



**HAL**  
open science

# An elementary approach to uniform in time propagation of chaos

Alain Durmus, Andreas Eberle, Arnaud Guillin, Raphael Zimmer

► **To cite this version:**

Alain Durmus, Andreas Eberle, Arnaud Guillin, Raphael Zimmer. An elementary approach to uniform in time propagation of chaos. Proceedings of the American Mathematical Society, 2020, 148 (12), pp.5387-5398. 10.1090/proc/14612 . hal-01801275

**HAL Id: hal-01801275**

**<https://hal.science/hal-01801275>**

Submitted on 28 May 2018

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# AN ELEMENTARY APPROACH TO UNIFORM IN TIME PROPAGATION OF CHAOS

ALAIN DURMUS, ANDREAS EBERLE, ARNAUD GUILLIN, RAPHAEL ZIMMER

ABSTRACT. Based on a coupling approach, we prove uniform in time propagation of chaos for weakly interacting mean-field particle systems with possibly non-convex confinement and interaction potentials. The approach is based on a combination of reflection and synchronous couplings applied to the individual particles. It provides explicit quantitative bounds that significantly extend previous results for the convex case.

## 1. INTRODUCTION

Let  $V, W : \mathbb{R}^d \rightarrow \mathbb{R}$  be twice continuously differentiable functions satisfying appropriate regularity and growth conditions. We consider the *kinetic Fokker-Planck equation*

$$(1) \quad \partial_t \mu_t = \nabla \cdot [\nabla \mu_t + (\nabla V + \nabla W * \mu_t) \mu_t]$$

and its probabilistic counterpart, the nonlinear stochastic differential equation

$$(2) \quad d\bar{X}_t = -\nabla V(\bar{X}_t) dt - \nabla W * \mu_t(\bar{X}_t) dt + \sqrt{2} dB_t, \quad \mu_t = \mathcal{L}(\bar{X}_t),$$

of *McKean-Vlasov* type. Here  $\mu_t$  is a time dependent probability measure on  $\mathbb{R}^d$  and  $*$  denotes the standard convolution operator. The function  $V$  corresponds to a *confinement potential* and the function  $W$  to an *interaction potential*. Variants of the equation occur for example in the modelling of granular media, cf. [1, 34].

Both in the probability and in the p.d.e. community, existence and uniqueness of (1) and (2) have attracted much attention, see [23, 16, 29, 24] for a few milestones, and [27] and [17] for two recent results. During the last twenty years, there has been a lot of progress on convergence to equilibrium of solutions  $(\mu_t)_{t \geq 0}$  of (1). Carrillo, McCann and Villani [7, 8] have proven an exponential convergence rate under the strict convexity condition  $\text{Hess}(V + 2W) \geq \rho \text{Id}$  with  $\rho > 0$ . They have also established a polynomial convergence rate in the case where  $V + 2W$  is only degenerately strictly convex with  $\text{Hess}(V + 2W)(x) = 0$  for some isolated points  $x \in \mathbb{R}^d$ . Malrieu [22] and Cattiaux et al. [11] have developed a probabilistic approach to these results that is based on an approximation by the mean-field particle system which is defined for  $N \in \mathbb{N}$  as the solution of the equations

$$(3) \quad dX_t^{i,N} = -\nabla V(X_t^{i,N}) dt - N^{-1} \sum_{j=1}^N \nabla W(X_t^{i,N} - X_t^{j,N}) dt + \sqrt{2} dB_t^i, \quad i = 1, \dots, N,$$

where the initial values  $X_0^1, \dots, X_0^N$  are i.i.d. random variables, and the processes  $(B_t^1)_{t \geq 0}, \dots, (B_t^N)_{t \geq 0}$  are independent Brownian motions. It is well-known [29, 24] that under

weak assumptions on  $V$  and  $W$ , the laws of the particles at time  $t$  converge to the solution  $\mu_t$  of (1) as  $N \rightarrow \infty$ . In [22, 11], both convergence to equilibrium for the nonlinear SDE (2) and uniform in time propagation of chaos for the particle system have been proven under convexity assumptions by using functional inequalities and synchronous couplings, respectively.

When  $V + 2W$  is not convex, the situation is much more delicate. Uniqueness of a stationary solution of (1) does not hold in general without additional conditions on  $V$  and  $W$ . Even in this case, only few results on convergence to equilibrium are known. By a direct study of the dissipation of the Wasserstein distance, Bolley, Gentil and Guillin [2] established an exponential trend to equilibrium in a weakly non-convex case. In a recent work by three of the authors [14], an exponential contraction property and thus convergence to equilibrium could be shown for a much broader class of potentials. The proof is based on a new coupling approach originating in [13], which is also the basis of this work. An interesting related problem arises when there are multiple invariant measures, see [30]. In this widely open case, a main interest are the relative basins of attraction of the equilibria.

The convergence of the empirical measure of the particle system (3) to the solution of (1), or, equivalently, the convergence of the pair empirical measure to the tensor product of two solutions of (1), has been stated under the name of *propagation of chaos* by Kac [20], and further developed by Sznitman [29]. However, the corresponding general results are only valid uniformly for a fixed time horizon. A crucial point is then not only to assert the convergence of the empirical measures but to quantify it. A remarkable analytic framework providing also quantitative propagation of chaos estimates has been developed by Mischler, Mouhot and their coauthors in connection with Kac's program in kinetic theory [25, 18, 26].

