lipschitz continuous (which, as discussed in Remark 4, also gives ($\mathcal{W}_2^2$**-curv-**$\mathcal{E}$) and (**Lip**)). We assume that V_0 is strongly convex outside a compact set, so that (**U-LSI**) holds (as explained in Section 2.1) and also (**o-s-Lip**). Assumptions 1, 2 and 3 are readily checked, and Assumption 4 follows from Propositions 2 and 6.

Using the notations introduced in Section 3.1 (except that we use m, as in *mean*, to denote the parameter, instead of ψ as in the general case), for $m \in \mathbb{R}^d$,

$$\rho_m(x) = \frac{e^{-\frac{1}{\sigma^2}\left[V(x)+\theta|x|^2-2\theta x\cdot m\right]}}{\int_{\mathbb{R}^d} e^{-\frac{1}{\sigma^2}\left[V(y)+\theta|y|^2-2\theta y\cdot m\right]}\mathrm{d}y} , \qquad f(m) = \int_{\mathbb{R}^d} x\rho_m(x)\mathrm{d}x ,$$

and

$$g(m) = \theta|m|^2 - \sigma^2 \ln \int_{\mathbb{R}^d} e^{-\frac{1}{\sigma^2}\left[V(x)+\theta|x|^2-2\theta x\cdot m\right]}\mathrm{d}x = -\sigma^2 \ln \int_{\mathbb{R}^d} e^{-\frac{1}{\sigma^2}\left[V(x)+\theta|x-m|^2\right]}\mathrm{d}x .$$

The graph of g and f are represented in Figures 1, 2 and 3, in dimension 1, either in the double-well case $V(x) = x^4/4 - x^2/2$ (Figures 1 and 3), or the convex case $V(x) = x^2/2$ (Figure 2), in each case with $W(x, y) = \theta|x - y|^2$ (with $\theta = 1$ in Figures 1 and 2 and $\theta = -1$ in Figure 3), at various temperatures.

3.2.1 The one-dimensional double-well case

In this section, $d = 1$, $\theta > 0$ and $V(x) = \frac{x^4}{4} - \frac{x^2}{2}$. In this one-dimensional attractive double-well case, the function f has been precisely studied in e.g. [46] (see also [25]), and the following holds (see Figure 1).

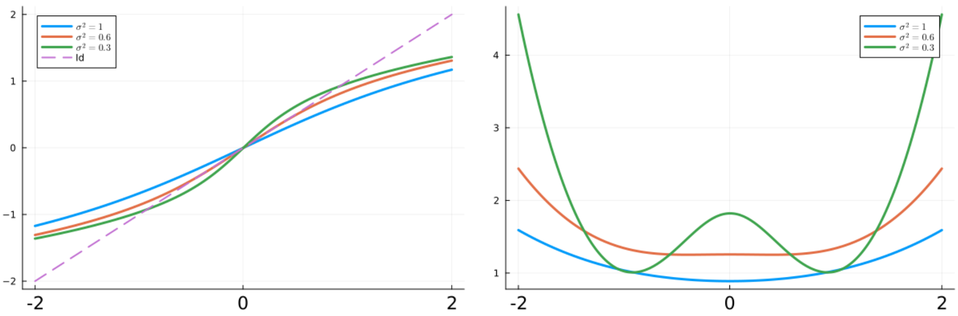

Figure 1: Graph of f (left) and g (right) in the double-well attractive case for $\sigma^2 \in \{1, 0.6, 0.3\}$

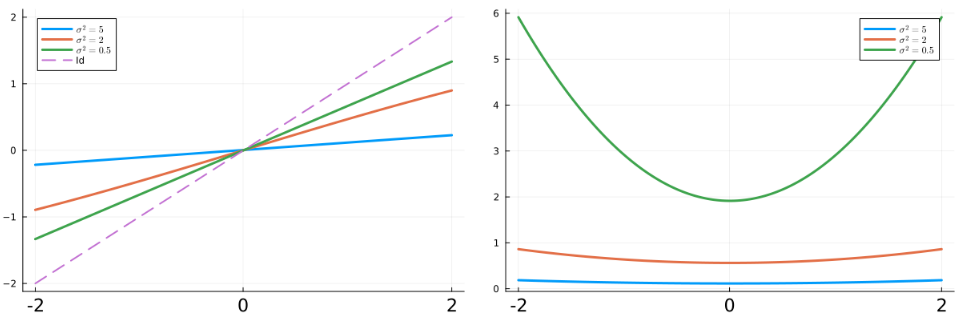

Figure 2: Graph of f (left) and g (right) in the single-well attractive case for $\sigma^2 \in \{5, 2, 0.5\}$

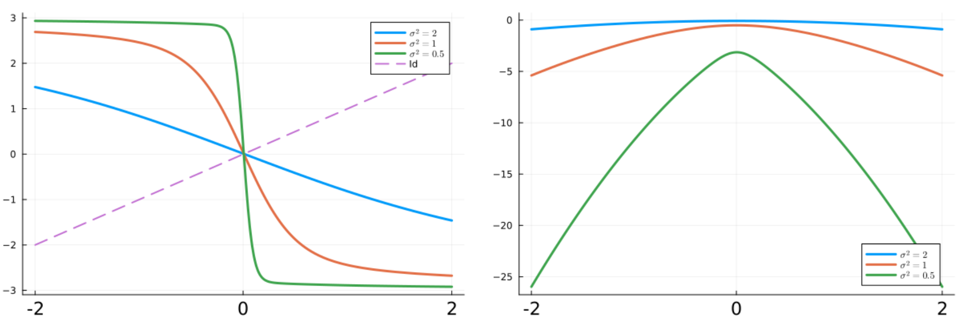

Figure 3: Graph of f (left) and g (right) in the double-well repulsive case for $\sigma^2 \in \{2, 1, 0.5\}$. At low temperature, $f'(0) < -1$, so that 0 is an unstable fixed point of $x \mapsto f(x)$.

- There exists $\sigma_c^2 > 0$ such that, for all $\sigma^2 \geqslant \sigma_c^2$, the unique fixed point of f is $m = 0$ and, for all $\sigma^2 < \sigma_c^2$, there are three such fixed points, $0, m_+, m_-$, with $m_+ > 0$ and $m_- = -m_+$.
- When $\sigma^2 < \sigma_c^2$, f is an increasing function with $0 < f'(m_+) = f'(m_-) < 1 < f'(0)$, $f'(m) \to 0$ as $|m| \to \infty$, $f(m) > m$ for $m \in (-\infty, m_-) \cup (0, m_+)$ and $f(m) < m$ for $m \in (m_-, 0) \cup (m_+, \infty)$.
- For $\sigma^2 = \sigma_c^2$, f is an increasing function with $f'(m) < 1$ for all $m \neq 0$, $f'(m) \to 0$ as $|m| \to \infty$, $f'(0) = 1$, $f''(0) = 0$ and $f^{(3)}(0) < 0$.
- For $\sigma^2 > \sigma_c^2$, f is an increasing function with $f'(m) \in (0,1)$ for all $m \in \mathbb{R}$ and $f'(m) \to 0$ as $|m| \to \infty$.

Moreover, since $g'(m) = 2\theta(m - f(m))$, the critical points of g are the fixed points of f and, for $\sigma^2 < \sigma_c^2$, g is decreasing on $(-\infty, m_-]$ and $[0, m_+]$ and increasing on $[m_-, 0]$ and $[m_+, \infty)$. As a consequence of all this, below the critical temperature, we get a non-linear log-Sobolev inequality for any non-centered $\rho \in \mathcal{P}_2(\mathbb{R}^d)$. We introduce the notation

$$m(\rho) = \int_{\mathbb{R}} x\rho(x)\mathrm{d}x.$$

Proposition 13. *In the double-well case with quadratic interaction* (2) *(with $\theta > 0$, $d = 1$) with $\sigma^2 < \sigma_c^2$, for any $\varepsilon > 0$, there exists $\alpha_\varepsilon \in [0,1)$ and $\overline{\eta}_\varepsilon > 0$ such that for any $m \geqslant \varepsilon$,*

$$|f(m) - m_+| \leqslant \alpha_\varepsilon |m - m_+| \tag{46}$$

and for any $\rho \in \mathcal{P}_2(\mathbb{R})$ with $|m(\rho)| \geqslant \varepsilon$,

$$\mathcal{F}(\rho) - \inf \mathcal{F} \leqslant \overline{\eta}_\varepsilon \sigma^4 \mathcal{I}\left(\rho|\Gamma(\rho)\right) . \tag{47}$$

In other words, for any $\varepsilon > 0$, (**NL-LSI**) *holds over $\mathcal{A}_\varepsilon = \{\rho \in \mathcal{P}_2(\mathbb{R}), |m(\rho)| \geqslant \varepsilon\}$.*

Remark 7. *In the present setting, Assumption 5 is satisfied and therefore Theorem 8 applies so that we get an exponential convergence starting from a $\mathcal{W}_2$-ball centered on ρ_{m_+} or ρ_{m_-}. This gives the first item of Proposition 1.*

Proof. Since $f(m) > m$ for $m \in (0, m_+)$ and $f(m) < m$ for $m \in (m_+, \infty)$, the function α given by

$$\alpha(m) = \frac{f(m) - m_+}{m - m_+}$$

for $m \neq m_+$ and $\alpha(m_+) = f'(m_+)$ is continuous on $\mathbb{R}_+$ with values in $(0,1)$ on $(0,\infty)$. Moreover, since $f'(m) \to 0$ as $m \to \infty$, so does α. As a consequence, for any $\varepsilon > 0$, $\alpha_\varepsilon := \sup_{m > \varepsilon} \alpha(m) < 1$, which proves (46). Proposition 9 then gives (47) for $\rho \in \mathcal{P}_2(\mathbb{R})$ with $m(\rho) \geqslant \varepsilon$. The case $m(\rho) \leqslant -\varepsilon$ is obtained by symmetry. ☐

Remark 8. *The same arguments work in the super-critical case where $\sigma^2 > \sigma_c^2$ because in that case $f'(0) < 1$. In that case the contraction $|f(m)| \leqslant \alpha |m|$ holds globally for $m \in \mathbb{R}$ for some $\alpha < 1$. In other words,* (**G-NL-LSI**) *holds for all $\sigma^2 > \sigma_c^2$.*

Remark 9. *To get explicit values for δ, λ, C in Proposition 1, it is sufficient to compute $f'(m_+)$ and to bound $\|f''\|_\infty$ (as in e.g. [25]) to get an explicit $\delta > 0$ such that $f'(m) \leqslant (1 + f'(m_+))/2 =: \alpha < 1$ for all m with $|m - m_+| \leqslant \delta$. From this, Theorem 8 gives explicit estimates in terms of $\sigma^2, \theta, \delta, \alpha$.*

Remark 10. *Denoting $\mathcal{A}_\varepsilon = \{\rho \in \mathcal{P}_2(\mathbb{R}), |m(\rho)| \geqslant \varepsilon\}$, the previous result does not mean that $\mathcal{F}(\rho_t) - \inf \mathcal{F} \leqslant e^{-t/\overline{\eta}_\varepsilon} (\mathcal{F}(\rho_t) - \inf \mathcal{F})$ for all $t \geqslant 0$ as soon as $\rho_0 \in \mathcal{A}_\varepsilon$, because the non-linear LSI (47) does not prevent $T_{\mathcal{A}_\varepsilon}(\rho_0) < \infty$. In fact, we provide a counter-example in the next statement.*

In the setting of Proposition 13, the precise determination of the basins of attraction of ρ_{m_+} and ρ_{m_-} for the McKean-Vlasov dynamics (8) remains an open question. The first item of Proposition 1 provides a sufficient condition, together with an exponential rate of convergence. On the other hand, in [45, 4], it is shown that if $\mathcal{F}(\rho_0)$ is sufficiently small, then the sign of $m(\rho_0)$ suffices to determine the limit of ρ_t. To complement these statements, we now show that without this condition on $\mathcal{F}(\rho_0)$, the sign of $m(\rho_t)$ may vary with t, so it is not enough to determine the limit of ρ_t. To our knowledge, this fact is known empirically but has never been explicitly evidenced.

