Trend to equilibrium and particle approximation for a weakly selfconsistent Vlasov-Fokker-Planck equation

Francois Bolley, Arnaud Guillin, Florent Malrieu

Trend to equilibrium and particle approximation for a weakly selfconsistent Vlasov-Fokker-Planck equation

François Bolley, Arnaud Guillin, Florent Malrieu

Preprint – June 7, 2009

Abstract

We consider a Vlasov-Fokker-Planck equation governing the evolution of the density of interacting and diffusive matter in the space of positions and velocities. We use a probabilistic interpretation to obtain convergence towards equilibrium in Wasserstein distance with an explicit exponential rate. We also prove a propagation of chaos property for an associated particle system, and give rates on the approximation of the solution by the particle system. Finally, a transportation inequality for the distribution of the particle system leads to quantitative deviation bounds on the approximation of the equilibrium solution of the equation by an empirical mean of the particles at given time.

Introduction and main results

We are interested in the long time behaviour and in a particle approximation of a distribution $f_t(x,v)$ in the space of positions $x \in \mathbb{R}^d$ and velocities $v \in \mathbb{R}^d$ (with $d \geqslant 1$) evolving according to the Vlasov-Fokker-Planck equation

$$\frac{\partial f_t}{\partial t} + v \cdot \nabla_x f_t - C *_x \rho[f_t](x) \cdot \nabla_v f_t = \Delta_v f_t + \nabla_v \cdot ((A(v) + B(x)) f_t), \qquad t > 0,\, x, v \in \mathbb{R}^d \tag{1}$$

where

$$\rho[f_t](x) = \int_{\mathbb{R}^d} f_t(x,v)\, dv$$

is the macroscopic density in the space of positions $x \in \mathbb{R}^d$ (or the space marginal of f_t). Here $a \cdot b$ denotes the scalar product of two vectors a and b in $\mathbb{R}^d$ and $*_x$ stands for the convolution with respect to $x \in \mathbb{R}^d$:

$$C *_x \rho[f_t](x) = \int_{\mathbb{R}^d} C(x-y)\, \rho[f_t](y)\, dy = \int_{\mathbb{R}^{2d}} C(x-y)\, f_t(y,v)\, dy\, dv.$$

Moreover ∇_x stands for the gradient with respect to the position variable $x \in \mathbb{R}^d$ whereas ∇_v, $\nabla_v \cdot$ and Δ_v respectively stand for the gradient, divergence and Laplace operators with respect to the velocity variable $v \in \mathbb{R}^d$.

The $A(v)$ term models the friction, the $B(x)$ term models an exterior confinement and the $C(x-y)$ term in the convolution models the interaction between positions x and y in the underlying physical system. For that reason we assume that C is an odd map on $\mathbb{R}^d$. This equation is used in the modelling of the distribution $f_t(x,v)$ of diffusive, confined and interacting stellar or charged matter when C respectively derives from the Newton and Coulomb potential (see [10] for instance). It has the following natural probabilistic interpretation: if f_0 is a density function, the solution f_t of (1) is the density of the law at time t of the $\mathbb{R}^{2d}$-valued process $(X_t, V_t)_{t \geqslant 0}$ evolving according to the mean field stochastic differential equation (diffusive Newton's equations)

$$\begin{cases} dX_t = V_t\, dt \\ dV_t = -A(V_t)\, dt - B(X_t)\, dt - C *_x \rho[f_t](X_t)\, dt + \sqrt{2}\, dW_t. \end{cases} \tag{2}$$

Here $(W_t)_{t\geqslant 0}$ is a Brownian motion in the velocity space $\mathbb{R}^d$ and f_t is the law of (X_t, V_t) in $\mathbb{R}^{2d}$, so that $\rho[f_t]$ is the law of X_t in $\mathbb{R}^d$.

Space homogeneous models of diffusive and interacting granular media (see [4]) have been studied by P. Cattiaux and the last two authors in particular [14], [19, 20], by means of a stochastic interpretation analogous to (2) and a particle approximation analogous to (4) below. They were interpreted as gradient flows in the space of probability measures by J. A. Carrillo, R. J. McCann and C. Villani [11, 12] (see also [26]), both approaches leading to explicit exponential (or algebraic for non uniformly convex potentials) *rates of convergence to equilibrium*. Also possibly time-uniform propagation of chaos was proven for the associated particle system.

Obtaining rates of convergence to equilibrium for (1) is much more complex, as the equation simultaneously presents hamiltonian and gradient flows aspects. Much attention has recently been called to the linear noninteracting case of (1), when $C = 0$, also known as the kinetic Fokker-Planck equation. First of all a probabilistic approach based on Lyapunov functionals, and thus easy to check conditions, lead D. Talay [24], L. Wu [27] or D. Bakry, P. Cattiaux and the second author [2] to exponential or subexponential convergence to equilibrium in total variation distance. The case when $A(v) = v$ and $B(x) = \nabla\Psi(x)$, and when the equilibrium solution is explicitely given by $f_\infty(x, v) = e^{-\Psi(x)-|v|^2/2}$ is studied in [16], [17] and [25, Chapter 7]: hypocoercivity analytic techniques are developed which, applied to this situation, give sufficient conditions, in terms of Poincaré or logarithmic Sobolev inequalities for the measure $e^{-\Psi}$, to L^2 or entropic convergence with an explicit exponential rate. We also refer to [23] for the evolution of two species, modelled by two coupled Vlasov-Fokker-Planck equations.

C. Villani's approach extends to the selfconsistent situation when C derives from a nonzero potential U (see [25, Chapter 17]): replacing the confinement force $B(x)$ by a periodic boundary condition, and for small and smooth potential U, he obtains an explicit exponential rate of convergence of all solutions toward the unique normalized equilibrium solution $e^{-|v|^2/2}$.

In this work we consider the case when the equation is set on the whole $\mathbb{R}^d$, with quadratic-like friction $A(v)$ and confinement $B(x)$ forces, and small Lipschitz interaction $C(x)$: in the whole paper we make the following

Assumption. We say that Assumption $(\mathcal{A})$ is fullfilled if there exist nonnegative constants $\alpha, \alpha', \beta, \gamma$ and δ such that

$$|A(v) - A(w)| \leqslant \alpha|v - w|, \quad (v - w)\cdot(A(v) - A(w)) \geqslant \alpha'|v - w|^2,$$

$$B(x) = \beta\, x + D(x) \quad \text{where } |D(x) - D(y)| \leqslant \delta|x - y|$$

and

$$|C(x) - C(y)| \leqslant \gamma|x - y|$$

for all x, y, v, w in $\mathbb{R}^d$.

Convergence of solutions will be measured in terms of Wasserstein distances: let $\mathcal{P}_2$ be the space of Borel probability measures μ on $\mathbb{R}^{2d}$ with finite second moment, that is, such that the integral $\displaystyle\int_{\mathbb{R}^{2d}} (|x|^2+|v|^2)\, d\mu(x, v)$ be finite. The space $\mathcal{P}_2$ is equipped with the (Monge-Kantorovich) Wasserstein distance d of order 2 defined by

$$d(\mu, \nu)^2 = \inf_{(X,V),\,(Y,W)} \mathbb{E}\Big(|X - Y|^2 + |V - W|^2\Big)$$

where the infimum runs over all the couples (X, V) and (Y, W) of random variables on $\mathbb{R}^{2d}$ with respective laws μ and ν. Convergence in this metric is equivalent to narrow convergence plus convergence of the second moment (see [26, Chapter 6] for instance).

The coefficients A, B and C being Lipschitz, existence and uniqueness for Equation (2) with square-integrable initial data are ensured by [21]. It follows that, for all initial data f_0 in $\mathcal{P}_2$, Equation (1) admits a unique measure solution in $\mathcal{P}_2$, that is, continuous on $[0, +\infty[$ with values in $\mathcal{P}_2$.

Assumption $(\mathcal{A})$ is made in the whole paper. Under an additional assumption on the smallness of γ and δ, we shall prove a quantitative exponential convergence of all solutions to a unique equilibrium:

Theorem 1. *Under Assumption $(\mathcal{A})$, for all positive α, α' and β there exists a positive constant c such that, if $0 \leqslant \gamma, \delta < c$, then there exist positive constants C and C' such that*

$$d(f_t, \bar{f}_t) \leqslant C' \, e^{-Ct} \, d(f_0, \bar{f}_0), \qquad t \geqslant 0 \tag{3}$$

for all solutions $(f_t)_{t \geqslant 0}$ and $(\bar{f}_t)_{t \geqslant 0}$ to (1) with respective initial data f_0 and $\bar{f}_0$ in $\mathcal{P}_2$.