In this article, we propose a very different and much more elementary, probabilistic approach to quantitative bounds for propagation of chaos. Our main result is a *uniform in time propagation of chaos* bound that takes the form

$$\mathcal{W}_{\ell^1(f)} \left( \mathcal{L}(X_t^{1,N}, \dots, X_t^{N,N}), \mu_t^{\otimes N} \right) \leq A \cdot N^{-1/2} \quad \text{for all } t \geq 0 \text{ and } n \in \mathbb{N},$$

if  $X_0^1, \dots, X_0^n$  are i.i.d. with initial law  $\mu_0$ , see Theorem 2 below. Here  $\mathcal{W}_{\ell^1(f)}$  is an  $L^1$  Wasserstein distance on  $(\mathbb{R}^d)^N$  and  $A$  is an explicit finite constant, see below for the precise definitions. The bound holds under similar assumptions as the quantitative bounds on convergence to equilibrium for McKean-Vlasov equations in [14]. In particular, it applies in non-convex cases provided the confinement potential  $V$  is strictly convex outside a ball, and the interaction potential  $W$  is symmetric and globally Lipschitz continuous with sufficiently small Lipschitz constant. Consequently, our results are significant extensions of the uniform in time propagation of chaos results in the convex case in [22, 11]. The main difference to the approach in these works is that we use a more refined coupling which is harder to analyse but much more powerful than synchronous coupling. The second main ingredient in our proofs is an adequately constructed  $L^1$  Wasserstein distance that is well adapted to the couplings we consider. Our probabilistic approach is very different from the analytic methods developed in [25, 18, 26], and the conditions required are not easily comparable. One advantage of the coupling approach presented

here is that it is very simple and quite robust. This might facilitate the application to other classes of models. For example, the same argument can be applied immediately if  $\nabla V$  is replaced by a non-gradient drift  $\beta$  that satisfies corresponding assumptions. Similarly,  $\nabla W$  can be replaced by a more general interaction term, see Remark 4 below.

In Section 2, we present our main hypotheses and results concerning the uniform in time propagation of chaos. The proofs are provided in Section 3. Our approach is based on an interplay between reflection and synchronous coupling. It relies heavily on the framework introduced in [13, 14]. Our results use properties of the confinement potential, which may nevertheless possess many wells. The interactions are mainly seen as a perturbation. An interesting and challenging problem to be taken up in forthcoming work is to prove propagation of chaos in situations where  $V = 0$  and  $W$  is not convex, but uniqueness of a stationary distribution holds.

## 2. UNIFORM IN TIME PROPAGATION OF CHAOS

In this section we state our main results. We will first state the precise assumptions on the potentials  $V$  and  $W$ . Moreover, we define some functions and parameters that will determine the particular Wasserstein distance that we consider.

**2.1. Hypotheses and definitions.** We first state our assumption on the confinement potential.

**H1.** *There is a continuous function  $\kappa : [0, \infty) \rightarrow \mathbb{R}$  satisfying  $\liminf_{r \rightarrow \infty} \kappa(r) > 0$  such that*

$$(4) \quad \langle \nabla V(x) - \nabla V(y), x - y \rangle \leq \kappa(\|x - y\|) \|x - y\|^2 \quad \text{for all } x, y \in \mathbb{R}^d .$$

Note that under **H1**, there exist  $m_V > 0$  and  $M_V \geq 0$  such that for all  $x, y \in \mathbb{R}^d$ ,

$$(5) \quad \langle \nabla V(x) - \nabla V(y), x - y \rangle \geq m_V \|x - y\|^2 - M_V .$$

Following the framework introduced in [13], we now define constants  $R_0$  and  $R_1$  as

$$(6) \quad R_0 := \inf \{s \in \mathbb{R}_+ : \kappa(r) \geq 0 \text{ for all } r \geq s\} ,$$

$$(7) \quad R_1 := \inf \{s \geq R_0 : s(s - R_0)\kappa(r) \geq 8 \text{ for all } r \geq s\} ,$$

and we consider the functions  $\varphi, \Phi, g : \mathbb{R}_+ \rightarrow \mathbb{R}_+$  defined by

$$(8) \quad \varphi(r) = \exp\left(-\frac{1}{4} \int_0^r s \kappa_-(s) ds\right) , \quad \Phi(r) = \int_0^r \varphi(s) ds ,$$

$$(9) \quad g(r) = 1 - \frac{c}{2} \int_0^{r \wedge R_1} \Phi(s) \varphi^{-1}(s) ds ,$$

where  $\kappa_- = \max(0, -\kappa)$ , and

$$(10) \quad c = 1 / \int_0^{R_1} \Phi(s) \varphi^{-1}(s) ds .$$

Note that  $\varphi(r) = \varphi(R_0)$  for  $r \geq R_0$ , and  $g(r) = 1/2$  for  $r \geq R_1$ . In addition, for all  $r \in \mathbb{R}_+$ ,  $g(r) \in [1/2, 1]$ . We now define an increasing function  $f : [0, \infty) \rightarrow [0, \infty)$  by

$$(11) \quad f(r) = \int_0^r \varphi(s) g(s) ds .$$

Since  $\varphi$  and  $g$  are decreasing,  $f$  is concave, for all  $r \geq 0$ ,

$$(12) \quad \varphi(R_0)r/2 \leq \Phi(r)/2 \leq f(r) \leq \Phi(r) \leq r .$$

Note that by (12),  $(x, y) \mapsto f(\|x - y\|)$  induces a distance that is equivalent to the Euclidean distance on  $\mathbb{R}^d$ . Below, we will use contraction properties in  $L^1$  Wasserstein distances based on the underlying distance  $f(\|x - y\|)$ . These contraction properties will be a consequence of the inequality

$$(13) \quad f''(r) - r\kappa(r)f'(r)/4 \leq -cf(r)/2 \quad \text{for all } r \in \mathbb{R}_+ \setminus \{R_1\} .$$

Indeed, notice that (13) is satisfied for  $r \in [0, R_1)$  by the definitions and since  $f \leq \Phi$ . Moreover, for  $r > R_1$ ,  $f(r) = f(R_1) + \varphi(R_0)(r - R_1)/2$ , and thus by definition of  $R_1$ ,

$$\begin{aligned} f''(r) - r\kappa(r)f'(r)/4 &= -r\kappa(r)\varphi(R_0)/8 \leq -r(R_1(R_1 - R_0))^{-1}\varphi(R_0) \\ &\leq -\Phi(r)(\Phi(R_1)(R_1 - R_0))^{-1}\varphi(R_0) \leq -\frac{1}{2}\Phi(r) \left/ \int_{R_0}^{R_1} \Phi(s)\varphi^{-1}(s)ds \right. \leq -cf(r)/2 . \end{aligned}$$

Here we have used that for  $r \geq R_0$ ,  $\varphi$  is constant,  $\Phi(r) = \Phi(R_0) + (r - R_0)\varphi(R_0)$ , and

$$\begin{aligned} \int_{R_0}^{R_1} \Phi(s)\varphi(s)^{-1} ds &= \Phi(R_0)\varphi(R_0)^{-1}(R_1 - R_0) + (R_1 - R_0)^2/2 \\ &\geq (R_1 - R_0)(\Phi(R_0) + (R_1 - R_0)\varphi(R_0))\varphi(R_0)^{-1}/2 = (R_1 - R_0)\Phi(R_1)\varphi(R_0)^{-1}/2 . \end{aligned}$$

Next, we state the assumptions on the interaction potential.