Proposition 14. *Consider the granular media equation with $d = 1$, $V(x) = \frac{x^4}{4} - \frac{x^2}{2}$ and $W(x, y) = \theta(x - y)^2$ for some $\theta > 0$. Then, for any $\sigma^2 > 0$, there exist solutions of (1) with $m(\rho_0) > 0$ and $m(\rho_{t_0}) < 0$ for some $t_0 > 0$.*

Proof. For $\varepsilon \in [0, 1]$, we consider an initial condition $\rho_0 = (1 - \varepsilon)\delta_{-1} + \varepsilon\delta_{2/\varepsilon-1}$ and write $m_t = \int_{\mathbb{R}} x\rho_t(x)\mathrm{d}x$. In particular, $m_0 = 1$ independently from ε. We consider the time-inhomogeneous SDE

$$\mathrm{d}X_t = -V'(X_t)\mathrm{d}t - (X_t - m_t)\mathrm{d}t + \sqrt{2}\sigma\mathrm{d}B_t = \left(-X_t^3 + m_t\right)\mathrm{d}t + \sqrt{2}\sigma\mathrm{d}B_t\,. \tag{48}$$

Let ρ_t^+ and ρ_t^- be the law of the solution of (48) with respective initial conditions $\delta_{2/\varepsilon-1}$ and δ_{-1}. Then, $m_t = (1 - \varepsilon)\rho_t^- + \varepsilon\rho_t^+$ for all $t \geqslant 0$. Assume by contradiction that $m_t \geqslant 0$ for all $t \in [0, 1]$. Consider the solution of

$$\mathrm{d}Y_t = -Y_t^3\mathrm{d}t + \sqrt{2}\sigma\mathrm{d}B_t$$

(with the same Brownian motion as (48)) with $Y_0 = -1$. Taking X_0 distributed according to m_0 in (48) (in particular, $Y_0 \leqslant X_0$ and $X_t \sim m_t$ for all $t \geqslant 0$), we get by monotonicity that $Y_t \leqslant X_t$ for all $t \in [0, 1]$, and thus

$$m_t \geqslant \mathbb{E}(Y_t)\,, \qquad \int_{\mathbb{R}} x^3\rho_t(\mathrm{d}x) \geqslant \mathbb{E}(Y_t^3)\,.$$

Notice that the law of Y is independent from ε. Now,

$$\partial_t m_t = -\int_{\mathbb{R}} x^3\rho_t(\mathrm{d}x) + m_t \leqslant -\mathbb{E}(Y_t^3) + m_t\,,$$

which together with $m_0 = 1$ shows that $M := \sup_{\varepsilon\in[0,1]}\sup_{t\in[0,1]} m_t < \infty$. By comparing the solution of (48) initialized with $X_0 = -1$ (so that $X_t \sim \rho_t^-$) to the solution of

$$\mathrm{d}Z_t = (-Z_t^3 + M)\mathrm{d}t + \sqrt{2}\sigma\mathrm{d}B_t$$

with $Z_0 = -1$ (whose law is again independent from ε), we get that $X_t \leqslant Z_t$ for all $t \in [0, 1]$ and since $t \mapsto \mathbb{E}(Z_t)$ is continuous we get that there exists $t_0 \in (0, 1]$, independent from $\varepsilon \in [0, 1]$, such that

$$\int_{\mathbb{R}} x\rho_{t_0}^-(\mathrm{d}x) \leqslant -\frac{1}{2}\,.$$

Now, writing $h(t) = \int_{\mathbb{R}} x^2 \rho_t^+(\mathrm{d}x)$, we see that

$$h'(t) = -2 \int_{\mathbb{R}} x^4 \rho_t^+(\mathrm{d}x) + 2m_t^2 + 2\sigma^2 \leqslant -2h^2(t) + 2(M + \sigma^2)\,.$$

As a consequence, if h^2 goes below $2(M + \sigma^2)$ at some time, it stays below afterwards, and while h^2 is above $2(M + \sigma^2)$ it holds

$$h'(t) \leqslant -h^2(t)\,,$$

hence $h(t) \leqslant (t + 1/h(0))^{-1} \leqslant 1/t$. In any case, for all $t \in [0,1]$,

$$h(t) \leqslant \frac{1}{t} + 2(M + \sigma^2)\,.$$

By Cauchy-Schwarz,

$$\sup_{\varepsilon \in [0,1]} \int_{\mathbb{R}} x \rho_{t_0}^+(\mathrm{d}x) \leqslant \sup_{\varepsilon \in [0,1]} \sqrt{h(t_0)} \leqslant \sqrt{\frac{1}{t_0} + 2(M + \sigma^2)}\,,$$

and thus

$$m_{t_0} \leqslant -\frac{1-\varepsilon}{2} + \varepsilon \sqrt{\frac{1}{t_0} + 2(M + \sigma^2)}\,,$$

which is negative for ε small enough, leading to a contradiction with the assumption that $m_t \geqslant 0$ for all $t \in [0,1]$. □

To conclude the discussion in the one-dimensional double-well case, let us notice that the critical temperature $\sigma^2 = \sigma_c^2$ provides an example of the degenerate situation addressed in Section 3.1.3.

Proposition 15. *In the double-well case with quadratic interaction* (2) *(with $\theta > 0$, $d = 1$) with $\sigma^2 = \sigma_c^2$, there exist $\beta > 0$ and $\overline{\eta} > 0$ such that for any $m \in \mathbb{R}$,*

$$|m| \leqslant \beta \left(|m - f(m)| + |m - f(m)|^{1/3} \right)\,, \tag{49}$$

and for any $\rho \in \mathcal{P}_2(\mathbb{R})$

$$\mathcal{F}(\rho) - \inf \mathcal{F} \leqslant \overline{\eta} \left[\mathcal{I}\left(\rho | \Gamma(\rho)\right) + \left(\mathcal{I}\left(\rho | \Gamma(\rho)\right)\right)^{1/3} \right]\,. \tag{50}$$

Remark 11. *By contrast to* (43)*, which may only hold locally, in* (49) *we add a linear term, because in this very simple case this inequality in fact holds globally for all $m \in \mathbb{R}$ (and the linear term becomes dominant at long range). This does not change the conclusion on the free energy, in the sense that* (50) *is similar to* (44)*.*

Proof. As in the proof of Proposition 13, we see that, for any $\varepsilon > 0$, $\alpha(m) = f(m)/m$ is continuous on $\mathbb{R}_+ \setminus [-\varepsilon, \varepsilon]$ with values in $(0,1)$, from which $\alpha_\varepsilon := \sup_{|m| \geqslant \varepsilon} \alpha(m) < 1$. Then, for $m \notin [-\varepsilon, \varepsilon]$, $|m| \leqslant |m - f(m)| + |f(m)| \leqslant |m - f(m)| + \alpha_\varepsilon |m|$, which proves that (49) holds for such m with $\beta = 1/(1 - \alpha_\varepsilon)$.

Denoting $s = -f^{(3)}(0) > 0$, we get that $f(m) = m - sm^3 + o(m^3)$ as $m \to 0$. Let $\varepsilon > 0$ be such that $|f(m) - m + sm^3| \leqslant s|m|^3/2$ for all $m \in [-\varepsilon, \varepsilon]$. Then, for such m,

$$|m|^3 \leqslant 2 \frac{|f(m) - m|}{s}\,,$$

which concludes the proof of (49).

The proof of (50) is then exactly the same as the proof of (44) from (43). □

Remark 12. *The second part of Proposition 1 is a consequence of Proposition 15, following the arguments that led to* (45) *(the only difference being that it is not necessary to check that the solution starts and remains in some neighborhood of the stationary solution since* (50) *holds uniformly over $\mathcal{P}_2(\mathbb{R}^d)$).*

3.2.2 Localization in the low-temperature regime

Consider now the granular media equation (1) on $\mathbb{R}^d$ with $W(x,y) = |x-y|^2$ (i.e. we take $\theta = 1$ for simplicity, which can always be enforced up to rescaling V and σ^2). Let $a \in \mathbb{R}^d$ be a non-degenerate local minimizer of V such that a is a global minimizer of $x \mapsto h_0(x) = V(x)+|x-a|^2$. It is proven in [47] that, under additional technical conditions, there exists a family $(\rho_{*,\sigma})_{\sigma>0}$ such that $\mathcal{W}_2(\rho_{*,\sigma}, \delta_a) \to 0$ as $\sigma \to 0$ and, for all $\sigma > 0$, $\rho_{*,\sigma}$ is a stationary solution of (1) (at temperature σ). We want to apply Corollary 10 in this context to get that $\rho_{*,\sigma}$ is stable and get local convergence rates.

In fact we will work in a slightly more general setting. Consider the general case (33). For simplicity, we restrict the study to the finite-dimensional case where $\mathbb{H} = \mathbb{R}^p$ for some $p \geqslant 1$ and $R(\psi) = -|\psi|^2$. Moreover, we write $V(x) = V_0(x) + R(\varphi(x))$. In other words, in all this section,

$$\begin{aligned} \mathcal{E}(\mu) &= \int_{\mathbb{R}^d} V(x)\mu(\mathrm{d}x) + R(\varphi(\mu)) - \int_{\mathbb{R}^d} R(\varphi(x))\,\mu(\mathrm{d}x) \\ &= \int_{\mathbb{R}^d} V(x)\mu(\mathrm{d}x) + \frac{1}{2}\int_{\mathbb{R}^d} |\varphi(x)-\varphi(y)|^2\mu(\mathrm{d}x)\mu(\mathrm{d}y)\,. \end{aligned}$$

Denoting by g_σ the function defined in (38) to emphasize the dependency in the temperature, we see that, for any $\psi \in \mathbb{H}$,

$$\begin{aligned} g_\sigma(\psi) &= -|\psi|^2 + 2|\psi|^2 - \sigma^2 \ln \int_{\mathbb{R}^d} e^{-\frac{1}{\sigma^2}\left[V(x)+|\varphi(x)|^2 - 2\psi\cdot\varphi(x)\right]}\mathrm{d}x \\ &= -\sigma^2 \ln \int_{\mathbb{R}^d} e^{-\frac{1}{\sigma^2}\left[V(x)+|\varphi(x)-\psi|^2\right]}\mathrm{d}x\,. \end{aligned}$$

As $\sigma \to 0$,

$$g_\sigma(\psi) \to g_0(\psi) := \inf_{x\in\mathbb{R}^d}\{V(x) + |\varphi(x) - \psi|^2\}\,,$$

and this convergence holds uniformly for ψ in compact sets of $\mathbb{R}^p$. We are interested in critical points $\rho_{*,\sigma}$ which are localized at low temperature at some point $a \in \mathbb{R}^d$, in the sense that $\rho_{*,\sigma} \to \delta_a$ as σ vanishes. This implies that $\varphi(\rho_{*,\sigma}) \to \varphi(a)$. On the other hand, for a fixed ψ, if $x \mapsto V(x) + |\varphi(x) - \psi|^2$ has a unique global minimum x_0, then ρ_ψ converges as σ vanishes to a Dirac mass at x_0. Hence, we expect $\rho_{*,\sigma}$ to converge to a Dirac mass at the minimizer of $x \mapsto V(x)+|\varphi(x)-\varphi(a)|^2$, and thus we need this minimizer to be a. We retrieve the condition of [47].

Remark 13. *If we do not assume that $R(\psi) = -|\psi|^2$ but still write $V(x) = V_0(x) + R(\varphi(x))$, the previous computation gives*

$$g_\sigma(\psi) = -\sigma^2 \ln \int_{\mathbb{R}^d} e^{-\frac{1}{\sigma^2}h(x,\psi)}\mathrm{d}x\,,$$

with a modulated energy

$$h(x,\psi) = R(\psi) - R(\varphi(x)) + \nabla R(\psi)\cdot(\varphi(x)-\psi) + V(x) = M(\psi,\varphi(x)) + V(x)\,,$$

with the notation of Remark 5. As σ vanishes, $g_\sigma(\psi)$ converges toward $\inf_{x\in\mathbb{R}^d} h(x,\psi)$.

In view of the previous discussion, we work under the following setting:

Assumption 7. *In the case* (33), $\mathbb{H} = \mathbb{R}^p$, $R(\psi) = -|\psi|^2$, $V_0(x) = V(x) - R(\varphi(x))$, $V \in \mathcal{C}^3(\mathbb{R}^d,\mathbb{R})$, $\varphi \in \mathcal{C}^3(\mathbb{R}^d,\mathbb{R}^p)$, *the following conditions hold:*

- *$a \in \mathbb{R}^d$ is a non-degenerate local minimizer of V (i.e. $\nabla^2 V(a) > 0$) and it is the unique global minimizer of $x \mapsto V(x) + |\varphi(x) - \varphi(a)|^2$.*

- *V goes to infinity at infinity and $\int_{\mathbb{R}^d} |x|^2 e^{-\beta V(x)} \mathrm{d}x < \infty$ for some $\beta > 0$.*

Proposition 16. *Under Assumptions 6 and 7, assuming moreover that* (**o-s-Lip**) *holds for some $\kappa_1 > 0$, there exist $r_0, \sigma_0^2 > 0$ and a family $(\rho_{*,\sigma})_{\sigma \in (0,\sigma_0]}$ with $\mathcal{W}_2(\rho_{*,\sigma}, \delta_a) \to 0$ as $\sigma \to 0$ such that the following holds: for each $\sigma \in (0, \sigma_0]$, $\rho_{*,\sigma}$ is a stationary solution of the McKean-Vlasov equation (9) at temperature σ^2 and there exist $C, \overline{\eta} > 0$ such that for all initial condition $\rho_0 \in \mathcal{B}_{\mathcal{W}_2}(\delta_a, r_0)$, along the flow (9) at temperature σ^2, for all $t \geqslant 1$,*

$$\mathcal{W}_2^2(\rho_t, \rho_{*,\sigma}) + \sigma^2 \mathcal{H}(\rho_t | \rho_{*,\sigma}) + \mathcal{F}(\rho_t) - \mathcal{F}(\rho_{*,\sigma}) \leqslant C e^{-t/\overline{\eta}} \mathcal{W}_2^2(\rho_0, \rho_{*,\sigma}) . \tag{51}$$

Moreover, if we assume furthermore that $x \mapsto V(x) + |\varphi(x) - \psi|^2$ is strongly convex uniformly over ψ in a neighborhood of $\varphi(a)$, then the previous statement holds with $\overline{\eta}, C$ which are independent from $\sigma \in (0, \sigma_0]$.