Moreover (1) admits a unique stationnary solution μ_∞ and all solutions $(f_t)_{t \geqslant 0}$ converge towards it, with

$$d(f_t, \mu_\infty) \leqslant C' e^{-Ct} d(f_0, \mu_\infty), \qquad t \geqslant 0.$$

For instance, for $\alpha = \alpha' = \beta = 1$, the general proof below shows that the nonnegative γ and δ with $\gamma + \delta < 0,26$ are admissible. In the linear free case when $\gamma = \delta = 0$, the convergence rate is given by $C = 1/3$, and for instance for γ and δ with $\gamma + \delta = 0,1$ we obtain $C \sim 0,27$.

Compared to Villani's results, convergence is here proven in the (weak) Wasserstein distance, not in L^1 norm, or relative entropy as in the noninteracting case - the latter being a stronger convergence since, in this specific situation, the equilibrium measure $e^{-\Psi(x) - |v|^2/2}$ satisfies a logarithmic Sobolev inequality, hence a transportation inequality. We refer to [1], [18] or [26, Chapter 22] for this and forthcoming notions.

However our result holds in the noncompact case with small Lipschitz interaction, and can be seen as a first attempt to deal with more general case. Moreover it shows existence and uniqueness of the equilibrium measure, and in particular does not use its explicit expression (which is unknown in our broader situation). It is also not only a result on the convergence to equilibrium, but also a stability result of all solutions. Let us finally note that it is based on the natural stochastic interpretation (2) and a simple coupling argument, and does not need any hypoelliptic regularity property of the solutions.

The *particle approximation* of solutions to (1) consists in the introduction of a large number N of $\mathbb{R}^{2d}$-valued processes $(X_t^{i,N}, V_t^{i,N})_{t \geqslant 0}$ with $1 \leqslant i \leqslant N$, no more evolving according to the force field $C *_x \rho[f_t]$ generated by the distribution f_t as in (2), but by the empirical measure $\hat{\mu}_t^N = \dfrac{1}{N} \displaystyle\sum_{i=1}^N \delta_{(X_t^{i,N}, V_t^{i,N})}$ of the system: if $(W_\cdot^i)_{i \geqslant 1}$ with $i \geqslant 1$ are independent standard Brownian motions on $\mathbb{R}^d$ and (X_0^i, V_0^i) with $i \geqslant 1$ are independent random vectors on $\mathbb{R}^{2d}$ with law f_0 in $\mathcal{P}_2$ and indepedent of $(W_\cdot^i)_{i \geqslant 1}$, we let $(X_t^{(N)}, V_t^{(N)})_{t \geqslant 0} = (X_t^{1,N}, \dots, X_t^{N,N}, V_t^{1,N}, \dots, V_t^{N,N})_{t \geqslant 0}$ be the solution of the following stochastic differential equation in $(\mathbb{R}^{2d})^N$:

$$\begin{cases} dX_t^{i,N} = V_t^{i,N} \, dt \\ dV_t^{i,N} = -A(V_t^{i,N}) \, dt - B(X_t^{i,N}) \, dt - \dfrac{1}{N} \displaystyle\sum_{j=1}^N C(X_t^{i,N} - X_t^{j,N}) \, dt + \sqrt{2} \, dW_t^i, \qquad 1 \leqslant i \leqslant N \\ (X_0^{i,N}, V_0^{i,N}) = (X_0^i, V_0^i). \end{cases} \tag{4}$$

The mean field force $C *_x \rho[f_t]$ in (2) is replaced by the pairwise actions $\dfrac{1}{N} C(X_t^{i,N} - X_t^{j,N})$ of particle j on particle i. Since this interaction is of order $1/N$, it may be reasonable that two of these interacting particles (or a fixed number k of them) become less and less correlated as N gets large.

In order to state this *propagation of chaos* property we let, for each $i \geqslant 1$, $(\bar{X}_t^i, \bar{V}_t^i)_{t \geqslant 0}$ be the solution of the kinetic McKean-Vlasov type equation on $\mathbb{R}^{2d}$

$$\begin{cases} d\bar{X}_t^i = \bar{V}_t^i \, dt \\ d\bar{V}_t^i = -A(\bar{V}_t^i) \, dt - B(\bar{X}_t^i) \, dt - C *_x \rho[\nu_t](\bar{X}_t^i) \, dt + \sqrt{2} \, dW_t^i, \\ (\bar{X}_0^i, \bar{V}_0^i) = (X_0^i, V_0^i). \end{cases} \tag{5}$$

where ν_t is the distribution of $(\bar{X}_t^i, \bar{V}_t^i)$. The processes $(\bar{X}_t^i, \bar{V}_t^i)_{t \geqslant 0}$ with $i \geqslant 1$ are independent since the initial conditions and driving Brownian motions are independent. Moreover they are identically distributed and their common law at time t evolves according to (1), so is the solution f_t of (1) with initial datum f_0. In this notation, and as N gets large, the N processes $(X_t^{i,N}, V_t^{i,N})_{t \geqslant 0}$ look more and more like the N independent processes $(\bar{X}_t^i, \bar{V}_t^i)_{t \geqslant 0}$:

Theorem 2 (Time-uniform propagation of chaos)**.** *Let (X_0^i, V_0^i) for $1 \leqslant i \leqslant N$ be N independent $\mathbb{R}^{2d}$-valued random variables with law f_0 in $\mathcal{P}_2(\mathbb{R}^{2d})$. Let also $(X_t^{i,N}, V_t^{i,N})_{t \geqslant 0, 1 \leqslant i \leqslant N}$ be the solution to* (2) *and $(\bar{X}_t^i, \bar{V}_t^i)_{t \geqslant 0}$ the solution to* (5) *with initial datum (X_0^i, V_0^i) for $1 \leqslant i \leqslant N$. Under Assumption $(\mathcal{A})$, for all positive α, α' and β there exists a positive constant c such that, if $0 \leqslant \gamma, \delta < c$, then there exists a positive constant C, independent of N, such that for $i = 1, \ldots, N$*

$$\sup_{t \geqslant 0} \mathbb{E} \left(\left| X_t^{i,N} - \bar{X}_t^i \right|^2 + \left| V_t^{i,N} - \bar{V}_t^i \right|^2 \right) \leqslant \frac{C}{N} \cdot$$

Here the constant C depends only on the coefficients of the equation and the second moment of f_0.

Remark 3. *In particular the law $f_t^{(1,N)}$ at time t of any $(X_t^{i,N}, V_t^{i,N})$ (by symmetry) converges to f_t as N goes to infinity, according to*

$$d(f_t^{(1,N)}, f_t)^2 \leqslant \mathbb{E} \left(\left| X_t^{i,N} - \bar{X}_t^i \right|^2 + \left| V_t^{i,N} - \bar{V}_t^i \right|^2 \right) \leqslant \frac{C}{N} \cdot$$

Propagation of chaos at the level of the *trajectories*, and not only of the time-marginals, is estimated in [8] and [21] for a broad class of equations, but with non time-uniform constants.

We finally turn to the approximation of the equilibrium solution of the Vlasov-Fokker-Planck equation (as given by Theorem 1) by the particle system at a given time T.

Since all solutions $(f_t)_{t \geqslant 0}$ to (1) with initial data in $\mathcal{P}_2$ converge to the equilibrium solution μ_∞, we let (x_0, v_0) in $\mathbb{R}^{2d}$ be given and we consider the Dirac mass $\delta_{(x_0,v_0)}$ at (x_0, v_0) as the initial datum f_0. We shall give precise bounds on the approximation of μ_∞ by the empirical measure of the particles $(X_t^{i,N}, V_t^{i,N})$ for $1 \leqslant i \leqslant N$, all of them initially at (x_0, v_0).

In the space homogeneous case of the granular media equation, this was performed by the third author [19, 20] by proving a logarithmic Sobolev inequality for the joint law $f_t^{(N)}$ of the N particles at time t. In turn this inequality was proved by a Bakry-Emery curvature criterion (see [3]). The argument does not work here as the particle system has $-\infty$ curvature, and we shall only prove a (Talagrand) T_2 transportation inequality for the joint law of the particles.

Remark 4. *At this stage we have to point out that, for instance when the force fields A, B and C are gradient of potentials, the invariant measure of the particle system, that is, the large time limit of the joint law of the N particles, is explicit and satisfies a logarithmic Sobolev inequality with carré du champ $|\nabla_x f|^2 + |\nabla_v f|^2$; however it does not satisfy a logarithmic Sobolev inequality with carré du champ $|\nabla_v f|^2$ (initiated by our dynamics), which would at once lead to exponential entropic convergence to equilibrium for the particle system.*

Let us recall that a probability measure μ on $\mathbb{R}^{2d}$ is said to satisfies a T_2 transportation inequality if there exists a constant D such that

$$d(\mu,\nu)^2 \leq D\ H(\nu|\mu)$$

for all probability measure ν; here

$$H(\nu|\mu) = \int \log\left(\frac{d\nu}{d\mu}\right) d\nu$$

if $\nu \ll \mu$ and $+\infty$ otherwise is the relative entropy of ν with respect to μ.