- H2.** (i)  $W$  is symmetric, i.e.,  $W(x) = W(-x)$  for all  $x \in \mathbb{R}^d$ .  
(ii) There exists  $\eta \in (0, c)$  such that for all  $x, y \in \mathbb{R}^d$ ,

$$\|\nabla W(x) - \nabla W(y)\| \leq \eta f(\|x - y\|) ,$$

where  $f$  and  $c$  are defined by (11) and (10).

Notice that by (12), a sufficient condition for **H2-(ii)** is that  $\nabla W$  is  $L$ -Lipschitz continuous, with  $L < 2c/\varphi(R_0)$ . By **H2-(i)**,  $\nabla W(0) = 0$ . Thus, since  $f' \leq 1$ , **H2** implies

$$(14) \quad \|\nabla W(x)\| \leq \eta \|x\| .$$

We consider also the following additional condition:

- H3.** There exists  $M_W \in [0, \infty)$  such that for all  $x, y \in \mathbb{R}^d$ ,

$$\langle \nabla W(x) - \nabla W(y), x - y \rangle \geq -M_W .$$

Under **H1** and **H2**, the equations (2) and (3) both have unique strong solutions  $(\bar{X}_t)_{t \geq 0}$  and  $\{(X_t^i)_{t \geq 0}, i = 1, \dots, N\}$  if we suppose sufficient integrability assumptions on the initial measures (say moments of order 4), see e.g. [11, Theorem 2.6]. In addition,

$$(15) \quad \sup_{t \in [0, T]} \mathbb{E} [\|\bar{X}_t\|^2 + \|X_t^1\|^2] < \infty \quad \text{for all } T \geq 0 .$$

**Example 1** (Double-well potential). A natural example which satisfies the assumptions above is the case of a double well confinement potential with quadratic interaction

$$V(x) = \|x\|^4 - a\|x\|^2, \quad W(x) = \pm\|x\|^2, \quad a > 0 ,$$

where the sign of the interaction implies attractiveness or repulsion. Using [13, Lemma 1 and Example 4], one has that the constant  $c$  in (13), which is crucial for the rate of

convergence to equilibrium, is of order  $\Theta(1)$  for small  $a$ , whereas for large  $a$ ,  $\log(c^{-1})$  is of order  $\Theta(a^2)$ .

**2.2. Main results.** Let  $\mathcal{W}_f$  denote the Kantorovich ( $L^1$  Wasserstein) distance on probability measures based on the underlying distance function  $f(\|x - y\|)$  on  $\mathbb{R}^d$ , i.e.,

$$\mathcal{W}_f(\nu, \mu) = \inf_{\xi \in \Pi(\nu, \mu)} \int_{\mathbb{R}^d \times \mathbb{R}^d} f(\|x - y\|) d\xi(x, y),$$

where  $\Pi(\nu, \mu)$  is the set of couplings of  $\mu$  and  $\nu$ , i.e., probability measures  $\xi$  on  $\mathcal{B}(\mathbb{R}^d \times \mathbb{R}^d)$  such that for all  $A \in \mathcal{B}(\mathbb{R}^d)$ ,  $\xi(A \times \mathbb{R}^d) = \mu(A)$  and  $\xi(\mathbb{R}^d \times A) = \nu(A)$ . In the case where  $f$  is the identity function  $f(r) \equiv r$ ,  $\mathcal{W}_f$  is the usual  $L^1$  Wasserstein distance  $\mathcal{W}_1$ . We also consider corresponding Kantorovich distances on probability measures  $\hat{\nu}, \hat{\mu}$  on the configuration space  $(\mathbb{R}^d)^N$ . Here we set

$$\mathcal{W}_{\ell^1(f)}(\hat{\nu}, \hat{\mu}) = \inf_{\xi \in \Pi(\hat{\nu}, \hat{\mu})} \int_{\mathbb{R}^{N \cdot d} \times \mathbb{R}^{N \cdot d}} \frac{1}{N} \sum_{i=1}^N f(\|x^i - y^i\|) d\xi(x, y).$$

For  $t \geq 0$  and  $\nu, \mu \in \mathcal{P}(\mathbb{R}^d)$ , we denote by  $\bar{\mu}_t^\nu$  and  $\mu_t^{\mu, N}$  the laws of  $\bar{X}_t$  and  $X_t^{i, N}$ , for an arbitrary  $i$ , with initial distributions  $\nu$  and  $\mu$  ( $\mu$  being the common distribution of the i.i.d. random variables  $X_0^{i, N}$ ,  $i \in \{1, \dots, N\}$ ). We now state our main result.

**Theorem 2.** *Assume **H1** and **H2**, and suppose that **H3** is satisfied or  $\eta \in (0, m_V/2)$ , where  $m_V$  is given by (5). Let  $\nu$  and  $\mu$  be probability measures on  $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$  which admit a finite fourth moment. Then for all  $t \geq 0$  and  $N \in \mathbb{N}$ , we have*

$$\mathcal{W}_f(\mu_t^{\mu, N}, \bar{\mu}_t^\nu) \leq e^{-2(c-\eta)t} \mathcal{W}_f(\nu, \mu) + (2(c-\eta))^{-1} C \eta N^{-1/2},$$

and, more generally,

$$\mathcal{W}_{\ell^1(f)}(\mathcal{L}(X_t^{1, N}, \dots, X_t^{N, N}), (\bar{\mu}_t^\nu)^{\otimes N}) \leq e^{-2(c-\eta)t} \mathcal{W}_f(\nu, \mu) + (2(c-\eta))^{-1} C \eta N^{-1/2}.$$

Here  $f$  and  $c$  are defined by (11) and (10), respectively, and  $C$  is a constant that depends only on the dimension  $d$ , the second moment of  $\nu$ , as well as  $V$  and  $W$ .