Remark 14. *In particular, if V has several local minimizers $a_1, \ldots, a_n$ and the interaction is strong enough so that a_j is the unique global minimizer of $V + |\varphi - \varphi(a_j)|^2$ for several j, then, at a sufficiently low temperature, there are several stable stationary solutions to the granular media equation. By contrast, as discussed in Section 2.2, in the same setting, at sufficiently high temperature or sufficiently weak interaction (i.e. multiplying φ by a sufficiently small factor $\varepsilon > 0$), uniqueness holds. This shows that phase transitions occur as the temperature varies. See [22, 12] on this topic.*

Proof. First, to get the existence of $\rho_{*,\sigma}$, we reason in a first step as in Proposition 11, namely we identify that g_σ should have a local minimizer close to $\varphi(a)$ for σ small enough. Second, in the rest of the proof, we show that Corollary 10 applies. This gives a local non-linear LSI and thus a convergence rate thanks to Theorem 8 (since, as discussed in Remark 4, in the present case where R is quadratic, the conditions ($\mathcal{W}_2^2$**-curv-**$\mathcal{E}$) and (**Lip**) of Assumption 5 are implied by Assumption 6).

Step 1. Fix $r \in (0, 1]$ (which later on will be assumed sufficiently small, but always independently from σ). For $\psi \in \mathbb{R}^p$, set

$$\zeta(\psi, r) := \inf\{V(x) + |\varphi(x) - \psi|^2,\ |x - a| \geqslant r\} .$$

Let $M \geqslant 1$ be such that $V(x) \geqslant \sup\{V(y),\ y \in \mathcal{B}(a, 1)\} + (\ell + 1)^2$ for all $x \in \mathbb{R}^d$ with $|x| \geqslant M$. Then, for $\psi \in \mathcal{B}(\varphi(a), 1)$ and $x, y \in \mathbb{R}^d$ with $|x| \geqslant M$ and $|y - a| \leqslant 1$,

$$\begin{aligned}
V(y) + |\varphi(y) - \psi|^2 &\leqslant V(y) + (|\varphi(y) - \varphi(a)| + |\varphi(a) - \psi|)^2 \\
&\leqslant V(y) + (\ell |y - a| + |\varphi(a) - \psi|)^2 \\
&\leqslant V(y) + (\ell + 1)^2 \\
&\leqslant V(x) \\
&\leqslant V(x) + |\varphi(x) - \psi|^2 .
\end{aligned}$$

As a consequence, since $r \in (0, 1]$, for $\psi \in \mathcal{B}(\varphi(a), 1)$,

$$\begin{aligned}
\zeta(\psi, r) &= \inf\{V(x) + |\varphi(x) - \psi|^2,\ |x - a| \geqslant r,\ |x| \leqslant M\} \\
&\geqslant \zeta(\varphi(a), r) - 2|\psi - \varphi(a)| \sup_{|x| \leqslant M} |\varphi(x) - \varphi(a)| ,
\end{aligned}$$

where, for the second line, we have used that for $x \in \mathcal{B}(0, M)$,

$$\begin{aligned} V(x) + |\varphi(x) - \psi|^2 &= V(x) + |\varphi(x) - \varphi(a)|^2 - 2\left(\varphi(x) - \varphi(a)\right) \cdot (\psi - \varphi(a)) + |\psi - \varphi(a)|^2 \\ &\geqslant V(x) + |\varphi(x) - \varphi(a)|^2 - 2|\psi - \varphi(a)| \sup_{|x| \leqslant M} |\varphi(x) - \varphi(a)| \,. \end{aligned}$$

On the other hand, since a is the unique global minimizer of $x \mapsto V(x) + |\varphi(x) - \varphi(a)|^2$, $\zeta(\varphi(a), r) > V(a)$. Thus, we can find $\varepsilon_0 > 0$ (which depends on r) such that for all $\varepsilon \in (0, \varepsilon_0]$ and $\psi \in \mathcal{B}(\varphi(a), \varepsilon)$,

$$\zeta\left(\psi, r\right) > V(a) + \varepsilon^2$$

(notice that the condition becomes weaker as ε decreases). As a consequence, for all $\psi \in \mathcal{B}(\varphi(a), \varepsilon)$,

$$g_0(\psi) = \inf_{x \in \mathcal{B}(a,r)} \{V(x) + |\varphi\left(x\right) - \psi|^2\} \tag{52}$$

(since, for $x \notin \mathcal{B}(a, r)$, $V(x) + |\varphi(x) - \psi|^2 \geqslant \zeta(\psi, r) > V(a) + |\varphi(x) - \psi|^2$). Assuming that r is small enough so that a is the unique global minimizer of V over $\mathcal{B}(a, r)$, we get that $g_0(\psi) > V(a) = g_0\left(\varphi(a)\right)$ for all $\psi \in \mathcal{B}(\varphi(a), \varepsilon) \setminus \{\varphi(a)\}$, and in particular by continuity

$$\inf\{g_0(\psi),\ |\psi - \varphi(a)| = \varepsilon\} - V(a) > 0\,.$$

By uniform convergence on compact sets of g_σ toward g_0 as σ vanishes, there exists $\sigma_0 > 0$ (depending on ε) such that

$$\kappa := \inf\{g_\sigma(\psi),\ |\psi - \varphi(a)| = \varepsilon, \sigma \in (0, \sigma_0]\} - V(a) > 0\,,$$

while $g_\sigma\left(\varphi(a)\right) \to V(a)$ as $\sigma \to 0$, and in particular $g_\sigma\left(\varphi(a)\right) \leqslant V(a) + \kappa/2$ for σ small enough. This means that, then, the minimum of g_σ over $\mathcal{B}\left(\varphi(a), \varepsilon\right)$ is not attained at its boundary, which implies that g_σ admits a critical point within this ball. Since we can take ε arbitrarily small, we can take at any $\sigma \in (0, \sigma_0]$ such a critical point $\psi_{*,\sigma}$ in such a way that $\psi_{*,\sigma} \to \varphi(a)$ as $\sigma \to 0$. Thanks to (39), $\rho_{*,\sigma} := \rho_{\psi_{*,\sigma}}$ is a stationary solution to (9). This concludes the first step of the proof.

Before proceeding with Step 2 of the proof, let us discuss a few points. First,

$$\mathcal{W}_2^2(\rho_{*,\sigma}, \delta_a) = \frac{\int_{\mathbb{R}^d} |x - a|^2 e^{-\frac{1}{\sigma^2}\left[V(x) + |\varphi(x) - \psi_{*,\sigma}|^2\right]} \mathrm{d}x}{\int_{\mathbb{R}^d} e^{-\frac{1}{\sigma^2}\left[V(x) + |\varphi(x) - \psi_{*,\sigma}|^2\right]} \mathrm{d}x}\,.$$

Using that $\psi_{*,\sigma} \to \varphi(a)$ and that a is the global minimizer of $V + |\varphi - \varphi(a)|^2$, it is then not difficult to see that $\mathcal{W}_2^2(\rho_{*,\sigma}, \delta_a)$ vanishes with σ (see also the end of the proof for details of the argument in a more complicated case).

It remains to show that Corollary 10 applies. To emphasize the dependency on the temperature, we write f_σ the function (35). We have to prove that $|\nabla f_\sigma(\psi_{*,\sigma})| < 1$ for σ small enough. From (41) and (42),

$$|\nabla f_\sigma(\psi_{*,\sigma})| = \frac{2}{\sigma^2} \sup_{u \in \mathbb{S}^{p-1}} \mathrm{var}_{\rho_{*,\sigma}}\left(u \cdot \varphi(X)\right)\,.$$

Note that by the variational characterization of the variance, for $u \in \mathbb{S}^{p-1}$,

$$\mathrm{var}_{\rho_{*,\sigma}}\left(u \cdot \varphi(X)\right) \leqslant \int_{\mathbb{R}^d} |u \cdot \left(\varphi(x) - \varphi(z)\right)|^2 \rho_{*,\sigma}(\mathrm{d}x)$$

for any $z \in \mathbb{R}^d$. Having in mind the Laplace approximation of $\rho_{*,\sigma}$ as σ vanishes, it is natural to take z as the minimizer of $x \mapsto h_\sigma(x) := V(x) + |\varphi(x) - \psi_{*,\sigma}|^2$. In the rest of the proof, first, we justify that, for σ small enough, h_σ admits a unique minimizer z_σ, which converges to a as σ vanishes (this is Step 2) and then (in Step 3), we apply the Laplace approximation (which still works although h_σ and z_σ depends on σ) to get that,

$$\lim_{\sigma\to 0}\frac{2}{\sigma^2}\sup_{u\in\mathbb{S}^{p-1}}\int_{\mathbb{R}^d}|u\cdot(\varphi(x)-\varphi(z_\sigma))|^2\rho_{*,\sigma}(\mathrm{d}x)<1\,. \tag{53}$$

This implies that the same holds for σ small enough, which concludes the proof. The technical justification of the Laplace approximation is postponed to Step 4.

Step 2. We write $\nabla\varphi=(\partial_{x_i}\varphi_j)_{(i,j)\in[\![1,d]\!]\times[\![1,p]\!]}$, where i (resp. j) is the index of the line (resp. the column). Denoting $h_0(x)=V(x)+|\varphi(x)-\varphi(a)|^2$, we compute that

$$\begin{aligned}\nabla^2 h_\sigma(x) &= \nabla^2 V(x)+2\nabla\varphi(x)\,(\nabla\varphi(x))^T+\sum_{j=1}^p\nabla^2\varphi_j(x)\,(\varphi_j(x)-\psi_{j,*,\sigma})\\ &= \nabla^2 h_0(x)+\epsilon(x,\sigma)\end{aligned}$$

with

$$\epsilon(x,\sigma)=\sum_{j=1}^p\nabla^2\varphi_j(x)\,(\varphi_j(a)-\psi_{j,*,\sigma})\ ,$$

so that $\epsilon(x,\sigma)$ vanishes with σ uniformly over $x\in\mathcal{B}(a,r)$. Besides,

$$\nabla^2 h_0(a)=\nabla^2 V(a)+2\nabla\varphi(a)\,(\nabla\varphi(a))^T$$

is positive definite, and thus we can assume that r is small enough so that $\nabla^2 h_0(x)$ is uniformly bounded below by a positive constant over $x\in\mathcal{B}(a,r)$. Thanks to the uniform convergence of $\nabla^2 h_\sigma$ to $\nabla^2 h_0$ over $\mathcal{B}(a,r)$, there is $\sigma_0',\kappa>0$ such that for all $\sigma\in(0,\sigma_0']$ and $x\in\mathcal{B}(a,r)$, $\nabla^2 h_\sigma(x)\geqslant 2\kappa$. On the other hand, we have seen when establishing (52) that there exists $\kappa'>0$ such that for σ small enough,

$$\inf\{h_\sigma(x),\ |x-a|\geqslant r\}-\inf\{h_\sigma(x),\ |x-a|\leqslant r\}\geqslant\kappa' \tag{54}$$

(here we use that for any $\varepsilon>0$, $\psi_{*,\sigma}\in\mathcal{B}(\varphi(a),\varepsilon)$ for σ small enough). The infimum of h_σ is thus attained (for σ small enough) in $\mathcal{B}(a,r)$, where it is strongly convex, and thus it admits a unique global minimizer z_σ. Since r can be taken arbitrarily small (which changes the threshold σ_0'), we get that $z_\sigma\to a$ as σ vanishes. Moreover, for σ small enough, for all $x\in\mathcal{B}(a,r)$,

$$h_\sigma(x)\geqslant h_\sigma(z_\sigma)+\kappa|x-z_\sigma|^2\,. \tag{55}$$

Since h_σ admits a unique global maximum which is non-degenerate, the Laplace approximation heuristics suggest that expectations with respect to $\rho_{*,\sigma}\propto\exp(-\frac{1}{\sigma^2}h_\sigma)$ are equivalent to expectations with respect to the Gaussian measure $\hat\rho_\sigma\propto\exp(-\frac{1}{\sigma^2}\hat h_\sigma)$ with

$$\hat h_\sigma(x)=h_\sigma(z_\sigma)+\frac{1}{2}(x-z_\sigma)^T\nabla^2 h_\sigma(z_\sigma)(x-z_\sigma)\,.$$