Theorem 5. *Under the assumptions of Theorem 1, for all positive α, α' and β there exists a positive constant c such that if $0 \leqslant \gamma, \delta < c$, then the joint law of the N particles $(X_T^{(i,N)}, V_T^{(i,N)})$ at given time T, all with deterministic starting points $(x_0, v_0) \in \mathbb{R}^{2d}$, satisfies a T_2 inequality with a constant D independent of the number N of particles, of time T and of the point (x_0, v_0).*
It follows that there exists a constant D' such that

$$\mathbb{P}\left(\frac{1}{N}\sum_{i=1}^{N} h(X_T^{i,N}, V_T^{i,N}) - \int_{\mathbb{R}^{2d}} h\, d\mu_\infty(h) \geq r + D'\left(\frac{1}{\sqrt{N}} + e^{-CT}\right)\right) \leq \exp\left(-\frac{Nr^2}{2D}\right)$$

for all $N, T, r \geqslant 0$ and all 1-Lipschitz observables h on $\mathbb{R}^{2d}$.

Here the constant C has been obtained in Theorem 1 and the constant D' depends only on the point (x_0, v_0) and the coefficients of the equation.

Remark 6. *Such single observable deviation inequalities were obtained in [19] for the space homogeneous granular media equation; they were upgraded in [9] to the very level of the measures, and to the level of the density of the equilibrium solution. The authors believe that such estimates can also be obtained in the present case.*

Remark 7. *Let us also point out that if we do not suppose a confinement/convexity assumption as in $(\mathcal{A})$ but only Lipschitz regularity on the drift fields A, B and C, then Theorems 1, 2 and 5 still hold but with constants growing exponentially fast with time T.*

Sections 1, 2 and 3 are respectively devoted to the proofs of Theorems 1, 2 and 5.

1 Long time behaviour for the Vlasov-Fokker-Planck equation

This section is devoted to the proof of Theorem 1, which is based on the stochastic interpretation (2) of (1) and a coupling argument. It uses the idea, also present in [24] and [25], of perturbing the Euclidean metric on $\mathbb{R}^{2d}$ in such a way that (2) is dissipative for this metric.

If Q is a positive quadratic form on $\mathbb{R}^{2d}$ and μ and ν are two probability measures in $\mathcal{P}_2$ we let

$$d_Q(\mu,\nu)^2 = \inf_{(X,V),\,(Y,W)} \mathbb{E}(Q((X,V)-(Y,W)))$$

where again the infimum runs over all the couples (X,V) and (Y,W) of random variables on $\mathbb{R}^{2d}$ with respective laws μ and ν; so that $d_Q = d$ if Q is the squared Euclidean norm on $\mathbb{R}^{2d}$. The key step in the proof is the following

Proposition 8. *Under the assumptions of Theorem 1, there exist a positive constant C and a positive quadratic form Q on $\mathbb{R}^{2d}$ such that*

$$d_Q(f_t, \bar{f}_t) \leqslant e^{-Ct}\, d_Q(f_0, \bar{f}_0), \qquad t \geqslant 0$$

for all solutions $(f_t)_{t\geqslant 0}$ and $(\bar{f}_t)_{t\geqslant 0}$ to (1) with respective initial data f_0 and $\bar{f}_0$ in $\mathcal{P}_2$.

Proof of Proposition 8. Let $(f_t)_{t\geqslant 0}$ and $(\bar f_t)_{t\geqslant 0}$ be two solutions to (1) with initial data f_0 and $\bar f_0$ in $\mathcal{P}_2$. Let also (X_0, V_0) and $(\bar X_0, \bar V_0)$ with respectively law f_0 and $\bar f_0$, evolving into (X_t, V_t) and $(\bar X_t, \bar V_t)$ according to (2), both with the same Brownian motion $(W_t)_{t\geqslant 0}$ in $\mathbb{R}^d$. Then, by difference, $(x_t, v_t) = (X_t - \bar X_t, V_t - \bar V_t)$ evolves according to

$$\left\{\begin{array}{lcl} dx_t &=& v_t\, dt \\ dv_t &=& -\big(A(V_t) - A(\bar V_t) + \beta\, x_t + D(X_t) - D(\bar X_t)\big)\, dt - \big(C *_x \rho[f_t](X_t) - C *_x \rho[\bar f_t](\bar X_t)\big)\, dt. \end{array}\right.$$

Then, if a and b are positive constants to be chosen later on,

$$\begin{array}{rcl} \dfrac{d}{dt}(a|x_t|^2 + 2\, x_t\cdot v_t + b|v_t|^2) &=& 2\, a\, x_t\cdot v_t + 2|v_t|^2 \\ && -2\, x_t\cdot\big(A(V_t) - A(\bar V_t) + \beta x_t + D(X_t) - D(\bar X_t)\big) \\ && -2\, b\, v_t\cdot\big(A(V_t) - A(\bar V_t) + \beta x_t + D(X_t) - D(\bar X_t)\big) \\ && -2\,(x_t + b\, v_t)\cdot\big(C *_x \rho[f_t](X_t) - C *_x \rho[\bar f_t](\bar X_t)\big). \end{array}$$

By the Cauchy-Schwarz inequality and assumptions on A and D, the first four terms are bounded by above by

$$2(a - b\beta)x_t\cdot v_t + 2(\alpha + b\delta)|x_t|\,|v_t| - 2\,(b\alpha' - 1)|v_t|^2 - 2\,(\beta - \delta)\,|x_t|^2.$$

Let now π_t be the law of $(X_t, V_t; \bar X_t, \bar V_t)$ on $\mathbb{R}^{2d}\times\mathbb{R}^{2d}$: then its marginals on $\mathbb{R}^{2d}$ are the respective distributions f_t and $\bar f_t$ of (X_t, V_t) and $(\bar X_t, \bar V_t)$, so that, since moreover C is odd:

$$\begin{array}{cl} & -2\,\mathbb{E}\, x_t\cdot\big(C *_x \rho[f_t](X_t) - C *_x \rho[\bar f_t](\bar X_t)\big) \\ = & -2\displaystyle\int_{\mathbb{R}^{8d}} (Y - \bar Y)\cdot\big(C(Y-y) - C(\bar Y - \bar y)\big)\, d\pi_t(y,w;\bar y,\bar w)\, d\pi_t(Y,W;\bar Y,\bar W) \\ = & -\displaystyle\int_{\mathbb{R}^{8d}} \big((Y-y) - (\bar Y - \bar y)\big)\cdot\big(C(Y-y) - C(\bar Y - \bar y)\big)\, d\pi_t(y,w;\bar y,\bar w)\, d\pi_t(Y,W;\bar Y,\bar W) \\ \leqslant & \gamma\displaystyle\int_{\mathbb{R}^{8d}} \big|(Y-y) - (\bar Y - \bar y)\big|^2\, d\pi_t(y,w;\bar y,\bar w)\, d\pi_t(Y,W;\bar Y,\bar W) \\ = & 2\,\gamma\Big[\displaystyle\int_{\mathbb{R}^{4d}} |y - \bar y|^2 d\pi_t(y,w;\bar y,\bar w) - \Big|\int_{\mathbb{R}^{4d}} (y - \bar y) d\pi_t(y,w;\bar y,\bar w)\Big|^2\Big] \\ \leqslant & 2\,\gamma\,\mathbb{E}|x_t|^2. \end{array}$$

In the same way, and by Young's inequality,

$$\begin{array}{cl} & -2\,\mathbb{E} v_t\cdot\big(C *_x \rho[f_t](X_t) - C *_x \rho[\bar f_t](\bar X_t)\big) \\ = & -2\displaystyle\int_{\mathbb{R}^{8d}} (W - \bar W)\cdot\big(C(Y-y) - C(\bar Y - \bar y)\big)\, d\pi_t(y,w;\bar y,\bar w)\, d\pi_t(Y,W;\bar Y,\bar W) \\ = & -\displaystyle\int_{\mathbb{R}^{8d}} \big((W - \bar W) - (w - \bar w)\big)\cdot\big(C(Y-y) - C(\bar Y - \bar y)\big)\, d\pi_t(y,w;\bar y,\bar w)\, d\pi_t(Y,W;\bar Y,\bar W) \\ \leqslant & \dfrac{\gamma}{2}\displaystyle\int_{\mathbb{R}^{8d}} \big|(W - \bar W) - (w - \bar w)\big|^2 + \big|(Y - \bar Y - (y - \bar y)\big|^2\, d\pi_t(y,w;\bar y,\bar w)\, d\pi_t(Y,W;\bar Y,\bar W) \\ \leqslant & \gamma\,\mathbb{E}[|x_t|^2 + |v_t|^2]. \end{array}$$

Collecting all terms leads to the bound

$$\begin{array}{rcl} \dfrac{d}{dt}\mathbb{E}\,(a|x_t|^2 + 2\, x_t\cdot v_t + b|v_t|^2) &\leqslant& 2(a - b\beta)\,\mathbb{E}\, x_t\cdot v_t + 2(\alpha + b\delta)\,\mathbb{E}|x_t|\,|v_t| \\ && -2\Big(\beta - \delta - \gamma - \dfrac{\gamma\, b}{2}\Big)\mathbb{E}|x_t|^2 - 2\Big(\alpha' b - 1 - \dfrac{\gamma\, b}{2}\Big)\mathbb{E}|v_t|^2 \end{array}$$

for every positive ε, and then (with $a = b\beta$) to

$$\frac{d}{dt}\mathbb{E}\,(b\beta\,|x_t|^2 + 2\, x_t\cdot v_t + b|v_t|^2) \leqslant -\Big(2\beta - 2\eta - \varepsilon - \eta b\Big)\mathbb{E}|x_t|^2 - \Big((2\alpha' - \eta)b - 2 - \frac{\alpha^2}{\varepsilon}\Big)\mathbb{E}|v_t|^2$$

by Young's inequality, where $\eta = \gamma + \delta$.