By (12) and the definition of  $\mathcal{W}_f$ , we immediately obtain:

**Corollary 3.** *Under the same hypotheses as in Theorem 2, for all  $t \geq 0$  and  $N \in \mathbb{N}$ ,*

$$\mathcal{W}_1(\mu_t^{\mu, N}, \bar{\mu}_t^\nu) \leq 2\varphi(R_0)^{-1} e^{-2(c-\eta)t} \mathcal{W}_f(\nu, \mu) + (\varphi(R_0)(c-\eta))^{-1} C \eta N^{-1/2},$$

where  $R_0$  and  $\varphi$  are defined by (6) and (8), respectively, and  $C$  is a constant only depending on  $d$ , the second moment of  $\nu$ , as well as  $V$  and  $W$ .

The dependence of the bounds on  $N$  is of the right order and consistent with both the non uniform in time estimates, and the uniform in time bounds under uniform strict convexity for  $L^2$  Wasserstein distances obtained for example in [29, 22].

**Remark 4.** *For the sake of clarity and comparison with other recent works on uniform propagation of chaos [22, 11], we have restricted ourselves to interaction drifts of the type  $b(x, \mu) = -\nabla W * \mu$ . Similarly to [14], the present result can be extended to  $b(x, \mu) = \int b(x, y) \mu(dy)$  if  $b$  satisfies a Lipschitz condition with sufficiently small Lipschitz constant. Similarly, the confinement drift  $-\nabla V$  can be replaced by a non-gradient drift satisfying a corresponding assumption as **H1**.*

**Remark 5.** In [7, 22, 11], it has been shown that for a fixed center of mass, exponential contractivity of the nonlinear SDE and uniform propagation of chaos hold if  $V + 2W$  is, for example, strictly uniformly convex. This suggests that convexity of the interaction potential  $W$  can make up for non-convexity of  $V$ , a fact that is not visible from our current approach. The problem is that a symmetrization trick is used to get a benefit from the convexity of the interaction potential. This trick does not carry over in the same form to the  $\ell_1$  type distances considered here. We will address this challenging issue in a future work.

**Remark 6.** By combining the result of Theorem 2 and the contraction result for mean-field particle systems in [13, Corollary 3.4], one can derive a contraction result for the non-linear equation (1), recovering essentially [14, Theorem 2.3].

### 3. PROOFS

The proof of the main result is based on three ingredients. As usual, the starting construction is to consider i.i.d. copies of the nonlinear process and to couple them with the system of particles. This will be done by considering a reflection coupling coordinate by coordinate. The second ingredient will then be to consider a specific Kantorovich ( $L^1$  Wasserstein) distance based on an  $\ell_1$  type metric on the product space that is particularly suited to this coupling. The last ingredient is a law of large numbers type control in  $L^2$ . Let us first define the reflection coupling between the independent nonlinear processes and the mean-field particle system.

**3.1. Coupling by reflection.** For every  $\delta > 0$ , we consider Lipschitz continuous functions  $\phi_s^\delta, \phi_r^\delta : \mathbb{R}^d \rightarrow \mathbb{R}$  satisfying

$$(16) \quad (\phi_s^\delta)^2(x) + (\phi_r^\delta)^2(x) = 1 \quad \text{for all } x \in \mathbb{R}^d, \quad \phi_r^\delta(x) = \begin{cases} 1 & \text{if } \|x\| \geq \delta, \\ 0 & \text{if } \|x\| \leq \delta/2. \end{cases}$$

Now fix  $\delta > 0$ , probability measures  $\nu, \mu$  on  $\mathbb{R}^d$  with finite fourth moment, and a coupling  $\xi \in \Pi(\nu, \mu)$ . We consider the coupling between the independent nonlinear processes and the mean-field particle system defined by the following system of stochastic differential equations:

$$\begin{aligned} d\bar{X}_s^i &= -\nabla V(\bar{X}_s^i)ds - \nabla W * \bar{\mu}_s^\nu(\bar{X}_s^i)ds + \sqrt{2} \left\{ \phi_r^\delta(E_t^i)dB_s^i + \phi_s^\delta(E_t^i)d\tilde{B}_s^i \right\}, \\ dX_s^{i,N} &= -\nabla V(X_s^{i,N})ds - \frac{1}{N} \sum_{j=1}^N \nabla W(X_s^{i,N} - X_s^{j,N})ds \\ &\quad + \sqrt{2} \left\{ \phi_r^\delta(E_t^i) (\text{Id} - 2e_t^i(e_t^i)^\top) dB_s^i + \phi_s^\delta(E_t^i)d\tilde{B}_s^i \right\}. \end{aligned}$$

Here we assume that  $(\bar{X}_0^i, X_0^{i,N})$ ,  $i \in \{1, \dots, N\}$ , are independent random variables with law  $\xi$ ,  $(B_s^i)_{s \geq 0}$  and  $(\tilde{B}_s^i)_{s \geq 0}$ ,  $i \in \{1, \dots, N\}$ , are independent standard Brownian motions in  $\mathbb{R}^d$  that are also independent of the initial conditions, and

$$(17) \quad E_t^i = \bar{X}_t^i - X_t^{i,N}, \quad e_t^i = n(E_t^i),$$

where  $n : \mathbb{R}^d \rightarrow \mathbb{R}^d$  is given for all  $x \in \mathbb{R}^d \setminus \{0\}$  by  $n(x) = x/\|x\|$ , and  $n(0) = 0$ .