As mentioned above, we postponed this technical justification to the end of the proof and, for now, take for granted that

$$\int_{\mathbb{R}^d}(\varphi-\varphi(z_\sigma))\,(\varphi-\varphi(z_\sigma))^T\,\rho_{*,\sigma}\underset{\sigma\to0}{\sim}\int_{\mathbb{R}^d}(\varphi-\varphi(z_\sigma))\,(\varphi-\varphi(z_\sigma))^T\,\hat\rho_\sigma\,. \tag{56}$$

Step 3. From this approximation, we have now to study the limit in (53). First, writing

$$\frac{1}{\sigma^2}\int_{\mathbb{R}^d}(\varphi(x)-\varphi(z_\sigma))(\varphi(x)-\varphi(z_\sigma))^T\hat\rho_\sigma(\mathrm{d}x)$$
$$=\mathbb{E}\left[\left(\frac{\varphi(z_\sigma+\sigma\sqrt{D_\sigma^{-1}}G)-\varphi(z_\sigma)}{\sigma}\right)\left(\frac{\varphi(z_\sigma+\sigma\sqrt{D_\sigma^{-1}}G)-\varphi(z_\sigma)}{\sigma}\right)^T\right],$$

with G a standard d-dimensional Gaussian variable and $D_\sigma := \nabla^2 h_\sigma(z_\sigma)$, the almost sure convergence

$$\lim_{\sigma\to 0}\frac{\varphi(z_\sigma+\sigma\sqrt{D_\sigma^{-1}}G)-\varphi(z_\sigma)}{\sigma}=(\nabla\varphi(a))^T\sqrt{D_0^{-1}}G,$$

together with the bound

$$\left|\frac{\varphi(z_\sigma+\sigma\sqrt{D_\sigma^{-1}}G)-\varphi(z_\sigma)}{\sigma}\right|^2\leq \ell^2\left|\sqrt{D_\sigma^{-1}}G\right|^2\underset{\sigma\to0}{\longrightarrow}\ell^2\left|\sqrt{D_0^{-1}}G\right|^2$$

give by dominated convergence that

$$\begin{aligned}
&\lim_{\sigma\to0}\mathbb{E}\left[\left(\frac{\varphi(z_\sigma+\sigma\sqrt{D_\sigma^{-1}}G)-\varphi(z_\sigma)}{\sigma}\right)\left(\frac{\varphi(z_\sigma+\sigma\sqrt{D_\sigma^{-1}}G)-\varphi(z_\sigma)}{\sigma}\right)^T\right]\\
&=\mathbb{E}\left[\left((\nabla\varphi(a))^T\sqrt{D_0^{-1}}G\right)\left((\nabla\varphi(a))^T\sqrt{D_0^{-1}}G\right)^T\right]\\
&=(\nabla\varphi(a))^T D_0^{-1}\nabla\varphi(a) \qquad (57)\\
&=A^T\left(2AA^T+\nabla^2V(a)\right)^{-1}A\,, \qquad (58)
\end{aligned}$$

where $A:=\nabla\varphi(a)$. Thanks to (56),

$$\frac{2}{\sigma^2}\int_{\mathbb{R}^d}\left(\varphi(x)-\varphi(z_\sigma)\right)\left(\varphi(x)-\varphi(z_\sigma)\right)^T\rho_{*,\sigma}(\mathrm{d}x)\underset{\sigma\to0}{\longrightarrow}2A^T\left(2AA^T+\nabla^2V(a)\right)^{-1}A\,.$$

It remains to show that this limit is strictly smaller than I_p (in the sense of quadratic forms), from which the left-hand side will be uniformly strictly less than I_p for σ small enough, which will conclude.

Let $u\in\mathbb{S}^{p-1}$ and $h=(2AA^T+\nabla^2V(a))^{-1}Au$. Then

$$h^TAu=h^T(2AA^T+\nabla^2V(a))h\geqslant(2+c)|A^Th|^2$$

for some $c>0$, using that $\nabla^2V(a)$ is definite positive. Then,

$$h^TAu\leqslant|u||A^Th|\leqslant\frac{1}{\sqrt{2+c}}|u|\sqrt{h^TAu}\,,$$

and thus

$$2u^TA^T\left(2AA^T+\nabla^2V(a)\right)^{-1}Au=2u^TA^Th\leqslant\frac{2}{2+c}\,.$$

As discussed above, this concludes the proof.

Notice that c is independent from σ, and thus, in Proposition 9, (36) holds on $\mathcal{A}'=\mathcal{B}(\varphi(a),r)$ with $r>0$ and $\alpha\in[0,1)$ independent from $\sigma\in(0,\sigma_0'']$ for some $\sigma_0''>0$. If, moroever, $x\mapsto V(x)+|\varphi(x)-\psi|^2$ is k-strongly convex for some $k>0$ for all $\psi\in\mathcal{A}'$, the

Bakry-Emery criterion shows that $\Gamma(\rho)$ satisfies an LSI with constant $\eta = k\sigma^2$ for all ρ such that $\varphi(\rho) \in \mathcal{A}'$, and thus for all $\rho \in \mathcal{A} := \mathcal{B}_{\mathcal{W}2}(\delta, r_0)$ provided r_0 is small enough (independently from σ). Following the proof of Proposition 9 (which in fact does not require (**U-LSI**) to hold globally on $\mathcal{P}(\mathbb{R}^d)$, but only on $\mathcal{A}$), we obtain for $\rho \in \mathcal{A}$ the local non-linear LSI

$$\mathcal{F}(\rho) - \mathcal{F}(\rho_{*,\sigma}) \leqslant k\sigma^4 \left(1 + \frac{4k\theta\ell^2}{(1-\alpha)^2}\right) \mathcal{I}\left(\rho|\Gamma(\rho)\right) ,$$

where θ, ℓ, α, k are independent from $\sigma \in (0, \sigma_0'']$, i.e. (**NL-LSI**) with $\overline{\eta}$ independent from σ. Similarly, using that $\eta = k\sigma^2$, we get that q_1 defined in (26) with $t = 1$ is bounded independently from σ small enough, for initial conditions $\rho_0 \in \mathcal{A}$. Applying Theorem 8 shows that (51) holds with $C, \overline{\eta}$ independent from $\sigma \in (0, \sigma_0'']$ in this case (again, the proof only uses an LSI for $\Gamma(\rho)$ uniformly over $\mathcal{A}$, not over $\mathcal{P}(\mathbb{R}^d)$).

Step 4. We now turn to the justification of (56). Write $\chi_0(x) = 1$ and, omitting the dependency in σ, $\chi_1(x) = (\varphi(x) - \varphi(z_\sigma))(\varphi(x) - \varphi(z_\sigma))^T$ (notice that $|\chi_1(x)| \leqslant C|x|^2 + C$ for some C independent from $\sigma \in (0, \sigma_0]$). For $i \in \{0,1\}$ and $k \in \{1,2,3,4\}$ we write

$$I_{i,k} = \int_{A_k} \chi_i \exp\left(-\frac{1}{\sigma^2}\left[h_\sigma - h_\sigma(z_\sigma)\right]\right) , \qquad \hat{I}_{i,k} = \int_{A_k} \chi_i \exp\left(-\frac{1}{\sigma^2}\left[\hat{h}_\sigma - h_\sigma(z_\sigma)\right]\right) ,$$

with

$$\begin{array}{ll} A_1 = \{|x - z_\sigma| \leqslant \sqrt{\sigma}\} & A_2 = \{\sqrt{\sigma} < |x - z_\sigma| \leqslant 2r\} \\ A_3 = \{2r < |x - z_\sigma| \leqslant M\} & A_4 = \{M < |x - z_\sigma|\} , \end{array}$$

where r is small enough for (54) to hold for some κ' and for (55) to hold for all $x \in \mathcal{B}(a, 3r)$ for some κ (with $r, \kappa, \kappa' > 0$ independent from σ) for all σ small enough, and $M > 1$ is large enough so that

$$\forall x \notin \mathcal{B}(a, M-1) , \qquad \frac{1}{2}V(x) \geqslant V(a) + 2 \tag{59}$$

For σ small enough, $|z_\sigma - a| < r$, and thus $\mathcal{B}(a, r) \subset \mathcal{B}(z_\sigma, 2r)$, which together with (54) gives that

$$\inf\{h_\sigma(x),\ |x - z_\sigma| \geqslant 2r\} \geqslant h_\sigma(z_\sigma) + \kappa' .$$

As a consequence, for $i \in \{0,1\}$, using that φ is bounded uniformly over $\mathcal{B}(a, M+1)$ (hence over $\mathcal{B}(z_\sigma, M)$ for σ small enough)

$$I_{i,3} = \underset{\sigma \to 0}{\mathcal{O}}\left(e^{-\frac{\kappa'}{\sigma^2}}\right) .$$

Similarly, for σ small enough, $\mathcal{B}(z_\sigma, 2r) \subset \mathcal{B}(a, 3r)$ and thus, thanks to (55) for $x \in \mathcal{B}(a, 2r)$, using that φ is bounded uniformly over $\mathcal{B}(a, 3r)$, for $i \in \{0,1\}$,

$$I_{i,2} = \underset{\sigma \to 0}{\mathcal{O}}\left(e^{-\frac{\kappa}{\sigma}}\right) .$$

Next, using that $\mathcal{B}(a, M-1) \subset \mathcal{B}(z_\sigma, M)$ and $V(a) + 1 \geqslant h_\sigma(z_\sigma)$ for σ small enough, (59) implies that

$$\forall x \notin \mathcal{B}(z_\sigma, M) , \qquad h_\sigma(x) \geqslant V(x) \geqslant \frac{1}{2}V(x) + 1 + h_\sigma(z_\sigma) .$$

This gives, for $\sigma \leqslant 1/(2\beta)$,

$$I_{i,4} \leqslant e^{-\frac{1}{\sigma^2}} \int_{|x - z_\sigma| > M} |\chi_i(x)| e^{-\frac{V(x)}{2\sigma^2}} \mathrm{d}x \leqslant e^{-\frac{1}{\sigma^2}} \int_{\mathbb{R}^d} |\chi_i(x)| e^{-\beta V(x)} \mathrm{d}x = \underset{\sigma \to 0}{\mathcal{O}}\left(e^{-\frac{1}{\sigma^2}}\right) ,$$

since $e^{-\beta V}$ admits a second moment (here for simplicity we have assumed without loss of generality that $V \geqslant 0$).

Finally, using that $\nabla^{(3)} h_\sigma$ converges $\nabla^{(3)} h_0$ uniformly over compact sets, we get that there exists $C > 0$ such that

$$|h_\sigma(x) - \hat{h}_\sigma(x)| \leqslant C|x - z_\sigma|^3$$

for all $x \in \mathcal{B}(a, 1)$ (hence all $x \in \mathcal{B}(z_\sigma, \sqrt{\sigma})$) for σ small enough. From this, for $i \in \{0, 1\}$,

$$I_{i,1} = \hat{I}_{i,1}\left(1 + \underset{\sigma \to 0}{\mathcal{O}}(\sigma)\right) .$$

On the other hand, denoting $c_\sigma = (2\pi)^{d/2}\sqrt{\det[(\nabla^2 h_\sigma(z_\sigma))^{-1}]}$ (which converges to some $c_0 > 0$ as $\sigma \to 0$), by usual computations for Gaussian distributions,

$$I_0 := \sum_{k=1}^{4} I_{0,k} = \hat{I}_{0,1}\left(1 + \underset{\sigma \to 0}{\mathcal{O}}(\sigma)\right) + \underset{\sigma \to 0}{\mathcal{O}}\left(e^{-\frac{\kappa}{\sigma}}\right) = \sigma^d c_\sigma + \underset{\sigma \to 0}{\mathcal{O}}\left(\sigma^{d+1}\right) ,$$

and then

$$\begin{aligned}
\frac{\sum_{k=1}^4 I_{1,k}}{I_0} &= \left(\sigma^d c_\sigma + \underset{\sigma \to 0}{\mathcal{O}}\left(\sigma^{d+1}\right)\right)^{-1}\left(\hat{I}_{1,1}\left(1 + \underset{\sigma \to 0}{\mathcal{O}}(\sigma)\right) + \underset{\sigma \to 0}{\mathcal{O}}\left(e^{-\frac{\kappa}{\sigma}}\right)\right) \\
&= \frac{\hat{I}_{1,1}}{\sigma^d c_\sigma}\left(1 + \underset{\sigma \to 0}{\mathcal{O}}(\sigma)\right) + \underset{\sigma \to 0}{\mathcal{O}}\left(e^{-\frac{\kappa}{2\sigma}}\right) \\
&= \frac{\sum_{k=1}^4 \hat{I}_{1,k}}{\sigma^d c_\sigma}\left(1 + \underset{\sigma \to 0}{\mathcal{O}}(\sigma)\right) + \underset{\sigma \to 0}{\mathcal{O}}\left(e^{-\frac{\kappa}{2\sigma}}\right) \\
&= \sigma^2 \left(\nabla\varphi(a)\right)^T \left(\nabla^2 h_0(a)\right)^{-1} \nabla\varphi(a) + \underset{\sigma \to 0}{o}\left(\sigma^2\right) ,
\end{aligned}$$

as we computed in (58). This concludes the proof of (56), hence of Proposition 16. $\square$