If $4 - 4\beta b^2 < 0$, that is, if $b > 1/\sqrt{\beta}$, then $Q : (x, v) \mapsto b\beta|x|^2 + 2x \cdot v + b|v|^2$ is a positive quadratic form on $\mathbb{R}^{2d}$. Then we look for b and ε such that

$$2\beta - 2\eta - \varepsilon - \eta b > 0 \quad \text{and} \quad (2\alpha' - \eta)b - 2 - \frac{\alpha^2}{\varepsilon} > 0, \tag{6}$$

in such a way that

$$\frac{d}{dt}\mathbb{E}\, Q(x_t, v_t) \leqslant -C\, \mathbb{E}\big[|x_t|^2 + |v_t|^2\big]$$

holds for a positive constant C.

Necessarily $\eta < 2\alpha'$, which is assumed in the sequel. Then, for instance for $\varepsilon = \beta$, the conditions (6) are equivalent to

$$\frac{2 + \alpha^2/\beta}{2\alpha' - \eta} < b < \frac{\beta - 2\eta}{\eta}, \quad \eta < 2\alpha'.$$

We look for η such that $\dfrac{2 + \alpha^2/\beta}{2\alpha' - \eta} < \dfrac{\beta - 2\eta}{\eta}$, that is, $2\eta^2 - \eta(2 + \dfrac{\alpha^2}{\beta} + \beta + 4\alpha') + 2\alpha'\beta 0$. This polynomial takes negative values at $\eta = 2\alpha'$, so it is positive on an interval $[0, \eta_0[$ for some $\eta_0 < 2\alpha'$. We further notice that $\eta_0 < \dfrac{\beta\sqrt{\beta}}{1 + 2\sqrt{\beta}}$, so that there exists b with all the above conditions for any $0 \leqslant \eta < \eta_0$.

Hence there exists a constant η_0, depending only on α, α' and β, such that, if $\gamma + \delta < \eta_0$, then there exist a positive quadratic form Q on $\mathbb{R}^{2d}$ and a constant C, depending only on $\alpha, \alpha', \beta, \gamma$ and δ such that

$$\frac{d}{dt}\mathbb{E}\, Q(x_t, v_t) \leqslant -C\, \mathbb{E}\big[|x_t|^2 + |v_t|^2\big]$$

for all $t \geqslant 0$. In turn, since $Q(x, v)$ and $|x|^2 + |v|^2$ are equivalent on $\mathbb{R}^{2d}$, this is bounded by $-C\, \mathbb{E}\, Q(x_t, v_t)$ for a new constant C, so that

$$\mathbb{E}\, Q\big((X_t, V_t) - (\bar{X}_t, \bar{V}_t)\big) \leqslant e^{-Ct}\, \mathbb{E}\, Q\big((X_0, V_0) - (\bar{X}_0, \bar{V}_0)\big)$$

for all $t \geqslant 0$ by integration. We finally optimize over (X_0, V_0) and $(\bar{X}_0, \bar{V}_0)$ with respective laws f_0 and $\bar{f}_0$ and use the relation $d_Q(f_t, \bar{f}_t) \leqslant \mathbb{E}\, Q((X_t, V_t) - (\bar{X}_t, \bar{V}_t))$ to deduce

$$d_Q(f_t, \bar{f}_t) \leqslant e^{-Ct}\, d_Q(f_0, \bar{f}_0).$$

This concludes the argument. □

Remark 9. *This coupling argument can also be performed for the (space homogeneous) granular media equation, for which it exactly recovers the contraction property in Wasserstein distance given in [12, Theorem 5], whence the statements which follow on the trend to equilibrium.*

We now turn to the

Proof of Theorem 1. First of all, the positive quadratic form $Q(x, v)$ on $\mathbb{R}^{2d}$ given by Proposition 8 is equivalent to $|x|^2 + |v|^2$, so there exist positive constants C'' and C' such that

$$d(f_t, \bar{f}_t) \leqslant C'' d_Q(f_t, \bar{f}_t) \leqslant C'' e^{-Ct} d_Q(f_0, \bar{f}_0) \leqslant C'\, e^{-Ct}\, d(f_0, \bar{f}_0), \qquad t \geqslant 0$$

for all solutions $(f_t)_{t\geqslant 0}$ and $(\bar{f}_t)_{t\geqslant 0}$ to (1) by the contraction property of Proposition 8: this proves the first assertion (3) of Theorem 1.

Now, if Q is the positive quadratic form on $\mathbb{R}^{2d}$ given by Proposition 8, then $\sqrt{Q}$ is a norm on $\mathbb{R}^{2d}$ so that the space $(\mathcal{P}_2, d_Q)$ is a complete metric space (see [7] or [26, Chapter 6] for instance).

Then Lemma 10 below (see [13, Lemma 7.3] for instance) and the contraction property of Proposition 8 ensure the existence of a unique stationary solution μ_∞ in $\mathcal{P}_2$ to (1):

Lemma 10. *Let $(S, dist)$ be a complete metric space and $(T(t))_{t\geqslant 0}$ be a continuous semigroup on $(S, dist)$ for which for all positive t there exists $L(t) \in]0,1[$ such that*

$$dist(T(t)(x), T(t)(y)) \leqslant L(t)\, dist(x,y)$$

for all positive t and x, y in S. Then there exists a unique stationary point x_∞ in S, that is, such that $T(t)(x_\infty) = x_\infty$ for all positive t.

Moreover all solutions $(f_t)_{t\geqslant 0}$ with initial data f_0 in $\mathcal{P}_2$ converge to this stationary solution μ_∞, with

$$d_Q(f_t, \mu_\infty) \leqslant e^{-Ct}\, d_Q(f_0, \mu_\infty), \qquad t \geqslant 0.$$

Finally, with $\bar{f}_0 = \mu_\infty$, (3) specifies into

$$d(f_t, \mu_\infty) \leqslant C' e^{-Ct} d(f_0, \mu_\infty), \qquad t \geqslant 0$$

which concludes the proof of Theorem 1. □

2 Particle approximation

The time-uniform propagation of chaos in Theorem 2 requires a time-uniform bound of the second moment of the solutions to (1).

Lemma 11. *Under Assumption $(\mathcal{A})$, for all positive α, α' and β there exists a positive constant c such that $\displaystyle\sup_{t\geqslant 0} \int_{\mathbb{R}^{2d}} (|x|^2 + |v|^2)\, f_t(x,v)\, dx\, dv$ is finite for all solutions $(f_t)_{t\geqslant 0}$ to (1) with initial datum f_0 in $\mathcal{P}_2$ and γ, δ in $[0, c)$.*