Under the fourth moment assumption on  $\nu$  and  $\mu$ , the system of SDEs has a unique strong solution. This can be shown similarly to [11, Theorem 2.6]. By (16) and Lévy's characterization of Brownian motion, the process  $((\bar{X}_t^1, \dots, \bar{X}_t^N)_{t \geq 0}, (X_t^{1,N}, \dots, X_t^{N,N})_{t \geq 0})$  is indeed a realization of a coupling between the system of independent nonlinear diffusions with initial law  $\nu^{\otimes N}$  and the mean-field particle system with initial law  $\mu^{\otimes N}$ . In particular,

$$(18) \quad \bar{\mu}_t^\nu = \mathcal{L}(\bar{X}_t^i) \quad \text{and} \quad \mu_t^{\mu, N} = \mathcal{L}(X_t^{i, N}) \quad \text{for all } t \geq 0 \text{ and } i \in \{1, \dots, N\}.$$

**Lemma 7.** *Assume H1 and H2. Then almost surely, for all  $t \geq 0$  and  $i \in \{1, \dots, N\}$ ,*

$$d \|E_t^i\| = - \langle \nabla V(\bar{X}_t^i) - \nabla V(X_t^{i, N}), e_t^i \rangle dt + A_t^i dt + 2\sqrt{2}\phi_r^\delta(E_t^i)(e_t^i)^\top dB_t^i,$$

where  $(A_t^i)_{t \geq 0}$  is an adapted stochastic process such that

$$(19) \quad A_t^i \leq \left\| \nabla W * \bar{\mu}_t^\nu(\bar{X}_t^i) - \frac{1}{N} \sum_{j=1}^N \nabla W(X_t^{i, N} - X_t^{j, N}) \right\|.$$

*Proof.* For all  $i \in \{1, \dots, N\}$ , using Itô's formula, we have

$$\begin{aligned} d \|E_t^i\|^2 &= -2 \langle \nabla V(\bar{X}_t^i) - \nabla V(X_t^{i, N}), E_t^i \rangle dt \\ &\quad - 2 \left\langle \nabla W * \bar{\mu}_t^\nu(\bar{X}_t^i) - N^{-1} \sum_{j=1}^N \nabla W(X_t^{i, N} - X_t^{j, N}), E_t^i \right\rangle dt \\ &\quad + 4\sqrt{2}\phi_r^\delta(E_t^i) \langle e_t^i, E_t^i \rangle (e_t^i)^\top dB_t^i + 8(\phi_r^\delta(E_t^i))^2 dt. \end{aligned}$$

Now let  $a > 0$ , and consider the function  $\psi_a(r) = (r + a)^{1/2}$  for  $r \geq 0$ . Note that  $\psi_a$  is infinitely continuously differentiable on  $(0, \infty)$ , and for all  $t \geq 0$ ,  $\lim_{a \rightarrow 0} \psi_a(t) = t^{1/2}$ . Therefore, using again Itô's formula, we get

$$(20) \quad \begin{aligned} d\psi_a \left( \|E_t^i\|^2 \right) &= -2\psi_a' \left( \|E_t^i\|^2 \right) \langle \nabla V(\bar{X}_t^i) - \nabla V(X_t^{i, N}), E_t^i \rangle dt \\ &\quad - 2\psi_a' \left( \|E_t^i\|^2 \right) \left\langle \nabla W * \bar{\mu}_t^\nu(\bar{X}_t^i) - N^{-1} \sum_{j=1}^N \nabla W(X_t^{i, N} - X_t^{j, N}), E_t^i \right\rangle dt \end{aligned}$$

$$(21) \quad + (\phi_r^\delta(E_t^i))^2 \left\{ 8\psi_a' \left( \|E_t^i\|^2 \right) + 16 \|E_t^i\|^2 \psi_a'' \left( \|E_t^i\|^2 \right) \right\} dt$$

$$(22) \quad + 4\sqrt{2}\psi_a' \left( \|E_t^i\|^2 \right) \phi_r^\delta(E_t^i) \langle e_t^i, E_t^i \rangle (e_t^i)^\top dB_t^i.$$

Note that  $2r\psi_a'(r^2) = r/\sqrt{r^2 + a} \leq 1/2$  for all  $a, r > 0$ . In particular,

$$\left| 2\psi_a' \left( \|E_t^i\|^2 \right) \langle \nabla V(\bar{X}_t^i) - \nabla V(X_t^{i, N}), E_t^i \rangle \right| \leq \left\| \nabla V(\bar{X}_t^i) - \nabla V(X_t^{i, N}) \right\|.$$

Therefore, by dominated convergence, we see that that almost surely for all  $T \geq 0$ ,

$$\lim_{a \rightarrow 0} \int_0^T 2\psi_a' \left( \|E_t^i\|^2 \right) \langle \nabla V(\bar{X}_t^i) - \nabla V(X_t^{i, N}), E_t^i \rangle dt = \int_0^T \langle \nabla V(\bar{X}_t^i) - \nabla V(X_t^{i, N}), e_t^i \rangle dt.$$

Furthermore, for any  $a > 0$ , the term in (20) is bounded from above by the expression on the right hand side of (19). Moreover, noting that  $\phi_r^\delta(z) = 0$  for  $z \in \mathbb{R}^d$  with  $\|z\| \leq \delta/2$ , and  $8\psi_a'(r^2) + 16r^2\psi_a''(r^2) = 4a/(r^2 + a)^{3/2} \leq 4a/r^3$  for  $a, r > 0$ , we see that for  $T \geq 0$ ,

$$\lim_{a \rightarrow 0} \int_0^T (\phi_r^\delta(E_t^i))^2 \left\{ 8\psi_a' \left( \|E_t^i\|^2 \right) + 16 \|E_t^i\|^2 \psi_a'' \left( \|E_t^i\|^2 \right) \right\} dt = 0.$$



Finally, by [28, Theorem 2.12], we have almost surely

$$\lim_{a \rightarrow 0} \int_0^T 4\sqrt{2}\psi'_a \left( \|E_t^i\|^2 \right) \phi_r^\delta(E_t^i) \langle e_t^i, E_t^i \rangle (e_t^i)^\top dB_t^i = \int_0^T 2\sqrt{2}\phi_r^\delta(E_t^i) (e_t^i)^\top dB_t^i,$$

for any  $T \geq 0$ . This completes the proof of the lemma.  $\square$

**3.2. A moment control.** The following uniform moment bound will be important for the proof of our main result.

**Lemma 8.** *Assume **H1** and **H2**, and suppose that  $\eta \in (0, m_V/2)$  or Assumption **H3** is satisfied. Let  $(\bar{X}_t)_{t \geq 0}$  be a solution of (2) with  $\mathbb{E} [\|\bar{X}_0\|^2] < \infty$ . Then there exists  $C \in (0, \infty)$  depending only on  $d, V, W$  and the second moment of  $\bar{X}_0$  such that*

$$\sup_{t \geq 0} \mathbb{E} [\|\bar{X}_t\|^2] \leq C.$$

*Proof.* The proof is quite standard but we include it here for completeness. First, by Itô's formula, we have