Remark 15. *In the case* (1) *with* $W(x, y) = \theta|x - y|^2$*, taking* $\theta > -\inf \nabla^2 V/2$ *(as in [49, Theorem 2.3]), we get that* (51) *holds with* $\overline{\eta}, C$ *that are uniform over* σ *small enough in a neighborhood of* δ_a*. At first this seems to be a desirable property for optimization (and in fact this is the idea underlying consensus-based optimization). This should be mitigated by the following observations.*

First, this condition implies that $x \mapsto b_a(x) = V(x) + \theta|x - a|^2$ *is strongly convex for all* a *and thus, any local minimizer* a *of* V *being a critical point of* b_a*, it is then the unique global minimizer of* b_a*. Hence, the Wasserstein gradient descent can be trapped in any local well of* V *(and thus it is not very different from a basic gradient descent for* V *in* $\mathbb{R}^d$*). By contrast, if* θ *is taken smaller, the condition that* a *has to be a global minimizer of* b_a *for a localization to occur can be used as a way to discard shallow local wells and select better solutions.*

Besides, the decay (51) *holds for solutions initialized in* $\mathcal{B}_{\mathcal{W}_2}(\delta_a, r_0)$ *for some small* r_0*. Now, assume for instance that we start close to* $\delta_{a'}$ *where* a' *is a local minimizer of* V *but not a global minimizer of* $b_{a'}$*, and* $a \neq a'$ *is the global minimizer of* V *and also the global minimizer of* b_x *for all* $x \in \mathbb{R}^d$*. For* σ *small enough, we expect all solutions to converge to a stationary solution close to* δ_a*. In particular, eventually, any trajectory will lie in* $\mathcal{B}_{\mathcal{W}_2}(\delta_a, r_0)$ *and thus the convergence rate* $\overline{\eta}$ *in* (51) *is independent from* σ*, as in Remark 1. However the constant* C *highly depends on the initial condition, reflecting the time needed to reach* $\mathcal{B}_{\mathcal{W}_2}(\delta_a, r_0)$*. Starting close to* $\delta_{a'}$*, the additional convexity due to the interaction will increase the energy barrier the process has to overcome to move from* a' *to* a*. More quantitatively, the Arrhenius law indicates that the typical time for the solution to put most of its mass around* a *will be of order* e^{D/σ^2}

where D is larger than in the case $\theta = 0$. In other words, an attractive interaction worsen the metastability of the process. If the goal is not to improve the local convergence rate but to enhance the exploration of the space by lowering energy barriers (as in the Adaptive Potential algorithms and related methods [6]), repulsive interaction (i.e. $\theta < 0$) seems more indicated, as studied in [13].

4 Vlasov-Fokker-Planck equation

In this section, we consider the kinetic Vlasov-Fokker-Planck equation, which reads

$$\partial_t \rho_t + v \cdot \nabla_x \rho_t = \nabla_v \cdot \left(\sigma^2 \nabla_v \rho_t + \left(v + \nabla E_{\rho_t^x}\right) \rho_t\right) , \tag{60}$$

where $\rho_t(x, v)$ is the probability density of particles at position $x \in \mathbb{R}^d$ with velocity $v \in \mathbb{R}^d$, $\rho^x(x) = \int_{\mathbb{R}^d} \rho(x, v) \mathrm{d}v$ is the marginal density of the position, E_{ρ^x} is the linear functional derivative of an energy $\mathcal{E}$ as in Section 2.1 (it only depends on the position x and the marginal density ρ^x).

Entropic long-time convergence for (60) have been established in some cases (with the energy corresponding to the granular media case (10)) in [40, 27] (using uniform LSI for the associated particle system), [14] (with global non-linear LSI) or [43] (with uniform conditional LSI for the associated particle system and a smallness assumption on the interaction). All these works establish global convergence toward a unique stationary solution and thus do not cover cases with several stationary solutions.

The Vlasov-Fokker-Planck is not the gradient flow of some free energy functional. However, we can still get a local convergence under the same assumptions as in the elliptic case (1). This is the main result of this section, stated in Theorem 20 below.

The free energy associated to (60) is

$$\mathcal{F}_k(\rho) = \sigma^2 \mathcal{H}(\rho) + \mathcal{E}(\rho^x) + \mathcal{E}_k(\rho)$$

with the kinetic energy

$$\mathcal{E}_k(\rho) = \int_{\mathbb{R}^{2d}} \frac{|v|^2}{2} \rho(x, v) \mathrm{d}x \mathrm{d}v \,.$$

In this section, Assumptions 1, 2 and 3 remain in force (which implies that $\mathcal{F}_k$ is also lower bounded, see Remark 16 below). The corresponding local equilibria are given by

$$\Gamma_k(\rho) = \Gamma(\rho^x) \otimes \mathcal{N}(0, \sigma^2 I_d) \tag{61}$$

where $\Gamma(\rho^x)$ is the corresponding overdamped local equilibrium (13) and $\mathcal{N}(0, \sigma^2 I_d)$ stands for the centered d-dimensional Gaussian distribution with variance $\sigma^2 I_d$. In other words,

$$\Gamma_k(\rho) \propto \exp\left(-\frac{1}{\sigma^2}\left(E_{\rho^x}(x) + \frac{1}{2}|v|^2\right)\right) \mathrm{d}x \mathrm{d}v \,.$$

Along the flow (60), under suitable regularity conditions,

$$\partial_t \mathcal{F}_k(\rho_t) = -\sigma^4 \int_{\mathbb{R}^{2d}} \left|\nabla_v \ln \frac{\rho_t}{\Gamma_k(\rho_t)}\right|^2 \mathrm{d}\rho_t \,. \tag{62}$$

Hence, $t \mapsto \mathcal{F}_k(\rho_t)$ is non-increasing, but the free energy dissipation can vanish even if $\rho \neq \Gamma_k(\rho)$. To avoid ambiguity we write $\mathcal{K}_k = \{\rho_* \in \mathcal{P}_2(\mathbb{R}^{2d}),\ \mathcal{F}_k(\rho_*) < \infty,\ \Gamma_k(\rho_*) = \rho_*\}$. There is a one-to-one correspondence between $\mathcal{K}_k$ (the set of critical points of $\mathcal{F}_k$ in $\mathcal{P}_2(\mathbb{R}^{2d})$) and $\mathcal{K}$

(the set of critical points of $\mathcal{F} = \sigma^2\mathcal{H} + \mathcal{E}$ in $\mathcal{P}_2(\mathbb{R}^d)$), since $\rho_* \in \mathcal{K}_k$ if and only if $\rho_*^x \in \mathcal{K}$ and $\rho_* = \rho_*^x \otimes \mathcal{N}(0, \sigma^2 I_d)$.

As in the elliptic case, we do not address well-posedness issues and work under the following conditions.

Assumption 8. *For all $\rho_0 \in \mathcal{P}_2(\mathbb{R}^{2d})$, (60) has a unique strong solution, continuous in time for $\mathcal{W}_2$, such that, for all $t > 0$, ρ_t has a continuous positive density, $\mathcal{F}_k(\rho_t)$, $\mathcal{H}(\rho_t|\Gamma_k(\rho_t))$ and $\mathcal{I}(\rho_t|\Gamma_k(\rho_t))$ are finite and (62) holds, with $t \mapsto \mathcal{I}(\rho_t|\Gamma(\rho_t))$ continuous over $\mathbb{R}_+^*$.*

To recover an entropy dissipation with a full Fisher information instead of (62), we work with a modified free energy of the form

$$\mathcal{L}(\rho_t) = \mathcal{F}_k(\rho_t) + a\sigma^4 \int_{\mathbb{R}^{2d}} \left|(\nabla_x + \nabla_v) \ln \frac{\rho_t}{\Gamma_k(\rho_t)}\right|^2 \mathrm{d}\rho_t\,, \tag{63}$$

for some $a > 0$ to be fixed.

Lemma 17. *Assume that*

$$M := \sup_{\rho \in \mathcal{P}_2(\mathbb{R}^d)} \|\nabla^2 E_\rho\|_\infty < \infty\,. \tag{$x-\mathbf{Lip}$}$$

Then, taking $a = (3 + 4M^2)^{-1}$ in the definition of $\mathcal{L}$, for all $t \geqslant 0$,

$$\partial_t \mathcal{L}(\rho_t) \leqslant -\frac{\sigma^4}{4} \mathcal{I}\left(\rho_t|\Gamma_k(\rho_t)\right)\,.$$

Proof. Set $h_t = \rho_t/\Gamma_k(\rho_t)$. As computed in the proof of [39, Proposition 3] (notice that the friction parameter γ of [39] is 1 in the present case, which besides can always be enforced by a suitable time rescaling, up to rescaling E_ρ and σ^2), for any $2d \times 2d$ matrix A of the form

$$A = \begin{pmatrix} a_{11} I_d & a_{12} I_d \\ a_{21} I_d & a_{22} I_d \end{pmatrix}\,,$$

writing $M' = M\sqrt{a_{11}^2 + a_{21}^2}$, it holds, for all $r > 0$,

$$\partial_t \int_{\mathbb{R}^{2d}} |A \nabla \ln h_t|^2 \, \mathrm{d}\rho_t \leqslant 2 \int_{\mathbb{R}^{2d}} \nabla \ln h_t \cdot B \nabla \ln h_t \mathrm{d}\rho_t$$

with

$$B = A^T A \begin{pmatrix} M' r I_d & 0 \\ -I_d & (M' r - 1) I_d \end{pmatrix} + \frac{M'}{r} \begin{pmatrix} 0 & 0 \\ 0 & I_d \end{pmatrix}\,.$$

Taking $a_{ij} = 1/\sqrt{2}$ for all $i, j \in \{1, 2\}$ we get that $M' = M$ and

$$\begin{aligned} B &= \begin{pmatrix} (Mr - 1) I_d & (Mr - 1) I_d \\ (Mr - 1) I_d & \left(Mr - 1 + \frac{M}{r}\right) I_d \end{pmatrix} \\ &\preceq \begin{pmatrix} -\frac{1}{4} I_d & 0 \\ 0 & \left(\frac{1}{2} + 2M^2 + 1\right) I_d \end{pmatrix} \end{aligned}$$

taking $r = 1/(2M)$, where $\preceq$ stands for the order over symmetric matrices and we used that $xv \leqslant \frac{1}{4}x^2 + v^2$ to bound the cross term. In other words, we have obtained that

$$\partial_t \int_{\mathbb{R}^d} |(\nabla_x + \nabla_v) \ln h_t|^2 \, \mathrm{d}\rho_t \leqslant -\frac{1}{4} \int_{\mathbb{R}^d} |\nabla_x \ln h_t|^2 \, \mathrm{d}\rho_t + \left(\frac{3}{2} + 2M^2\right) \int_{\mathbb{R}^d} |\nabla_v \ln h_t|^2 \, \mathrm{d}\rho_t\,.$$

Conclusion follows from (62). $\square$

As a corollary, assuming in the kinetic case the local non-linear LSI

$$\forall \rho \in \mathcal{A}\,, \qquad \mathcal{F}_k(\rho) - \mathcal{F}_k(\rho_*) \leqslant \overline{\eta}\sigma^4 \mathcal{I}(\rho|\Gamma_k(\rho)) \tag{k-NL-LSI}$$

for some $\mathcal{A} \subset \mathcal{P}_2(\mathbb{R}^{2d})$, $\overline{\eta} > 0$ and $\rho_* \in \mathcal{K}_k$, which, using that $|(\nabla_x + \nabla_v)f|^2 \leqslant 2|\nabla f|^2$, in turn implies

$$\forall \rho \in \mathcal{A}\,, \qquad \mathcal{L}(\rho) - \mathcal{L}(\rho_*) \leqslant (\overline{\eta} + 2a)\,\sigma^4 \mathcal{I}(\rho|\Gamma_k(\rho))\,,$$

we get an exponential decay of $\mathcal{L}(\rho_t) - \mathcal{L}(\rho_*)$ for times $t \leqslant T_{\mathcal{A}}(\rho_0)$.