Proof of Lemma 11. Let $(f_t)_{t\geqslant 0}$ be a solution to (1) with initial datum f_0 in $\mathcal{P}_2$, and let a and b be positive numbers to be chosen later on. Then

$$\begin{aligned}\frac{d}{dt}\int_{\mathbb{R}^{2d}} &\big(a\,|x|^2 + 2\,x\cdot v + b\,|v|^2\big)\, df_t(x,v) \\ &= 2bd + 2\int_{\mathbb{R}^{2d}} v\cdot(ax+v) - (x+bv)\cdot\big(A(v) + B(x) + C *_x \rho[f_t](x)\big)\, df_t(x,v)\end{aligned}$$

where, by Young's inequality and assumption on A, B and C,

$$-2\,x\cdot A(v) = -2\,x\cdot\big(A(v) - A(0)\big) - 2\,x\cdot A(0) \leqslant 2\,\alpha\,|x|\,|v| - 2x\cdot A(0),$$

$$-2\,x\cdot B(x) = -2\,x\cdot\big(\beta\,x + D(x) - D(0) + D(0)\big) \leqslant -(2\beta - 2\delta)|x|^2 - 2\,x\cdot D(0),$$

$$-2\,b\,v\cdot A(v) = -2\,b\,v\cdot(A(v) - A(0) + A(0)\big) \leqslant -2\,b\,\alpha'|v|^2 - 2\,b\,v\cdot A(0),$$

$$-2\,b\,v\cdot B(x) = -2\,b\,v\cdot\big(\beta\,x + D(x) - D(0) + D(0)\big) \leqslant -2b\beta\,v\cdot x + 2b\delta\,|v|\,|x| - 2b\,v\cdot D(0),$$

$$\begin{aligned}-2\int_{\mathbb{R}^{2d}} x\cdot C *_x \rho[f_t](x)\, df_t(x,v) \quad &= \quad -\int_{\mathbb{R}^{4d}} (x-y)\cdot C(x-y)\, df_t(x,v)\, df_t(y,w) \\ &\leqslant \quad \gamma\int_{\mathbb{R}^{4d}} |x-y|^2\, df_t(x,v)\, df_t(y,w) \\ &\leqslant \quad 2\,\gamma\int_{\mathbb{R}^{2d}} |x|^2\, df_t(x,v)\end{aligned}$$

and

$$\begin{aligned}
-2\,b\int_{\mathbb{R}^{2d}} v\cdot C*_x\rho[f_t](x)\,df_t(x,v) &= -b\int_{\mathbb{R}^{4d}}(v-w)\cdot C(x-y)\,df_t(x,v)\,df_t(y,w)\\
&\leqslant b\,\gamma\int_{\mathbb{R}^{4d}}|v-w|\,|x-y|\,df_t(x,v)\,df_t(y,w)\\
&\leqslant \frac{b\gamma}{2}\int_{\mathbb{R}^{4d}}\big(|v-w|^2+|x-y|^2\big)\,df_t(x,v)\,df_t(y,w)\\
&\leqslant b\,\gamma\int_{\mathbb{R}^{2d}}(|x|^2+|v|^2)df_t(x,v).
\end{aligned}$$

Collecting all terms and using Young's inequality we obtain, with $a=\beta b$ and $\eta=\gamma+\delta$,

$$\begin{aligned}
&\frac{d}{dt}\int_{\mathbb{R}^{2d}}(\beta\,b\,|x|^2+2\,x\cdot v+b\,|v|^2)\,df_t(x,v)\\
\leqslant\;& 2bd+(2\alpha+2b\delta)\int|x|\,|v|\,df_t(x,v)+\big[b\gamma+2\gamma-2\beta+2\delta\big]\int|x|^2df_t(x,v)\\
&+\big[2+\gamma b-2\alpha' b\big]\int|v|^2\,df_t(x,v)-2\big(A(0)+D(0)\big)\cdot\Big(\int x\,df_t(x,v)+b\int v\,df_t(x,v)\Big)\\
\leqslant\;& 2bd+\Big(\frac{2}{\varepsilon}+\frac{b^2\varepsilon}{2\alpha^2}\Big)|A(0)+D(0)|^2\\
&-\big[2\beta-2\eta-\varepsilon-\eta b\big]\int|x|^2\,df_t(x,v)-\Big[(2\alpha'-\eta)b-2-\frac{4\alpha^2}{\varepsilon}\Big]\int|v|^2\,df_t(x,v)
\end{aligned}$$

for all positive ε.

Now, as in the proof of Proposition 8, with α replaced by 2α, we get the existence of a positive constant η_0, depending only on α,α' and β, such that for all $0\leqslant\gamma+\delta<\eta_0$ there exist b (and ε) such that $Q(x,v)=\beta\,b\,|x|^2+2\,x\cdot v+b\,|v|^2$ be a positive quadratic form on $\mathbb{R}^{2d}$ and such that

$$\frac{d}{dt}\int_{\mathbb{R}^{2d}}Q(x,v)\,f_t(x,v)\,dx\,dv\leqslant C_1-C_2\int_{\mathbb{R}^{2d}}(|x|^2+|v|^2)\,f_t(x,v)\,dx\,dv\leqslant C_1-C_3\int_{\mathbb{R}^{2d}}Q(x,v)\,f_t(x,v)\,dx\,dv$$

for positive constants C_i. It follows that

$$\sup_{t\geqslant 0}\int_{\mathbb{R}^{2d}}Q(x,v)\,f_t(x,v)\,dx\,dv<+\infty$$

if initially $\displaystyle\int_{\mathbb{R}^{2d}}Q(x,v)\,f_0(x,v)\,dx\,dv<+\infty$, that is,

$$\sup_{t\geqslant 0}\int_{\mathbb{R}^{2d}}(|x|^2+|v|^2)\,f_t(x,v)\,dx\,dv<+\infty$$

if initially f_0 belongs to $\mathcal{P}_2$. This concludes the argument. □

We now turn to the

Proof of Theorem 2. For each $1\leqslant i\leqslant N$ the law f_t of $(\bar X_t^i,\bar V_t^i)$ is the solution to (1) with f_0 as initial datum and the processes $(\bar X_t^i,\bar V_t^i)_{t\geqslant 0}$ and $(X_t^{i,N},V_t^{i,N})_{t\geqslant 0}$ are driven by the same Brownian motion. In particular the differences $x_t^i=X_t^{i,N}-\bar X_t^i$ and $v_t^i=V_t^{i,N}-\bar V_t^i$ evolve according to

$$\left\{\begin{aligned}
dx_t^i &= v_t^i\,dt\\
dv_t^i &= -\big(A(V_t^{i,N})-A(\bar V_t^i)+\beta\,x_t^i+D(X_t^{i,N})-D(\bar X_t^i)\big)\,dt-\frac{1}{N}\sum_{j=1}^N\big(C(X_t^{i,N}-X_t^{j,N})-C*_x\rho[f_t](\bar X_t^i)\big)\,dt
\end{aligned}\right.$$

with $(x_0^i, v_0^i) = (0,0)$.

Then, if a and b are positive constants to be chosen later on,

$$\begin{aligned}\frac{d}{dt}(a|x_t^i|^2 + 2\,x_t^i\cdot v_t^i + b|v_t^i|^2) \;&=\; 2\,a\,x_t^i\cdot v_t^i + 2|v_t^i|^2 - 2\,x_t^i\cdot\big(A(V_t^{i,N}) - A(\bar V_t^i) + \beta x_t^i + D(X_t^{i,N}) - D(\bar X_t^i)\big)\\ &\quad -2\,b\,v_t^i\cdot\big(A(V_t^{i,N}) - A(\bar V_t^i) + \beta x_t^i + D(X_t^{i,N}) - D(\bar X_t^i)\big)\\ &\quad -\frac{2}{N}\sum_{j=1}^N (x_t^i + b\,v_t^i)\cdot\big(C(X_t^{i,N} - X_t^{j,N}) - C *_x \rho[f_t](\bar X_t^i)\big).\end{aligned}$$

By the Young inequality and assumptions on A and D, for all positive ε the third and fourth terms are bounded by above according to

$$-2\,x_t^i\cdot(A(V_t^{i,N}) - A(\bar V_t^i)) \leqslant 2|x_t^i|\,|A(V_t^{i,N}) - A(\bar V_t^i)| \leqslant 2\,\alpha\,|x_t^i|\,|v_t^i| \leqslant \frac{\varepsilon}{2}|x_t^i|^2 + \frac{2\alpha^2}{\varepsilon}|v_t^i|^2,$$

$$-2\,x_t^i\cdot(D(X_t^{i,N}) - D(\bar X_t^i)) = -2b(X_t^i - \bar X_t^i)\cdot(D(X_t^{i,N}) - D(\bar X_t^i)) \leqslant 2\,\delta\,|x_t^i|^2,$$
$$-2\,b\,v_t^i\cdot(A(V_t^{i,N}) - A(\bar V_t^i)) = -2\,b(V_t^{i,N} - \bar V_t^i)\cdot(A(V_t^{i,N}) - A(\bar V_t^i)) \leqslant -2\,b\,\alpha|v_t^i|^2$$

and

$$-2\,b\,v_t^i\cdot(D(X_t^{i,N}) - D(\bar X_t^i)) \leqslant 2\,b\,\delta|v_t^i|\,|x_t^i| \leqslant b\,\delta(|x_t^i|^2 + |v_t^i|^2).$$

Hence, with $a = \beta\,b$,

$$\begin{aligned}\frac{d}{dt}(\beta\,b|x_t^i|^2 + 2\,x_t^i\cdot v_t^i + b|v_t^i|^2) \;&\leqslant\; \big(\frac{\varepsilon}{2} - 2\beta + 2\delta + \delta b\big)|x_t^i|^2 + \big(2 + \frac{2\alpha^2}{\varepsilon} - 2\alpha' b + \delta b\big)|v_t^i|^2\\ &\quad -\frac{2}{N}\sum_{j=1}^N (x_t^i + b\,v_t^i)\cdot\big(C(X_t^{i,N} - X_t^{j,N}) - C *_x \rho[f_t](\bar X_t^i)\big).\end{aligned}$$