$$(1/2)d\|\bar{X}_t\|^2 = -\langle \bar{X}_t, \nabla V(\bar{X}_t) \rangle dt - \langle \bar{X}_t, \nabla W * \bar{\mu}'_t(\bar{X}_s) \rangle dt + d dt + \sqrt{2}\bar{X}_t^\top dB_t,$$

where  $\nu$  denotes the law of  $\bar{X}_0$ . Let  $(\tilde{X}_t)_{t \geq 0}$  be an independent copy of  $(\bar{X}_t)_{t \geq 0}$ . Then by symmetrization and (15), we get taking the expectation that

$$\frac{d}{dt} \mathbb{E} [\|\bar{X}_t\|^2] = -2\mathbb{E} [\langle \bar{X}_t, \nabla V(\bar{X}_t) \rangle] - \mathbb{E} [\langle \bar{X}_t - \tilde{X}_t, \nabla W(\bar{X}_t - \tilde{X}_t) \rangle] + 2d.$$

Now suppose first that **H1**, **H2** and **H3** are satisfied. Then by (5) and since  $\nabla W(0) = 0$ ,

$$\frac{d}{dt} \mathbb{E} [\|\bar{X}_t\|^2] \leq 2M_V + M_W + 2d - 2m_V \mathbb{E} [\|\bar{X}_t\|^2] + 2\|\nabla V(0)\| \mathbb{E} [\|\bar{X}_t\|^2]^{1/2}.$$

Hence Gronwall's lemma concludes the proof.

Alternatively, assume only that **H1** and **H2** are satisfied. Since  $f(r) \leq r$ , we obtain

$$\frac{d}{dt} \mathbb{E} [\|\bar{X}_t\|^2] \leq 2M_V + 2d + 2(\eta - m_V) \mathbb{E} [\|\bar{X}_t\|^2] + 2\|\nabla V(0)\| \mathbb{E} [\|\bar{X}_t\|^2]^{1/2}.$$

Hence for  $\eta < m_V/2$ , we can still apply Gronwall's lemma to conclude the proof.  $\square$

**3.3. Proof of main results.** We can now prove our main result.

*Proof of Theorem 2.* We fix  $\delta > 0$  and a coupling  $\xi \in \Pi(\nu, \mu)$ , and we consider the reflection coupling with initial law  $\xi^{\otimes N}$  between the system of nonlinear processes and the mean-field particle system as introduced in Subsection 3.1. Since  $f$  is continuously differentiable and  $f'$  is absolutely continuous, by Lemma 7 and the Itô-Tanaka formula, we obtain

$$df(\|E_t^i\|) = (C_t^i + A_t^i) f'(\|E_t^i\|) dt + 4f''(\|E_t^i\|) (\phi_r^\delta(E_t^i))^2 dt + \sqrt{8}\phi_r^\delta(E_t^i) (e_t^i)^\top dB_t^i,$$

where  $C_t^i = -\langle \nabla V(\bar{X}_t^i) - \nabla V(X_t^{i,N}), e_t^i \rangle$ . Define  $\omega : \mathbb{R}_+ \rightarrow \mathbb{R}_+$  by

$$(23) \quad \omega(r) = \sup_{s \in [0, r]} s \kappa(s)^-.$$

By definition of  $\kappa$  given by (4), and by (13),

$$(24) \quad \begin{aligned} C_t^i f'(\|E_t^i\|) + 4f''(\|E_t^i\|) (\phi_r^\delta(E_t^i))^2 \\ \leq -\|E_t^i\| \kappa(\|E_t^i\|) f'(\|E_t^i\|) + 4f''(\|E_t^i\|) (\phi_r^\delta(E_t^i))^2 \end{aligned}$$

$$(25) \quad \begin{aligned} \leq -2cf(\|E_t^i\|) (\phi_r^\delta(E_t^i))^2 + \omega(\delta) \\ \leq -2cf(\|E_t^i\|) + \omega(\delta) + 2cf(\delta). \end{aligned}$$

Moreover, by (19), we can estimate  $A_t^i \leq N^{-1} \sum_{j=1}^N \Xi_t^{i,j} + \Upsilon_t^i$ , where

$$\begin{aligned} \Xi_t^{i,j} &= \left\| \nabla W(\bar{X}_t^{i,N} - \bar{X}_t^{j,N}) - \nabla W(X_t^{i,N} - X_t^{j,N}) \right\|, \\ \Upsilon_t^i &= \left\| \nabla W * \bar{\mu}_t^\nu(\bar{X}_t^i) - N^{-1} \sum_{j=1}^N \nabla W(\bar{X}_t^i - \bar{X}_t^j) \right\|. \end{aligned}$$

Therefore, and since  $(\int_0^t \phi_r^\delta(E_s^i) (e_s^i)^\top dB_s^i)_{t \geq 0}$  is a martingale and  $f' \leq 1$ , we obtain

$$(26) \quad \begin{aligned} \frac{1}{N} \sum_{i=1}^N \frac{d}{dt} \mathbb{E}[f(\|E_t^i\|)] &\leq -\frac{2c}{N} \sum_{i=1}^N \mathbb{E}[f(\|E_t^i\|)] + \frac{1}{N} \sum_{i=1}^N \mathbb{E} \left[ \frac{1}{N} \sum_{j=1}^N \Xi_t^{i,j} + \Upsilon_t^i \right] \\ &\quad + \omega(\delta) + 2cf(\delta) \end{aligned}$$

for a.e.  $t \geq 0$ . By Assumption **H2**, for all  $i, j \in \{1, \dots, N\}$  and  $t \geq 0$ ,

$$(27) \quad \Xi_t^{i,j} \leq \eta f(\|E_t^i\| + \|E_t^j\|) \leq \eta \left( f(\|E_t^i\|) + f(\|E_t^j\|) \right).$$

Moreover, in order to control  $\Upsilon_t^i$ , we remark that given  $\bar{X}_t^i$ , the random variables  $\bar{X}_t^j$ ,  $j \neq i$ , are i.i.d. with law  $\bar{\mu}_t^\nu$ . In particular,

$$\mathbb{E} \left[ \nabla W(\bar{X}_t^i - \bar{X}_t^j) | \bar{X}_t^i \right] = \nabla W * \bar{\mu}_t^\nu(\bar{X}_t^i).$$