Remark 16. *For $\rho \in \mathcal{P}_2(\mathbb{R}^{2d})$,*

$$\begin{aligned}\mathcal{F}_k(\rho) &= \sigma^2 \mathcal{H}\left(\rho|\Gamma_k(\rho)\right) - \sigma^2 \ln\left(Z_{\rho^x}(2\pi\sigma^2)^{d/2}\right) + \mathcal{E}(\rho^x) - \int_{\mathbb{R}^d} \rho^x E_{\rho^x} \\ &= \sigma^2 \mathcal{H}\left(\rho|\Gamma_k(\rho)\right) - \frac{d}{2}\sigma^2 \ln\left(2\pi\sigma^2\right) + \mathcal{G}\left(\rho^x\right)\end{aligned}$$

with

$$\mathcal{G}(\rho^x) = \mathcal{E}(\rho^x) - \int_{\mathbb{R}^d} \rho^x E_{\rho^x} - \sigma^2 \ln Z_{\rho^x}\,,$$

which is also such that, in the elliptic case, $\mathcal{F}(\rho^x) = \sigma^2\mathcal{H}(\rho^x|\Gamma(\rho^x)) + \mathcal{G}(\rho^x)$. In particular, by the extensivity property of the relative entropy, $\mathcal{F}_k(\rho) \geqslant \mathcal{F}(\rho^x) - d\sigma^2/2\ln(2\pi\sigma^2)$.

If $\Gamma(\rho^x)$ satisfies a LSI uniformly over ρ, so does $\Gamma_k(\rho)$ by tensorization of LSI, and then both (**k-NL-LSI**) *and* (**NL-LSI**) *follow from a bound of the form $\mathcal{G}(\rho^x) - \mathcal{G}(\rho^x_*) \leqslant \overline{\eta}'\mathcal{I}(\rho^x|\Gamma(\rho^x))$ for some $\overline{\eta}' > 0$, as has been established in various situations in Section 3.*

In order to follow the argument that led in the elliptic case to Theorem 8, we need to provide an analogue of Corollary 7 in the present hypoelliptic case. This is done by using the Wasserstein-to-entropy short time regularization of [42, Theorem 4.1] instead of Proposition 6 and the entropy-to-Fisher short time regularization of [14, Proposition 5.5], as we now detail.

The following is from [14, Proposition 5.5] (the proof is based on computations similar to Lemma 17).

Proposition 18. *Assuming* (**Lip**) *and* ($x-$**Lip**)*, there exists $\varepsilon \in (0,1]$ (which depends only on M, L and σ^2) such that for any solution of* (60) *with $\mathcal{F}_k(\rho_0) < \infty$, writing*

$$\mathcal{E}(t,\rho) := \mathcal{F}_k(\rho) + \varepsilon t \int_{\mathbb{R}^{2d}} \left[\left|\nabla_v \ln \frac{\rho}{\Gamma_k(\rho)}\right|^2 + \left|(\varepsilon t \nabla_x + \nabla_v) \ln \frac{\rho}{\Gamma_k(\rho)}\right|^2 \right] \mathrm{d}\rho\,,$$

then $t \mapsto \mathcal{E}(t,\rho_t)$ is non-increasing over $t \in [0,1]$.

The following is a particular case of [42, Theorem 4.1].

Proposition 19. *Assuming* (**Lip**) *and* ($x-$**Lip**)*, there exists $K > 0$ (which depends only on M, L and σ^2) such that for any solution of* (60) *with $\rho_0 \in \mathcal{P}_2(\mathbb{R}^{2d})$ and any $\rho_* \in \mathcal{K}_k$ and $t \geqslant 0$,*

$$\mathcal{H}\left(\rho_{t+1}|\rho_*\right) \leqslant K\mathcal{W}_2^2(\rho_t,\rho_*)\,.$$

We can now state the analogue to Theorem 8.

Assumption 9. *The conditions* ($\mathcal{W}_2^2$**-curv-**$\mathcal{E}$)*,* (**Lip**)*,* ($x-$**Lip**) *and* (**U-LSI**) *hold for some $\lambda, L, M, \eta > 0$, and* (**k-NL-LSI**) *holds on some non-empty set $\mathcal{A} \subset \mathcal{P}_2(\mathbb{R}^{2d})$ for some $\overline{\eta} > 0$. Furthermore there exist $\rho_* \in \mathcal{A}\cap\mathcal{K}_k$ and $\delta > 0$ such that $\mathcal{B}_{\mathcal{W}_2}(\rho_*,\delta) \subset \mathcal{A}$ and $\mathcal{B}_{\mathcal{W}_2}(\rho_*,\delta)\cap\mathcal{K}_k = \{\rho_*\}$. The function $\mathcal{L}$ is given by* (63) *with $a = (3+4M^2)^{-1}$.*

Theorem 20. *Under Assumption 9, there exist $\delta', c, C > 0$ (which depend only on λ, M, L, δ, η, $\overline{\eta}$ and σ^2) such that, for all $\rho_0 \in \mathcal{B}_{\mathcal{W}_2}(\rho_*, \delta')$, $T_{\mathcal{A}}(\rho_0) = \infty$; for all $t \geqslant 0$,*

$$\mathcal{W}_2^2(\rho_t, \rho_*) \leqslant Ce^{-ct}\mathcal{W}_2^2(\rho_0, \rho_*) \tag{64}$$
$$\mathcal{F}_k(\rho_t) - \mathcal{F}_k(\rho_*) \leqslant Ce^{-ct}\left(\mathcal{F}_k(\rho_0) - \mathcal{F}_k(\rho_*)\right) \tag{65}$$
$$\mathcal{H}(\rho_t|\rho_\infty) \leqslant Ce^{-ct}\mathcal{H}(\rho_0|\rho_\infty) \tag{66}$$

and for all $t \geqslant 1$,

$$\begin{aligned}&\max\left(\mathcal{W}_2^2(\rho_0, \rho_*)\,,\, \mathcal{H}(\rho_0|\rho_\infty)\,,\, \mathcal{F}_k(\rho_0) - \mathcal{F}_k(\rho_*)\right)\\ &\qquad \leqslant Ce^{-ct}\min\left(\mathcal{W}_2^2(\rho_0, \rho_*)\,,\, \mathcal{H}(\rho_0|\rho_\infty)\,,\, \mathcal{F}_k(\rho_0) - \mathcal{F}_k(\rho_*)\right)\,.\end{aligned} \tag{67}$$

Proof. First, without loss of generality we can assume that (**U-LSI**) holds with $\eta \geqslant \sigma^2$, in which case the same inequality holds for all $\nu, \rho \in \mathcal{P}_2(\mathbb{R}^{2d})$ when Γ is replaced by Γ_k, due to the LSI satisfied by $\mathcal{N}(0, \sigma^2 I_d)$ and the tensorization property of LSI. In particular, ρ_* satisfies a LSI and thus a Talagrand inequality with constant η. Moreover, notice that ($\mathcal{W}_2^2$**-curv-**$\mathcal{E}$) implies the same condition with $\mathcal{E}$ replaced by $\mathcal{E} + \mathcal{E}_k$, since $\mu \mapsto \mathcal{E}_k(\mu)$ is linear. In particular, in view of these two points, we can apply Lemmas 3 and 5 in the kinetic case (namely with $\mathcal{F}$ replaced by $\mathcal{F}_k$, Γ by Γ_k, $\rho_* \in \mathcal{K}_k$ and $\mu_0, \mu_1, \rho \in \mathcal{P}_2(\mathbb{R}^{2d})$). Using all these points, for any $\rho \in \mathcal{P}_2(\mathbb{R}^{2d})$,

$$\begin{aligned}\mathcal{F}_k(\rho) - \mathcal{F}_k(\rho_*) &\leqslant \sigma^2\mathcal{H}\left(\rho|\Gamma_k(\rho)\right) + \lambda\mathcal{W}_2^2(\rho, \rho_*)\\ &\leqslant \sigma^2\left(1 + L\eta + \frac{4\lambda\eta}{\sigma^2} + \frac{L^2\eta^2}{2}\right)\mathcal{H}(\rho|\rho_*)\,.\end{aligned} \tag{68}$$

Similarly, we can apply Theorem 8 to the elliptic equation on $\mathbb{R}^{2d}$

$$\partial_t\tilde{\rho}_t = \nabla\cdot\left(\tilde{\rho}_t\nabla\frac{\delta\mathcal{F}_k(\tilde{\rho}_t)}{\delta\mu}\right)\,,$$

from which (27) reads

$$\mathcal{W}_2^2(\rho, \rho_*) \leqslant 4\overline{\eta}\left(\mathcal{F}_k(\rho) - \mathcal{F}_k(\rho_*)\right) \tag{69}$$

for all $\rho \in \mathcal{B}_{\mathcal{W}_2}(\rho_*, \delta_1')$ for $\delta_1' \leqslant \delta$ small enough. Using that $t \mapsto \mathcal{F}_k(\rho_t)$ is non-increasing and Proposition 19, gathering the previous bounds gives that

$$\mathcal{W}_2^2(\rho_t, \rho_*) \leqslant 4\overline{\eta}\sigma^2\left(1 + L\eta + \frac{4\lambda\eta}{\sigma^2} + \frac{L^2\eta^2}{2}\right)K\mathcal{W}_2^2(\rho_0, \rho_*)\,.$$

for all $t \geqslant 1$ such that $\rho_t \in \mathcal{B}_{\mathcal{W}_2}(\rho_*, \delta_1')$. Using a synchronous coupling as in the proof of Theorem 8, it is easily seen that

$$\mathcal{W}_2(\rho_t, \rho_*) \leqslant e^{Ct}\mathcal{W}_2(\rho_0, \rho_*) \tag{70}$$

for some C depending only on M and L. As a consequence, we can take $\delta_2' \in (0, \delta_1']$ such that for all $\rho_0 \in \mathcal{B}_{\mathcal{W}_2}(\rho_*, \delta_2')$, it holds that $\rho_t \in \mathcal{B}_{\mathcal{W}_2}(\rho_*, \delta_1')$ for all $t \in [0, 1]$ and

$$\mathcal{W}_2(\rho_t, \rho_*) \leqslant \frac{\delta_1'}{2}$$

for all $t \geqslant 1$ such that $\rho_t \in \mathcal{B}_{\mathcal{W}_2}(\rho_*, \delta_1')$. Using that $t \mapsto \rho_t$ is continuous with respect to $\mathcal{W}_2$, we get by contradiction that $\rho_t \in \mathcal{B}_{\mathcal{W}_2}(\rho_*, \delta_1')$ for all $t \geqslant 0$ if $\rho_0 \in \mathcal{B}_{\mathcal{W}_2}(\rho_*, \delta_2')$, and in particular $T_{\mathcal{A}}(\rho_0) = +\infty$.

In the remaining of the proof we consider an initial condition $\rho_0 \in \mathcal{B}_{\mathcal{W}_2}(\rho_*, \delta_2')$. As we used in (22), (**k-NL-LSI**) implies that ρ_* is a minimizer of $\mathcal{F}_k$ over $\mathcal{B}_{\mathcal{W}_2}(\rho_*, \delta)$, so that $\mathcal{F}_k(\rho_t) \geqslant \mathcal{F}_k(\rho_*)$ for all $t \geqslant 0$. Applying Proposition 18 (but with initial condition ρ_t instead of ρ_0), we get that, for all $t \geqslant 0$,

$$\begin{aligned}\frac{\varepsilon^3}{2}\mathcal{I}\left(\rho_{t+1}|\Gamma(\rho_{t+1})\right) &\leqslant \mathcal{E}(1,\rho_{t+1}) - \mathcal{F}_k(\rho_{t+1}) \\ &\leqslant \mathcal{E}(0,\rho_t) - \mathcal{F}_k(\rho_*) \;=\; \mathcal{F}_k(\rho_t) - \mathcal{F}_k(\rho_*)\,,\end{aligned}$$

and thus, using that $t \mapsto \mathcal{F}_k(\rho_t)$ is non-increasing, there exists $R > 0$ such that

$$\mathcal{L}(\rho_{t+1}) \leqslant R\left(\mathcal{F}_k(\rho_t) - \mathcal{F}_k(\rho_*)\right)$$

for all $t \geqslant 0$. Since Lemma 17 together with (**k-NL-LSI**) implies an exponential decay of $\mathcal{L}(\rho_t)$ at some rate $c > 0$, we can bound, for $t \geqslant 1$,

$$\begin{aligned}\mathcal{F}_k(\rho_t) - \mathcal{F}_k(\rho_*) &\leqslant \mathcal{L}(\rho_t) \\ &\leqslant e^{-c(t-1)}\mathcal{L}(\rho_1) \;\leqslant\; Re^{-c(t-1)}\left(\mathcal{F}_k(\rho_0) - \mathcal{F}_k(\rho_*)\right)\,.\end{aligned}$$