Moreover, by symmetry, $\mathbb{E}|x_t^i|^2$, $\mathbb{E}x_t^i\cdot v_t^i$, ... are independent of $i = 1,\dots,N$, so that, by averaging on i,

$$\begin{aligned}\frac{d}{dt}\mathbb{E}\big[\beta\,b|x_t^1|^2 + 2\,x_t^1\cdot v_t^1 + b|v_t^1|^2\big] \;&\leqslant\; -\big(2\beta - 2\delta - \frac{\varepsilon}{2} - \delta b\big)\mathbb{E}|x_t^1|^2 - \big((2\alpha' - \delta)b - 2 - \frac{2\alpha^2}{\varepsilon}\big)\mathbb{E}|v_t^1|^2\\ &\quad -\frac{2}{N^2}\sum_{i,j=1}^N \mathbb{E}\Big[(x_t^i + b\,v_t^i)\cdot\big(C(X_t^{i,N} - X_t^{j,N}) - C *_x \rho[f_t](\bar X_t^i)\big)\Big].\end{aligned} \tag{7}$$

We decompose the last term in (7) according to

$$C(X_t^{i,N} - X_t^{j,N}) - C *_x \rho[f_t](\bar X_t^i) = C(X_t^{i,N} - X_t^{j,N}) - C(\bar X_t^i - \bar X_t^j) + C(\bar X_t^i - \bar X_t^j) - C *_x \rho[f_t](\bar X_t^i)$$

which leads to estimating four terms:

1. By symmetry and assumption on C,

$$\begin{aligned}
-\sum_{i,j=1}^{N}\mathbb{E}\Big[x_t^i\cdot\big(C(X_t^{i,N}-X_t^{j,N})-C(\bar X_t^i-\bar X_t^j)\big)\Big] &= -\sum_{i,j=1}^{N}\mathbb{E}\left[(X_t^{i,N}-\bar X_t^i)\cdot\big(C(X_t^{i,N}-X_t^{j,N})-C(\bar X_t^i-\bar X_t^j)\big)\right]\\
&= -\frac{1}{2}\sum_{i,j=1}^{N}\mathbb{E}\Big[\big((X_t^{i,N}-X_t^{j,N})-(\bar X_t^i-\bar X_t^j)\big)\\
&\qquad\cdot\big(C(X_t^{i,N}-X_t^{j,N})-C(\bar X_t^i-\bar X_t^j)\big)\Big]\\
&\leqslant \frac{\gamma}{2}\sum_{i,j=1}^{N}\mathbb{E}\big|(X_t^{i,N}-X_t^{j,N})-(\bar X_t^i-\bar X_t^j)\big|^2\\
&= \frac{\gamma}{2}\sum_{i,j=1}^{N}\mathbb{E}\big|(X_t^{i,N}-\bar X_t^i)-(\bar X_t^{j,N}-\bar X_t^j)\big|^2\\
&= \gamma\sum_{i,j=1}^{N}\mathbb{E}|x_t^i|^2-\gamma\mathbb{E}\Big|\sum_{i=1}^{N}(X_t^{i,N}-\bar X_t^i)\Big|^2\\
&\leqslant \gamma N^2|x_t^1|^2.
\end{aligned}$$

2. By assumption on C and the Young inequality,

$$\begin{aligned}
-\sum_{i,j=1}^{N}\mathbb{E}\,v_t^i\cdot\big(C(X_t^{i,N}-X_t^{j,N})-C(\bar X_t^i-\bar X_t^j)\big) &= -\frac{1}{2}\sum_{i,j=1}^{N}\mathbb{E}\big[(v_t^i-v_t^j)\cdot\big(C(X_t^{i,N}-X_t^{j,N})-C(\bar X_t^i-\bar X_t^j)\big)\big]\\
&\leqslant \frac{\gamma}{2}\sum_{i,j=1}^{N}\mathbb{E}\big[|v_t^i-v_t^j|\,\big|(X_t^{i,N}-X_t^{j,N})-(\bar X_t^i-\bar X_t^j)\big|\big]\\
&\leqslant \frac{\gamma}{2}\sum_{i,j=1}^{N}\mathbb{E}\left[\frac{1}{2}|v_t^i-v_t^j|^2+\frac{1}{2}\big|(X_t^{i,N}-\bar X_t^i)-(X_t^{j,N}-\bar X_t^j)\big|^2\right]\\
&\leqslant \frac{\gamma}{2}N^2\,\mathbb{E}(|v_t^1|^2+|x_t^1|^2).
\end{aligned}$$

3. For each $i=1,\dots,N$, and again by the Young inequality

$$-2\,\mathbb{E}\left[x_t^i\cdot\sum_{j=1}^{N}\Big(C(\bar X_t^i-\bar X_t^j)-C*_x\rho[f_t](\bar X_t^i)\Big)\right]\leqslant LN\mathbb{E}\,|x_t^i|^2+\frac{1}{LN}\mathbb{E}\Big|\sum_{j=1}^{N}\Big(C(\bar X_t^i-\bar X_t^j)-C*_x\rho[f_t](\bar X_t^i)\Big)\Big|^2$$

for any positive constant L, where the last expectation is

$$\sum_{j=1}^{N}\mathbb{E}\big|C(\bar X_t^i-\bar X_t^j)-C*_x\rho[f_t](\bar X_t^i)\big|^2+\sum_{j\neq k}\mathbb{E}\Big[\big(C(\bar X_t^i-\bar X_t^j)-C*_x\rho[f_t](\bar X_t^i)\big)\cdot\big(C(\bar X_t^i-\bar X_t^k)-C*_x\rho[f_t](\bar X_t^i)\big)\Big].$$

First of all, C is odd, so $C(0)=0$ and hence $|C(z)|\leqslant\gamma\,|z|$. Then, for each $j=1,\dots,N$,

$$\begin{aligned}
\mathbb{E}\big|C(\bar X_t^i-\bar X_t^j)-C*_x\rho[f_t](\bar X_t^i)\big|^2 &\leqslant 2\,\mathbb{E}\big|C(\bar X_t^i-\bar X_t^j)\big|^2+2\,\mathbb{E}\big|C*_x\rho[f_t](\bar X_t^i)\big|^2\\
&\leqslant 2\,\gamma^2\Big[\mathbb{E}\big|\bar X_t^i-\bar X_t^j\big|^2+\int_{\mathbb{R}^{4d}}|y-x|^2f_t(x,v)\,f_t(y,w)\,dx\,dv\,dy\,dw\Big]\\
&\leqslant 8\,\gamma^2\int_{\mathbb{R}^{2d}}|x|^2\,f_t(x,v)\,dx\,dv\\
&\leqslant M
\end{aligned}$$

for a constant M, provided γ and δ are small enough for the conclusion of Lemma 11 to hold. The constant M depends on the initial moment $\displaystyle\int_{\mathbb{R}^{2d}} \left(|x|^2+|v|^2\right) f_0(x,v)\,dx\,dv$ and the coefficients of the equation, but not on t or N.

Then, for all $j \neq k$,

$$
\begin{aligned}
\mathbb{E}\Big[&\big(C(\bar X^i_t-\bar X^j_t)-C*_x\rho[f_t](\bar X^i_t)\big)\cdot\big(C(\bar X^i_t-\bar X^k_t)-C*_x\rho[f_t](\bar X^i_t)\big)\Big]\\
&=\mathbb{E}_{\bar X^i_t}\Big[\Big(\mathbb{E}_{\bar X^j_t}\big[C(\bar X^i_t-\bar X^j_t)-C*_x\rho[f_t](\bar X^i_t)\big]\Big)\cdot\Big(\mathbb{E}_{\bar X^k_t}\big[C(\bar X^i_t-\bar X^k_t)-C*_x\rho[f_t](\bar X^i_t)\big]\Big)\Big]\\
&=\mathbb{E}_{\bar X^i_t}\big[0\big]=0
\end{aligned}
$$

since $\bar X^j_t$ and $\bar X^k_t$ are independent and have law $\rho[f_t]$.