Since by **H2**,  $\nabla W(0) = 0$  and  $\nabla W$  is Lipschitz with constant  $\eta$ , we obtain

$$\begin{aligned} \mathbb{E} \left[ \left\| \nabla W * \bar{\mu}_t^\nu(\bar{X}_t^i) - \frac{1}{N-1} \sum_{j=1}^N \nabla W(\bar{X}_t^i - \bar{X}_t^j) \right\|^2 \middle| \bar{X}_t^i \right] \\ = \frac{1}{N-1} \text{Var}_{\bar{\mu}_t^\nu}(\nabla W(\bar{X}_t^i - \cdot)) \leq \frac{\eta^2}{N-1} \int \|x\|^2 \bar{\mu}_t^\nu(dx). \end{aligned}$$

Hence, by the Cauchy-Schwarz inequality and (14),

$$\begin{aligned} \mathbb{E}[\Upsilon_t^i] &\leq \mathbb{E} \left[ \left\| \nabla W * \bar{\mu}_t^\nu(\bar{X}_t^i) - \frac{1}{N-1} \sum_{j=1}^N \nabla W(\bar{X}_t^i - \bar{X}_t^j) \right\| \right] \\ &\quad + \left( \frac{1}{N-1} - \frac{1}{N} \right) \sum_{j=1}^N \mathbb{E} \left[ \eta \left\| \bar{X}_t^i - \bar{X}_t^j \right\| \right] \\ &\leq \eta \left( \frac{1}{\sqrt{N-1}} + \frac{\sqrt{2}}{N} \right) \left( \int \|x\|^2 \bar{\mu}_t^\nu(dx) \right)^{1/2}. \end{aligned}$$

By Lemma 8, we conclude that there is an explicit finite constant  $C$  such that for  $N \geq 2$ ,

$$(28) \quad \sup_{t \geq 0} \mathbb{E}[\Upsilon_t^i] \leq C\eta N^{-1/2}, \quad i = 1, \dots, N.$$

Now, combining (26), (27) and (28), we finally obtain

$$\frac{d}{dt} \frac{1}{N} \sum_{i=1}^N \mathbb{E} [f (\|E_t^i\|)] \leq -2 \frac{c-\eta}{N} \sum_{i=1}^N \mathbb{E} [f (\|E_t^i\|)] + \omega(\delta) + 2cf(\delta) + \frac{C\eta}{\sqrt{N}}$$

for a.e.  $t \geq 0$ . Assuming  $\eta < c$ , we can conclude that

$$\begin{aligned} \frac{1}{N} \sum_{i=1}^N \mathbb{E} [f (\|E_t^i\|)] &\leq e^{-2(c-\eta)t} \frac{1}{N} \sum_{i=1}^N \mathbb{E} [f (\|E_0^i\|)] \\ &\quad + (2(c-\eta))^{-1} \left( \omega(\delta) + 2cf(\delta) + C\eta N^{-1/2} \right). \end{aligned}$$

Noting that  $\mathbb{E} [f (\|E_0^i\|)] \leq \int \|x - y\| d\xi(x, y)$  for all  $i \in \{1, \dots, d\}$ , we obtain

$$\begin{aligned} \mathcal{W}_{\ell^1(f)} \left( \mathcal{L}(X_t^{1,N}, \dots, X_t^{N,N}), (\bar{\mu}_t^\nu)^N \right) \\ \leq e^{-2(c-\eta)t} \int \|x - y\| d\xi(x, y) + (2(c-\eta))^{-1} \left( \omega(\delta) + 2cf(\delta) + C\eta N^{-1/2} \right). \end{aligned}$$

By (23) and **H 1**,  $\lim_{\delta \rightarrow 0^+} \omega(\delta) = 0$ . Hence taking the limit as  $\delta$  goes to 0, and the infimum over all couplings  $\xi \in \Pi(\nu, \mu)$  concludes the proof of the second inequality in Theorem 2. The first inequality follows similarly, noting that by (18), for all  $i \in \{1, \dots, d\}$  and  $t \geq 0$ ,

$$\mathcal{W}_f(\bar{\mu}_t^\nu, \mu_t^{\mu, N}) \leq \mathbb{E} \left[ f \left( \left\| \bar{X}_t^i - X_t^{i,N} \right\| \right) \right] = \mathbb{E} [f (\|E_t^i\|)] .$$

□

**Acknowledgment.** This work was initiated through a Procope project which is greatly acknowledged. The work was partially supported by ANR-17-CE40-0030, by the Hausdorff Center for Mathematics, and by the grant 346300 for IMPAN from the Simons Foundation and the matching 2015-2019 Polish MNiSW fund.

## REFERENCES

- [1] D. Benedetto, E. Caglioti, J. A. Carrillo, and M. Pulvirenti. A non Maxwellian steady distribution for one-dimensional granular media. *J. Stat. Physics*, 91(5/6):979–990, 1998.
- [2] F. Bolley, I. Gentil, and A. Guillin. Uniform convergence to equilibrium for granular media. *Arch. Ration. Mech. Anal.*, 208(2):429–445, 2013.
- [3] F. Bolley, A. Guillin, and F. Malrieu. Trend to equilibrium and particle approximation for a weakly selfconsistent Vlasov-Fokker-Planck equation. *Math. Mod. Num. Anal.*, 44 (5):867–884, 2010.
- [4] F. Bolley, A. Guillin, and C. Villani. Quantitative concentration inequalities for empirical measures on non compact spaces. *Prob. Theor. Rel. Fields*, 137(3-4):541–593, 2007.
- [5] Francois Bolley, Ivan Gentil, and Arnaud Guillin. Convergence to equilibrium in Wasserstein distance for Fokker-Planck equations. *J. Funct. Anal.*, 263(8):2430–2457, 2012.
- [6] O. A. Butkovsky. On ergodic properties of nonlinear Markov chains and stochastic McKean-Vlasov equations. *Theory Probab. Appl.*, 58(4):661–674, 2014.
- [7] J. A. Carrillo, R. J. McCann, and C. Villani. Kinetic equilibration rates for granular media and related equations: entropy dissipation and mass transportation estimates. *Rev. Mat. Iberoam.*, 19(3):971–1018, 2003.
- [8] J. A. Carrillo, R. J. McCann, and C. Villani. Contractions in the 2-Wasserstein length space and thermalization of granular media. *Arch. Rational Mech. Anal.*, 179:217–263, 2006.