For $t \in [0,1]$ we can simply use that $t \mapsto \mathcal{F}_k(\rho_t)$ is non-increasing, and this concludes the proof of (65). The results for $\mathcal{W}_2$ and the relative entropy then follow by combining (65) with the various bounds relating the different quantities. More specifically, (64) is obtained simply by (70) for $t \in [0,1]$ and, for $t \geqslant 1$, by combining (69), (65), (68) and Proposition 19 to bound

$$\mathcal{W}_2^2(\rho_t,\rho_*) \leqslant 4\overline{\eta}Ce^{-c(t-1)}\left(\mathcal{F}_k(\rho_1) - \mathcal{F}_k(\rho_*)\right) \leqslant C'e^{-ct}\mathcal{W}_2^2(\rho_0,\rho_*)$$

for some C'. Similarly, (66) is obtained by using first (20), then (64) and (65) and finally (68) and (69) to bound

$$\begin{aligned}\sigma^2\mathcal{H}(\rho_t|\rho_*) &\leqslant \mathcal{F}_k(\rho_t) - \mathcal{F}_k(\rho_*) + \lambda\mathcal{W}_2^2(\rho_t,\rho_*) \\ &\leqslant Ce^{-ct}\left(\mathcal{F}_k(\rho_0) - \mathcal{F}_k(\rho_*) + \lambda\mathcal{W}_2^2(\rho_0,\rho_*)\right) \\ &\leqslant C'e^{-ct}\mathcal{H}(\rho_0|\rho_*)\end{aligned}$$

for some C'. The proof of (67) uses the same ingredients. □

Example 3. *In the one-dimensional double-well case, Proposition 1 remains true if the granular media equation* (1) *is replaced by the Vlasov-Fokker-Planck equation* (60)*. In the subcritical case, it directly follows from Theorem 20, thanks to Proposition 13 and Remark 16. In the critical case, the arguments are as in Remark 12.*

5 Fast free energy decay for interacting particles

5.1 Settings and results

We consider the particle approximations of (8) and (60), namely the overdamped Langevin process $(X_t)_{t\geq 0}$ solving

$$\forall i \in [\![1,N]\!]\,, \qquad \mathrm{d}X_t^i = -\nabla E_{\pi(X_t)}(X_t^i)\mathrm{d}t + \sqrt{2}\sigma \mathrm{d}B_t^i \tag{71}$$

and the kinetic Langevin process $(Y_t,V_t)_{t\geq 0}$ solving

$$\forall i \in [\![1,N]\!]\,, \qquad \left\{\begin{array}{rcl}\mathrm{d}Y_t^i &=& V_t^i\mathrm{d}t \\ \mathrm{d}V_t^i &=& -\nabla E_{\pi(Y_t)}(Y_t^i)\mathrm{d}t - V_t^i\mathrm{d}t + \sqrt{2}\sigma\mathrm{d}B_t^i\,,\end{array}\right. \tag{72}$$

where in both cases $B^1, \dots, B^N$ are independent d-dimensional Brownian motions and

$$\pi(x) = \frac{1}{N} \sum_{i=1}^{N} \delta_{x_i}$$

stands for the empirical distribution of $x = (x_1, \dots, x_N) \in \mathbb{R}^{dN}$. Under (**Lip**) and, respectively, (**o-s-Lip**) for (71) and ($x-$**Lip**) for (72), both equations admit a unique global strong solution.

Considering first the overdamped case, denote by ρ_t^N the law of $X_t = (X_t^1, \dots, X_t^N)$ solving (71). Assuming that

$$\int_{\mathbb{R}^{dN}} \exp\left(-\frac{N}{\sigma^2}\mathcal{E}(\pi(x))\right) \mathrm{d}x < \infty \,, \tag{73}$$

we consider the Gibbs measure

$$\rho_\infty^N \propto \exp\left(-\frac{N}{\sigma^2}\mathcal{E}(\pi(x))\right) \mathrm{d}x \,.$$

Using that

$$\forall i \in [\![1, N]\!], \qquad \nabla_{x_i} N\mathcal{E}(\pi(x)) = \nabla E_{\pi(x)}(x) \,,$$

by the properties of the linear derivative, we see that (71) is an overdamped Langevin diffusion on $\mathbb{R}^{dN}$ with potential $N\mathcal{E}(\pi(x))$ and temperature σ^2. Its invariant measure is ρ_∞^N, and $t \mapsto \mathcal{H}(\rho_t^N|\rho_\infty^N)$ is decreasing (by Jensen inequality) and goes to zero as $t \to \infty$ under mild assumptions. Defining the N-particle free energy of $\rho^N \in \mathcal{P}_2(\mathbb{R}^{dN})$ as

$$\mathcal{F}^N(\rho^N) = \sigma^2 \mathcal{H}(\rho^N) + N \int_{\mathbb{R}^{dN}} \mathcal{E}(\pi(x)) \rho^N(\mathrm{d}x) \,,$$

we see that

$$\sigma^2 \mathcal{H}(\rho^N|\rho_\infty^N) = \mathcal{F}^N(\rho^N) - \mathcal{F}^N(\rho_\infty^N).$$

In particular, ρ_∞^N is the global minimizer of $\mathcal{F}^N$ and $t \mapsto \mathcal{F}^N(\rho_t^N|\rho_\infty^N)$ is non-increasing.

As studied in [22], when the non-linear dynamics (8) admits several stable stationary solutions, the log-Sobolev constant of ρ_∞^N (assuming that a LSI holds, for instance if $x \mapsto \mathcal{E}(\pi(x))$ is uniformly convex outside a compact set) goes to infinity as $N \to \infty$, and thus we cannot expect a fast convergence of $\mathcal{F}^N(\rho_t^N)$ to its infimum (which typically occurs at a time-scale of order e^{aN} for some $a > 0$). However, in short time, initialized with independent initial conditions, the law of the particle system stays close to the solution of the non-linear problem (8) and thus the local convergence of the latter to some critical point $\rho_* \in \mathcal{K}$ drives the initial behavior of the particle system. This gives the following.

Proposition 21. *Under Assumption 5, assume furthermore* (73) *and that there exists* $\lambda' \geqslant 0$ *such that for all* $\mu_0, \mu_1 \in \mathcal{P}_2(\mathbb{R}^d)$ *and* $t \in [0, 1]$,

$$\mathcal{E}(t\mu_0 + (1-t)\mu_1) \geqslant t\mathcal{E}(\mu_0) + (1-t)\mathcal{E}(\mu_1) - t(1-t)\lambda' \mathcal{W}_2^2(\mu_0, \mu_1) \,. \tag{74}$$

Let $\rho_0 \in \mathcal{P}_2(\mathbb{R}^d)$ *be such that the solution of* (8) *converges to* ρ_* *in long time. There exist* $C, \beta > 0$ *such that, for all* $N \in \mathbb{N}$ *and* $t \geqslant 1$, *considering the particle system* (71) *with initial distribution* $\rho_0^N = \rho_0^{\otimes N}$,

$$\frac{1}{N}\mathcal{F}^N(\rho_t^N) - \mathcal{F}(\rho_*) \leqslant C\left(N^{-\beta} + e^{-t/\overline{\eta}}\right) \,.$$

Notice that, in view of (17), (74) holds for instance in the case (16) if the functions r_k are L_k-Lipschitz with $\sum_{k\in\mathbb{N}} L_k^2 < \infty$.

Remark 17. *When the solution of the non-linear McKean-Vlasov equations converges to some ρ_*, under suitable conditions (in particular of regularity), a quantitative weak convergence at rate $1/N$ for the empirical distribution of the corresponding system of N interacting particles is shown in [21, Theorem 3.1] to hold over times of order N^p for any p. This requires the initial condition to start close to ρ_* in the total variation sense. However, thanks to Proposition 6, if the initial condition is close to ρ_* in $\mathcal{W}_2$, then it becomes close in the total variation sense after a time* 1.

The same occurs in the kinetic case. For $\rho^N \in \mathcal{P}_2(\mathbb{R}^{2dN})$, define the N-particle kinetic free energy as

$$\mathcal{F}_k^N(\rho^N) = \sigma^2 \mathcal{H}(\rho^N) N \int_{\mathbb{R}^{2dN}} \left[\mathcal{E}(\pi(x)) + \mathcal{E}_k(\pi(v))\right] \rho^N(\mathrm{d}x\mathrm{d}v)\,.$$

Proposition 22. *Under Assumption 9, assume furthermore that there exists $\lambda' \geqslant 0$ such that* (74) *holds for all $\mu_0, \mu_1 \in \mathcal{P}_2(\mathbb{R}^d)$ and $t \in [0,1]$. Let $\rho_0 \in \mathcal{P}_2(\mathbb{R}^{2d})$ be such that the solution of* (60) *converges to ρ_* in long time. There exist $C, \beta, \gamma > 0$ such that, for all $N \in \mathbb{N}$ and $t \geqslant 1$, considering the kinetic particle system* (72) *with initial distribution $\rho_0^N = \rho_0^{\otimes N}$,*

$$\frac{1}{N}\mathcal{F}_k^N(\rho_t^N) - \mathcal{F}_k(\rho_*) \leqslant C\left(N^{-\beta} + e^{-\gamma t}\right)\,.$$

5.2 Proofs

We mostly focus on the overdamped case in this section, denoting by ρ_t^N the law of (71). We start with the following variation of Proposition 6.

Proposition 23. *For any $\rho_* \in \mathcal{K}$, there exists $C > 0$ such that for all $N \in \mathbb{N}$, $t \geqslant 0$ and any initial distribution $\rho_0^N \in \mathcal{P}_2(\mathbb{R}^{dN})$,*

$$\mathcal{H}\left(\rho_{t+1}^N | \rho_*^{\otimes N}\right) \leqslant C\mathcal{W}_2^2\left(\rho_t^N, \rho_*^{\otimes N}\right) + C\,.$$

Similarly to Proposition 6, this follows from a Girsanov transform, see [30, Theorem 2.3] or [14, Lemma 5.3].

To relate $\mathcal{H}\left(\rho_t^N | \rho_*^{\otimes N}\right)$ to $\mathcal{F}^N(\rho_t^N)$, we can rely on the following.

Lemma 24. *Assume that there exists $\lambda' \geqslant 0$ such that* (74) *holds for all $\mu_0, \mu_1 \in \mathcal{P}_2(\mathbb{R}^d)$ and $t \in [0,1]$. Then, for all $\rho^N \in \mathcal{P}_2(\mathbb{R}^{dN})$ and $\rho_* \in \mathcal{K}$,*

$$\mathcal{F}^N(\rho^N) - N\mathcal{F}(\rho_*) \leqslant \sigma^2 \mathcal{H}\left(\rho^N | \rho_*^{\otimes N}\right) + 2\lambda' \mathcal{W}_2^2\left(\rho^N, \rho_*^{\otimes N}\right) + 2\lambda' N \alpha(N)$$

where

$$\alpha(N) = \int_{\mathbb{R}^{dN}} \mathcal{W}_2^2(\pi(x), \rho_*) \rho_*^{\otimes N}(\mathrm{d}x)\,.$$

Thanks to [24, Theorem 1], $\alpha(N) \leqslant CN^{-2/d}$ for some C (but possibly for some particular ρ_* it may vanish faster).

Proof. Dividing by t and sending t to zero in (74) yields

$$\int_{\mathbb{R}^d} E_{\mu_1}(\mu_0 - \mu_1) + \mathcal{E}(\mu_1) \geqslant \mathcal{E}(\mu_0) - \lambda \mathcal{W}_2^2(\mu_0, \mu_1)\,.$$

Then, using this with $\mu_0 = \pi(x)$ and $\mu_1 = \rho_*$,

$$\begin{aligned}
&\mathcal{F}^N(\rho^N) - N\mathcal{F}(\rho_*)\\
&\quad = N\int_{\mathbb{R}^{dN}} \left[\mathcal{E}(\pi(x)) - \mathcal{E}(\rho_*)\right]\rho^N(\mathrm{d}x) + \sigma^2\mathcal{H}(\rho^N) - N\sigma^2\mathcal{H}(\rho_*)\\
&\quad \leqslant N\int_{\mathbb{R}^{dN}} \left[\lambda'\mathcal{W}_2^2(\pi(x),\rho_*) + \int_{\mathbb{R}^{2d}} E_{\rho_*}(y)(\pi(x)-\rho_*)(\mathrm{d}y)\right]\rho^N(\mathrm{d}x) + \sigma^2\mathcal{H}(\rho^N) - N\sigma^2\mathcal{H}(\rho_*).
\end{aligned}$$

Using that $E_{\rho_*} = -\sigma^2(\ln\rho_* + \ln Z_{\rho_*})$,

$$\begin{aligned}
N\int_{\mathbb{R}^{dN}}\int_{\mathbb{R}^{2d}} E_{\rho_*}(y)(\pi(x)-\rho_*)(\mathrm{d}y)\rho^N(\mathrm{d}x) &= N\sigma^2\mathcal{H}(\rho_*) - \sigma^2\sum_{i=1}^N \int_{\mathbb{R}^{dN}} \ln\rho_*(x_i)\rho^N(\mathrm{d}x)\\
&= N\sigma^2\mathcal{H}(\rho_*) - \sigma^2\int_{\mathbb{R}^{dN}} \ln\rho_*^{\otimes N}(x)\rho^N(\mathrm{d}x)\,.
\end{aligned}$$

Plugging this in the previous inequality reads

$$\mathcal{F}^N(\rho^N) - N\mathcal{F}(\rho_*) \leqslant \sigma^2\mathcal{H}\left(\rho_N|\rho_*^{\otimes N}\right) + N\lambda'\int_{\mathbb{R}^{dN}} \mathcal{W}_2^2(\pi(x),\rho_*)\rho^N(\mathrm{d}x)\,.$$