To sum up,

$$
-2\sum_{i,j=1}^N \mathbb{E}\left[x^i_t\cdot\big(C(\bar X^i_t-\bar X^j_t)-C*_x\rho[f_t](\bar X^i_t)\big)\right]\leqslant L\,N^2\,\mathbb{E}|x^1_t|^2+\frac{M}{L}N.
$$

4. In the same way for any positive L' we obtain the bound

$$
-2\sum_{i,j=1}^N \mathbb{E}\left[v^i_t\cdot\big(C(\bar X^i_t-\bar X^j_t)-C*_x\rho[f_t](\bar X^i_t)\big)\right]\leqslant L'\,N^2\,\mathbb{E}|v^1_t|^2+\frac{M}{L'}N.
$$

Collecting all terms and letting for instance $L=\dfrac{\varepsilon}{2}$ and $L'=\dfrac{2\alpha^2}{b\varepsilon}$, it follows from (7) that there exists a positive constant c such that for all γ,δ in $[0,c)$ there exists a constant M such that

$$
\begin{aligned}
&\frac{d}{dt}\mathbb{E}\big[\beta\,b|x^1_t|^2+2\,x^1_t\cdot v^1_t+b|v^1_t|^2\big]\\
&\qquad\leqslant-(2\beta-\varepsilon-2\eta-\eta b)\mathbb{E}|x^1_t|^2-\big((2\alpha'-\eta)b-2-\frac{4\alpha^2}{\varepsilon}\big)\mathbb{E}|v^1_t|^2+\frac{M}{N}\Big(\frac{2}{\varepsilon}+\frac{\varepsilon b^2}{2\alpha^2}\Big)\cdot
\end{aligned}
$$

for all positive t,b and ε, where $\eta=\gamma+\delta$.

Now, as in the proof of Proposition 8, with α replaced by 2α, we get the existence of a positive constant η_0, depending only on α,α' and β, such that for all $0\leqslant\gamma+\delta<\eta_0$ there exist b (and ε) such that $Q(x,v)=\beta\,b\,|x|^2+2\,x\cdot v+b\,|v|^2$ be a positive quadratic form on $\mathbb{R}^{2d}$ and such that

$$
\frac{d}{dt}\mathbb{E}\,Q(x^1_t,v^1_t)\leqslant-C_1\,\mathbb{E}\big[|x^1_t|^2+|v^1_t|^2\big]+\frac{C_2}{N}
$$

for all $t\geqslant 0$ and for positive constants C_1 and C_2, also depending on f_0 through its second moment, but not on N. In turn this is bounded by $-C_3\,\mathbb{E}\,Q(x^1_t,v^1_t)+\dfrac{C_2}{N}$, so that

$$
\sup_{t\geqslant 0}\,\mathbb{E}\,Q(x^1_t,v^1_t)\leqslant\frac{C_4}{N}
$$

and finally

$$
\sup_{t\geqslant 0}\mathbb{E}\left[|X^{1,N}_t-\bar X^1_t|^2+|V^{1,N}_t-\bar V^1_t|^2\right]\leqslant\frac{C}{N}
$$

where the constant C depends on the parameters of the equation and on the second moment of f_0, but not on N. This concludes the proof of Theorem 2. □

Remark 12. *One can prove a contraction property for the particle system, similar to Proposition 8 for the Vlasov-Fokker-Planck equation: if f_0 is an initial datum in $\mathcal{P}_2$ we let $f_t^{(1,N)}$ be the common law at time t of any of the N particles $(X_t^{i,N}, V_t^{i,N})$. Then there exists a positive constant c such that, if $0 \leqslant \gamma, \delta < c$, then there exist a positive constant C and a positive quadratic form Q on $\mathbb{R}^{2d}$ such that*

$$d_Q(f_t^{(1,N)}, \tilde{f}_t^{(1,N)}) \leqslant e^{-Ct}\, d_Q(f_0^{(1,N)}, \tilde{f}_0^{(1,N)}) = e^{-Ct}\, d_Q(f_0, \tilde{f}_0)$$

for all t and all initial data f_0 and $\tilde{f}_0$ in $\mathcal{P}_2$. Here the form Q and the constants c and C depend only on the coefficients of the equations, and not on N. From this and Remark 3, and following [14], one can recover the contraction property of Proposition 8, whence Theorem 1.

3 Transportation inequality and deviation result

This final section is devoted to the proof of Theorem 5. It is based on the idea, borrowed to [15], of proving a T_2 transportation inequality not only for the law $f_T^{(N)}$ at time T, but for the whole trajectory up to time T; this transportation inequality will be proved by means of stochastic calculus, a coupling argument, a clever formulation of the relative entropy of two trajectory laws and a change of metric as in the previous sections; it will imply the announced transportation inequality by projection at time T.

We only sketch the proof, emphasizing the main steps and refering to the previous sections and to [15] for further details.

We equip the space $\mathcal{C}$ of $\mathbb{R}^{2dN}$-valued continuous functions on $[0,T]$ with the L^2 norm and consider the space $\mathcal{P}(\mathcal{C})$ of Borel probability measures on $\mathcal{C}$, equipped with the Wasserstein distance defined by the cost $\|\gamma_1 - \gamma_2\|_{L^2}^2$ for $\gamma_1, \gamma_2 \in \mathcal{C}$.

We write Equation (4) on the particle system $(X_t^{(N)}, V_t^{(N)})_{t\geqslant 0}$ in the form

$$d(X_t^{(N)}, V_t^{(N)}) = \sigma^{(N)}(X_t^{(N)}, V_t^{(N)}) dW_t^{(N)} + b^{(N)}(X_t^{(N)}, V_t^{(N)}) dt$$

for some coefficients $\sigma^{(N)}$ and $b^{(N)}$.

Let $\mathbb{P} \in \mathcal{P}(\mathcal{C})$ be the law of the trajectory $(X^{(N)}, V^{(N)}) = (X_t^{(N)}, V_t^{(N)})_{0\leqslant t\leqslant T}$ of the particles, all of them starting from the deterministic point $(x_0, v_0) \in \mathbb{R}^{2d}$.

The transportation inequality for $\mathbb{P}$, which it is sufficient to prove for laws $\mathbb{Q}$ absolutely continuous with respect to $\mathbb{P}$, will obtained in two steps.

Step 1. Following [15, Proof of Theorem 5.6], for every trajectory law $\mathbb{Q} \in \mathcal{P}(\mathcal{C})$, there exists $(\beta_t)_{t\in[0,T]} \in L^2([0,T], R^{2dN})$ such that $H(\mathbb{Q}, \mathbb{P}) = \frac{1}{2}\mathbb{E}^{\mathbb{Q}} \int_0^T |\beta_t|^2 dt$; moreover

$$d(X_t^{(N)}, V_t^{(N)}) = \sigma^{(N)}(X_t^{(N)}, V_t^{(N)}) d\tilde{W}_t^{(N)} + b^{(N)}(X_t^{(N)}, V_t^{(N)}) dt + \sigma^{(N)}(X_t^{(N)}, V_t^{(N)}) \beta_t dt$$

under the law $\mathbb{Q}$, where $\tilde{W}_t^{(N)} = W_t^{(N)} - \displaystyle\int_0^t \beta_s ds$ is a Brownian motion under $\mathbb{Q}$. We now build a coupling between $\mathbb{Q}$ and $\mathbb{P}$ by letting $(\tilde{X}^{(N)}, \tilde{V}^{(N)}) = (\tilde{X}_t^{(N)}, \tilde{V}_t^{(N)})_{0\leqslant t\leqslant T}$ be the solution (under $\mathbb{Q}$) of

$$d(\tilde{X}_t^{(N)}, \tilde{V}_t^{(N)}) = \sigma^{(N)}(\tilde{X}_t^{(N)}, \tilde{V}_t^{(N)}) d\tilde{W}_t^{(N)} + b^{(N)}(\tilde{X}_t^{(N)}, \tilde{V}_t^{(N)}) dt,$$

whose law under $\mathbb{Q}$ is exactly $\mathbb{P}$.

Step 2. In order to prove the T_2 inequality, and as in the previous sections, we change the metric induced on $\mathcal{C}$ by the L^2 norm and consider an equivalent positive quadratic form $Q(x,v) = a|x|^2 + 2\, x\cdot v + b|v|^2$. We control the quantity

$$\mathbb{E}^{\mathbb{Q}} Q\big((X_t^{(N)}, V_t^{(N)}) - (\tilde{X}_t^{(N)}, \tilde{V}_t^{(N)})\big)$$

by proving the existence of a positive constant D, independent of N and T, such that

$$\mathbb{E}^{\mathbb{Q}}\big[a|x_t|^2+2x_t\cdot v_t+b|v_t|^2\big]\leq -D\int_0^t\mathbb{E}^{\mathbb{Q}}\big[|x_s|^2+|v_s|^2\big]ds+\int_0^t\mathbb{E}^{\mathbb{Q}}\big[\nabla Q(x_s,v_s)\cdot\sigma^{(N)}(X_s^{(N)},V_s^{(N)})\beta_s\big]ds$$

in the notation $x_t=X_t^{(N)}-\tilde{X}_t^{(N)}$ and $v_t=V_t^{(N)}-\tilde{V}_t^{(N)}$. Then we bound the last term by

$$\varepsilon\int_0^t\mathbb{E}^{\mathbb{Q}}\big[|x_s|^2+|v_s|^2\big]ds+\frac{1}{\varepsilon}\int_0^t\mathbb{E}^{\mathbb{Q}}|\beta_s|^2ds$$

and the transportation inequality for the trajectory law $\mathbb{P}$ follows again by Gronwall's lemma, with a new constant D independent of T.