- [9] J. A. Carrillo and G. Toscani. Wasserstein metric and large-time asymptotics of nonlinear diffusion equations. In *New trends in math. physics*. World Sci., Singapore, 2005.
- [10] J. A. Carrillo and G. Toscani. Contractive probability metrics and asymptotic behavior of dissipative kinetic equations. *Riv. Mat. Univ. Parma*, 7(6):75–198, 2007.
- [11] P. Cattiaux, A. Guillin, and F. Malrieu. Probabilistic approach for granular media equations in the non uniformly convex case. *Prob. Theor. Rel. Fields*, 140(1-2):19–40, 2008.
- [12] P. Del Moral and J. Tugaut. Uniform propagation of chaos and creation of chaos for a class of nonlinear diffusions. 2015.
- [13] A. Eberle. Reflection couplings and contraction rates for diffusions. *Probability theory and related fields*, 166(3-4):851–886, 2016.
- [14] A. Eberle, A. Guillin, and R. Zimmer. Quantitative Harris type theorems for diffusions and McKean-Vlasov processes. To appear in *Trans. Am. Math. Soc.*, 2018.
- [15] Andreas Eberle. Reflection coupling and Wasserstein contractivity without convexity. *C. R. Math. Acad. Sci. Paris*, 349(19-20):1101–1104, 2011.
- [16] Tadahisa Funaki. A certain class of diffusion processes associated with nonlinear parabolic equations. *Z. Wahrsch. Verw. Gebiete*, 67(3):331–348, 1984.
- [17] W. Hammersley, D. Siska, and L. Szpruch. McKean-Vlasov SDEs under Measure Dependent Lyapunov Conditions. <https://arxiv.org/abs/1802.03974>, 2018.
- [18] Maxime Hauray and Stéphane Mischler. On Kac’s chaos and related problems. *J. Funct. Anal.*, 266(10):6055–6157, 2014.
- [19] S. Herrmann and J. Tugaut. Non-uniqueness of stationary measures for self-stabilizing processes. *Stochastic Process. Appl.*, 120(7):1215–1246, 2010.
- [20] M. Kac. Foundations of kinetic theory. In *Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability, 1954–1955, vol. III*, pages 171–197. University of California Press, Berkeley and Los Angeles, 1956.
- [21] Torgny Lindvall and L. C. G. Rogers. Coupling of multidimensional diffusions by reflection. *Ann. Probab.*, 14(3):860–872, 1986.
- [22] F. Malrieu. Convergence to equilibrium for granular media equations and their Euler schemes. *Ann. Appl. Probab.*, 13(2):540–560, 2003.
- [23] H. P. McKean, Jr. A class of Markov processes associated with nonlinear parabolic equations. *Proc. Nat. Acad. Sci. U.S.A.*, 56:1907–1911, 1966.
- [24] Sylvie Méléard. Asymptotic behaviour of some interacting particle systems; McKean-Vlasov and Boltzmann models. In *Probabilistic models for nonlinear partial differential equations (Montecatini Terme, 1995)*, volume 1627 of *Lecture Notes in Math.*, pages 42–95. Springer, Berlin, 1996.
- [25] Stéphane Mischler and Clément Mouhot. Kac’s program in kinetic theory. *Invent. Math.*, 193(1):1–147, 2013.
- [26] Stéphane Mischler, Clément Mouhot, and Bernt Wennberg. A new approach to quantitative propagation of chaos for drift, diffusion and jump processes. *Probab. Theory Related Fields*, 161(1-2):1–59, 2015.
- [27] S. Mishura and A.Y. Veretennikov. Existence and uniqueness theorems for solutions of mckean-vlasov stochastic equations. <https://arxiv.org/abs/1603.02212v5>, 2016.
- [28] D. Revuz and M. Yor. *Continuous martingales and Brownian motion*, volume 293 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, third edition, 1999.
- [29] Alain-Sol Sznitman. Topics in propagation of chaos. In *École d’Été de Probabilités de Saint-Flour XIX—1989*, volume 1464 of *Lecture Notes in Math.*, pages 165–251. Springer, Berlin, 1991.
- [30] Julian Tugaut. Convergence to the equilibria for self-stabilizing processes in double-well landscape. *Ann. Probab.*, 41(3A):1427–1460, 05 2013.
- [31] Julian Tugaut. Self-stabilizing processes in multi-wells landscape in  $\mathbb{R}^d$ -invariant probabilities. *J. Theoret. Probab.*, 27(1):57–79, 2014.
- [32] C. Villani. *Topics in optimal transportation*, volume 58 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2003.

- [33] C. Villani. *Optimal transport, Old and new*, volume 338 of *Grund. Math. Wiss.* Springer, Berlin, 2009.
- [34] Cédric Villani. Mathematics of granular materials. *J. Stat. Phys.*, 124(2-4):781–822, 2006.

ENS PARIS-SACLAY, CENTRE DE MATHÉMATIQUES ET LEURS APPLICATIONS, 61 AVENUE DU PRÉSIDENT  
WILSON, 94235 CACHAN, FRANCE  
*E-mail address:* `alain.durmus@cmla.ens-cachan.fr`

UNIVERSITÄT BONN, INSTITUT FÜR ANGEWANDTE MATHEMATIK, ENDENICHER ALLEE 60, 53115  
BONN, GERMANY  
*E-mail address:* `eberle@uni-bonn.de`, `Raphael@infoZimmer.de`  
*URL:* `http://www.uni-bonn.de/~eberle`

LABORATOIRE DE MATHÉMATIQUES BLAISE PASCAL, CNRS - UMR 6620, UNIVERSITÉ CLERMONT-  
AUVERGNE, AVENUE DES LANDAIS,, 63177 AUBIERE CEDEX, FRANCE  
*E-mail address:* `guillin@math.univ-bpclermont.fr`  
*URL:* `http://math.univ-bpclermont.fr/~guillin`