Let $\mu(\mathrm{d}x\mathrm{d}y)$ be an optimal $\mathcal{W}_2$ coupling of ρ^N and $\rho_*^{\otimes N}$. We bound

$$\begin{aligned}
\int_{\mathbb{R}^{dN}} \mathcal{W}_2^2(\pi(x),\rho_*)\rho^N(\mathrm{d}x) &\leqslant 2\int_{\mathbb{R}^{dN}} \mathcal{W}_2^2(\pi(x),\pi(y))\mu(\mathrm{d}x\mathrm{d}y) + 2\alpha(N)\\
&\leqslant \frac{2}{N}\sum_{i=1}^N \int_{\mathbb{R}^{dN}} |x_i - y_i|^2\mu(\mathrm{d}x\mathrm{d}y) + 2\alpha(N)\\
&= 2\mathcal{W}_2^2\left(\rho^N,\rho_*^{\otimes N}\right) + 2\alpha(N)\,,
\end{aligned}$$

which completes the proof. □

Proof of Proposition 21. Thanks to (31) and standard finite propagation of chaos estimates (obtained by synchronous coupling), we bound

$$\begin{aligned}
\mathcal{W}_2^2\left(\rho_t^N,\rho_*^{\otimes N}\right) &\leqslant 2\mathcal{W}_2^2\left(\rho_t^N,\rho_t^{\otimes N}\right) + 2\mathcal{W}_2^2\left(\rho_t^{\otimes N},\rho_*^{\otimes N}\right)\\
&\leqslant Ce^{Ct} + 2NC_0e^{-t/\overline{\eta}}
\end{aligned}$$

for some $C > 0$. Proposition 23 and Lemma 24 yield

$$\mathcal{F}^N(\rho_t^N) - N\mathcal{F}(\rho_*) \leqslant C'\left(e^{Ct} + Ne^{-t/\overline{\eta}} + N\alpha(N)\right)$$

for all $t \geqslant 1$ for some $C' > 0$. Conclusion follows by using that $t \mapsto \mathcal{F}^N(\rho_t^N)$ is non-decreasing (distinguishing whether t is above or under $\ln N/(2C)$) and applying [24, Theorem 1] to bound $\alpha(N)$. □

Proof of Proposition 22. The proof is the same as the proof of Proposition 21. Proposition 23 still holds in the kinetic case (with $\rho_* \in \mathcal{K}_k$; this is [14, Lemma 5.3]), and so does Lemma 24 (replacing $\mathcal{F}$ by $\mathcal{F}_k$, with the same proof, noticing that the kinetic energy is linear and thus satisfies (74) with $\lambda' = 0$). We use Theorem 20 instead of Theorem 8. □

Acknowledgments

The research of P. Monmarché is supported by the projects SWIDIMS (ANR-20-CE40-0022) and CONVIVIALITY (ANR-23-CE40-0003) of the French National Research Agency. The research of J. Reygner is supported by the projects QUAMPROCS (ANR-19-CE40-0010) and CONVIVIALITY (ANR-23-CE40-0003) of the French National Research Agency.

References

[1] Luigi Ambrosio, Nicola Gigli, and Giuseppe Savaré. *Gradient flows: in metric spaces and in the space of probability measures.* Springer Science & Business Media, 2005.

[2] Dominique Bakry, Ivan Gentil, and Michel Ledoux. *Analysis and geometry of Markov diffusion operators*, volume 348 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences].* Springer, Cham, 2014.

[3] Florent Barret, Anton Bovier, and Sylvie Méléard. Uniform Estimates for Metastable Transition Times in a Coupled Bistable System. *Electronic Journal of Probability*, 15(none):323 – 345, 2010.

[4] Kaveh Bashiri. On the long-time behaviour of McKean-Vlasov paths. *Electronic Communications in Probability*, 25(none):1 – 14, 2020.

[5] Kaveh Bashiri and Georg Menz. Metastability in a continuous mean-field model at low temperature and strong interaction. *Stochastic Processes and their Applications*, 134:132–173, 2021.

[6] Michel Benaïm, Charles-Edouard Bréhier, and Pierre Monmarché. Analysis of an adaptive biasing force method based on self-interacting dynamics. *Electron. J. Probab.*, 25:28 pp., 2020.

[7] Jean-David Benamou and Yann Brenier. A computational fluid mechanics solution to the monge-kantorovich mass transfer problem. *Numerische Mathematik*, 84(3):375–393, 2000.

[8] Evan Camrud, Aalin Oliviero Durmus, Pierre Monmarché, and Gabriel Stoltz. Second order quantitative bounds for unadjusted generalized Hamiltonian Monte Carlo. *arXiv e-prints*, 2023.

[9] José A Carrillo, Katy Craig, Li Wang, and Chaozhen Wei. Primal dual methods for wasserstein gradient flows. *Foundations of Computational Mathematics*, pages 1–55, 2022.

[10] José A Carrillo, Rishabh S Gvalani, and Jeremy S-H Wu. An invariance principle for gradient flows in the space of probability measures. *Journal of Differential Equations*, 345:233–284, 2023.

[11] José A. Carrillo, Robert J. McCann, and Cédric Villani. Kinetic equilibration rates for granular media and related equations: entropy dissipation and mass transportation estimates. *Revista Matemática Iberoamericana*, 19(3):971 – 1018, 2003.

[12] Jose Antonio Carrillo, Rishabh S Gvalani, Grigorios A Pavliotis, and Andre Schlichting. Long-time behaviour and phase transitions for the mckean–vlasov equation on the torus. *Archive for Rational Mechanics and Analysis*, 235(1):635–690, 2020.

[13] Paul-Eric Chaudru de Raynal, Ma Hong Duong, Pierre Monmarché, Milica Tomašević, and Julian Tugaut. Reducing exit-times of diffusions with repulsive interactions. *ESAIM: PS*, 27:723–748, 2023.

[14] Fan Chen, Yiqing Lin, Zhenjie Ren, and Songbo Wang. Uniform-in-time propagation of chaos for kinetic mean field Langevin dynamics. *arXiv e-prints*, page arXiv:2307.02168, July 2023.

[15] Fan Chen, Zhenjie Ren, and Songbo Wang. Uniform-in-time propagation of chaos for mean field langevin dynamics. *arXiv preprint arXiv:2212.03050*, 2022.

[16] Lénaïc Chizat. Mean-field langevin dynamics : Exponential convergence and annealing. *Transactions on Machine Learning Research*, 2022.

[17] Lenaic Chizat and Francis Bach. On the global convergence of gradient descent for over-parameterized models using optimal transport. *Advances in neural information processing systems*, 31, 2018.

[18] Giovanni Conforti, Daniel Lacker, and Soumik Pal. Projected Langevin dynamics and a gradient flow for entropic optimal transport. *arXiv e-prints*, page arXiv:2309.08598, September 2023.

[19] Quentin Cormier. On the stability of the invariant probability measures of mckean-vlasov equations. *arXiv preprint arXiv:2201.11612*, 2022.

[20] Donald A Dawson and J Gärtner. Large deviations and tunnelling for particle systems with mean field interaction. *CR Math. Rep. Acad. Sci. Canada*, 8(6):387–392, 1986.

[21] François Delarue and Alvin Tse. Uniform in time weak propagation of chaos on the torus. *arXiv preprint arXiv:2104.14973*, 2021.

[22] Matías G Delgadino, Rishabh S Gvalani, Grigorios A Pavliotis, and Scott A Smith. Phase transitions, logarithmic sobolev inequalities, and uniform-in-time propagation of chaos for weakly interacting diffusions. *Communications in Mathematical Physics*, pages 1–49, 2023.

[23] Jean Dolbeault, Bruno Nazaret, and Giuseppe Savaré. A new class of transport distances between measures. *Calculus of Variations and Partial Differential Equations*, 34(2):193–231, 2009.

[24] Nicolas Fournier and Arnaud Guillin. On the rate of convergence in wasserstein distance of the empirical measure. *Probability theory and related fields*, 162(3-4):707–738, 2015.

[25] Arnaud Guillin, Pierre Le Bris, and Pierre Monmarché. Some remarks on the effect of the Random Batch Method on phase transition. *arXiv e-prints*, page arXiv:2308.01624, August 2023.

[26] 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(3):1590 – 1614, 2022.

[27] Arnaud Guillin and Pierre Monmarché. Uniform long-time and propagation of chaos estimates for mean field kinetic particles in non-convex landscapes. *Journal of Statistical Physics*, 185:1–20, 2021.

[28] Rishabh S Gvalani and André Schlichting. Barriers of the mckean–vlasov energy via a mountain pass theorem in the space of probability measures. *Journal of Functional Analysis*, 279(11):108720, 2020.

[29] Kaitong Hu, Zhenjie Ren, David Šiška, and Łukasz Szpruch. Mean-field Langevin dynamics and energy landscape of neural networks. *Annales de l'Institut Henri Poincaré, Probabilités et Statistiques*, 57(4):2043 – 2065, 2021.

[30] Xing Huang. Coupling by Change of Measure for Conditional McKean-Vlasov SDEs and Applications. *arXiv e-prints*, page arXiv:2303.04369, March 2023.

[31] Lucas Journel and Pierre Monmarché. Convergence of the kinetic annealing for general potentials. *Electronic Journal of Probability*, 27:1–37, 2022.

[32] Daniel Lacker. Independent projections of diffusions: Gradient flows for variational inference and optimal mean field approximations. *arXiv e-prints*, page arXiv:2309.13332, September 2023.

[33] Marc Lambert, Sinho Chewi, Francis Bach, Silvère Bonnabel, and Philippe Rigollet. Variational inference via wasserstein gradient flows. *Advances in Neural Information Processing Systems*, 35:14434–14447, 2022.

[34] Lingxiao Li, Qiang Liu, Anna Korba, Mikhail Yurochkin, and Justin Solomon. Sampling with mollified interaction energy descent. In *The Eleventh International Conference on Learning Representations*, 2023.

[35] Song Mei, Theodor Misiakiewicz, and Andrea Montanari. Mean-field theory of two-layers neural networks: dimension-free bounds and kernel limit. In *Conference on Learning Theory*, pages 2388–2464. PMLR, 2019.

[36] Song Mei, Andrea Montanari, and Phan-Minh Nguyen. A mean field view of the landscape of two-layer neural networks. *Proceedings of the National Academy of Sciences*, 115(33):E7665–E7671, 2018.

[37] Pierre Monmarché. Hypocoercivity in metastable settings and kinetic simulated annealing. *Probability Theory and Related Fields*, 172(3):1215–1248, 2018.

[38] Pierre Monmarché. An entropic approach for Hamiltonian Monte Carlo: the idealized case. *to appear in Annals of Applied Probability*, 2023.

[39] Pierre Monmarché. A note on a Vlasov-Fokker-Planck equation with non-symmetric interaction. *To appear in Annales de la Faculté des Sciences de Toulouse*, 2024.

[40] Pierre Monmarché. Long-time behaviour and propagation of chaos for mean field kinetic particles. *Stochastic Processes and their Applications*, 127(6):1721–1737, 2017.

[41] Felix Otto and Cédric Villani. Generalization of an inequality by talagrand and links with the logarithmic sobolev inequality. *Journal of Functional Analysis*, 173(2):361–400, 2000.

[42] Zhongmin Qian, Panpan Ren, and Feng-Yu Wang. Entropy Estimate for Degenerate SDEs with Applications to Nonlinear Kinetic Fokker-Planck Equations. *arXiv e-prints*, page arXiv:2308.16373, August 2023.

[43] Panpan Ren and Feng-Yu Wang. Exponential convergence in entropy and wasserstein for mckean–vlasov sdes. *Nonlinear Analysis*, 206:112259, 2021.

[44] Bin Shi, Weijie Su, and Michael I Jordan. On learning rates and schrödinger operators. *Journal of Machine Learning Research*, 24(379):1–53, 2023.

[45] Julian Tugaut. Convergence to the equilibria for self-stabilizing processes in double-well landscape. *Ann. Probab.*, 41(3A):1427–1460, 2013.

[46] Julian Tugaut. Phase transitions of mckean–vlasov processes in double-wells landscape. *Stochastics An International Journal of Probability and Stochastic Processes*, 86(2):257–284, 2014.

[47] Julian Tugaut. Self-stabilizing processes in multi-wells landscape in $\mathbb{R}^d$-invariant probabilities. *Journal of Theoretical Probability*, 27:57–79, 2014.

[48] Julian Tugaut. Granular media equation with double-well external landscape: limiting steady state. *arXiv preprint arXiv:2202.13395*, 2022.

[49] Julian Tugaut. On the steady states for the granular media equation: existence, local uniqueness, local stability and rate of convergence. *preprint HAL*, 2023.

[50] Feng-Yu Wang. Distribution dependent SDEs for Landau type equations. *Stochastic Processes and their Applications*, 128(2):595–621, 2018.