The transportation inequality for the law $f_T^{(N)}$ at time T finally follows by projection at time T.

We now turn to the deviation inequality in Theorem 5. First of all, if h is a 1-Lipschitz function on $\mathbb{R}^{2d}$, then

$$\begin{aligned}\frac{1}{N}\sum_{i=1}^N h(X_T^{i,N},V_T^{i,N})&-\int_{\mathbb{R}^{2dN}}\frac{1}{N}\sum_{i=1}^N h(x_i,v_i)\,df_T^{(N)}(x_1,\ldots,v_N)\\&=\frac{1}{N}\sum_{i=1}^N h(X_T^{i,N},V_T^{i,N})-\int hd\mu_\infty+\int h\,d\mu_\infty-\int h\,df_T+\int h\,df_T-\int h\,df_T^{(1,N)}\\&\geq\frac{1}{N}\sum_{i=1}^N h(X_T^{i,N},V_T^{i,N})-\int h\,d\mu_\infty-d(f_T,\mu_\infty)-d(f_T,f_T^{(1,N)})\end{aligned}$$

by exchangeability. But, by Theorem 1 with $f_0=\delta_{(x_0,v_0)}$, there exist two constants C and C', depending only on the coefficients of the equation, such that

$$d(f_T,\mu_\infty)\leqslant C'e^{-CT}\,d(f_0,\mu_\infty)$$

where f_T is the solution at time T of Equation (1) with initial datum $f_0=\delta_{(x_0,v_0)}$. Moreover, by Remark 3, there exists a constant C'', depending only on the equation and on (x_0,v_0), such that

$$d(f_T,f_T^{(1,N)})\leqslant\frac{C''}{\sqrt{N}}.$$

Hence

$$\begin{aligned}\mathbb{P}&\left[\frac{1}{N}\sum_{i=1}^N h(X_T^{i,N},V_T^{i,N})-\int_{\mathbb{R}^{2d}}h\,d\mu_\infty\geq r+D'\left(\frac{1}{\sqrt{N}}+e^{-CT}\right)\right]\\&\leq\mathbb{P}\left[\frac{1}{N}\sum_{i=1}^N h(X_T^{i,N},V_T^{i,N})-\int_{\mathbb{R}^{2dN}}\frac{1}{N}\sum_{i=1}^N h(x_i,v_i)\,df_T^{(N)}(x_1,\ldots,v_N)\geqslant r\right]\end{aligned}$$

where $D'=\max(C'd(f_0,\mu_\infty),C'')$ depends on (x_0,v_0).

Now the law $f_T^{(N)}$ satisfies a T_2 inequality on $\mathbb{R}^{2dN}$ with constant D, hence a Gaussian deviation inequality for Lipschitz functions (see [6]); moreover the map $(x_1,\ldots,v_N)\mapsto\frac{1}{N}\sum_{i=1}^N h(x_i,v_i)$ is $\frac{1}{\sqrt{N}}$-Lipschitz on $\mathbb{R}^{2dN}$, so the probability on the right-hand side is bounded by

$$\exp\left(-\frac{Nr^2}{2D}\right)$$

for all $r\geqslant 0$. This concludes the proof of Theorem 5.

References

[1] C. Ané, S. Blachère, D. Chafaï, P. Fougères, I. Gentil, F. Malrieu, C. Roberto, and G. Scheffer, *Sur les inégalités de Sobolev logarithmiques*, Panoramas et Synthèses 10, Société Math. de France, Paris, 2000.

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

[3] D. Bakry and M. Émery, *Diffusions hypercontractives*, Sém. de probabilités XIX, 1983/84, Lecture Notes in Math. 1123 (1985).

[4] D. Benedetto, E. Caglioti, J. A. Carillo, and M. Pulvirenti, *A non Maxwellian steady distribution for one-dimensional granular media*, J. Stat. Physics **91** (1998), no. 5/6, 979–990.

[5] S.G. Bobkov, I. Gentil, and M. Ledoux, *Hypercontractivity of Hamilton-Jacobi equations*, J. Math. Pures Appli. **80** (2001), no. 7, 669–696.

[6] S.G. Bobkov and F. Götze, *Exponential integrability and transportation cost related to logarithmic Sobolev inequalities*, J. Funct. Anal. **163** (1999), no. 1, 1–28.

[7] F. Bolley, *Separability and completeness for the Wasserstein distance*, Sém. de probabilités XLI, Lecture Notes in Math. 1934 (2008).

[8] ———, *Quantitative concentration inequalities on sample path space for mean field interaction*, Esaim Prob & Stat., to appear (2009).

[9] F. Bolley, A. Guillin, and C. Villani, *Quantitative concentration inequalities for empirical measures on non compact spaces*, Prob. Theor. Rel. Fields **137** (2007), no. 3-4, 541–593.

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

[11] 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. Iberoamericana **19** (2003), no. 3, 971–1018.

[12] ———, *Contractions in the 2-Wasserstein length space and thermalization of granular media*, Arch. Rat. Mech. Anal. **179** (2006), no. 2, 217–263.

[13] J. A. Carrillo and G. Toscani, *Contractive probability metrics and asymptotic behavior of dissipative kinetic equations*, Riv. Mat. Univ. Parma (7) **6** (2007), 75–198.

[14] P. Cattiaux, A. Guillin, and F. Malrieu, *Probabilistic approach for granular media equations in the non uniformly convex case*, Prob. Theor. Rel. Fields **140** (2008), no. 1-2, 19–40.

[15] H. Djellout, A. Guillin, and L. Wu, *Transportation cost-information inequalities and applications to random dynamical systems and diffusions*, Ann. Probab. **32** (2004), no. 3B, 2702–2732.

[16] F. Hérau, *Short and long time behavior of the Fokker-Planck equation in a confining potential and applications*, J. Funct. Anal. **244** (2007), no. 1, 95–118.

[17] F. Hérau and F. Nier, *Isotropic hypoellipticity and trend to the equilibrium for the Fokker-Planck equation with high degree potential*, Arch. Rat. Mech. Anal. **2** (2004), no. 171, 151–218.

[18] M. Ledoux, *The concentration of measure phenomenon*, Math. Surveys and Monographs 89, Amer. Math. Society, Providence, 2001.

[19] F. Malrieu, *Logarithmic Sobolev inequalities for some nonlinear PDE's*, Stochastic Process. Appl. **95** (2001), no. 1, 109–132.

[20] ———, *Convergence to equilibrium for granular media equations and their Euler schemes*, Ann. Appl. Probab. **13** (2003), no. 2, 540–560.

[21] S. Meléard, *Asymptotic behaviour of some interacting particle systems: Mc Kean-Vlasov and Boltzmann models*, Probabilistic models for non linear partial differential equations. Lecture Notes in Math. 1627 (1995), 42–95.

[22] F. Otto and C. Villani, *Generalization of an inequality by Talagrand, and links with the logarithmic Sobolev inequality*, J. Funct. Anal. **173** (2000), no. 2, 361–400.

[23] Esposito R., Guo Y., and Marra R., *Stability of the front under a vlasov-fokker-planck dynamics*, Arch. Rat. Mech. Anal., to appear (2009).

[24] D. Talay, *Stochastic Hamiltonian dissipative systems: exponential convergence to the invariant measure, and discretization by the implicit Euler scheme*, Mark. Proc. Rel. Fields **8** (2002), no. 2, 163–198.

[25] C. Villani, *Hypocoercivity*, Memoirs of the Amer. Math. Society, to appear, 2009.

[26] ———, *Optimal transport, old and new*, Grund. der Math. Wissenschaften, vol. 338, Springer-Verlag, Berlin, 2009.

[27] L. Wu, *Large and moderate deviations and exponential convergence for stochastic damping Hamiltonian systems*, Stoch. Proc. Appl. **91** (2001), 205–238.

François Bolley, mailto:bolley(AT)ceremade.dauphine(DOT)fr

Ceremade, UMR CNRS 7534
Université Paris-Dauphine, Place du Maréchal De Lattre De Tassigny, F-75775 Paris Cedex 16

Arnaud Guillin, corresponding author, mailto:guillin(AT)math.univ-bpclermont(DOT)fr

UMR CNRS 6620, Laboratoire de Mathématiques
Universite Blaise Pascal, avenue des Landais, F-63177 Aubiere Cedex

Florent Malrieu, mailto:florent.malrieu(AT)univ-rennes1(DOT)fr

UMR CNRS 6625, Institut de Recherche Mathématique de Rennes (IRMAR) ;
Université de Rennes I, Campus de Beaulieu, F-35042 Rennes Cedex

Compiled June 7, 2009.