# TREND TO EQUILIBRIUM AND PARTICLE APPROXIMATION FOR A WEAKLY SELFCONSISTENT VLASOV-FOKKER-PLANCK EQUATION 

François Bolley ${ }^{1}$, Arnaud Guillin ${ }^{2}$ and Florent Malrieu $^{3}$


#### 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.


Mathematics Subject Classification. 65C35, 35K55, 65C05, 82C22, 26D10, 60 E 15.
Received June 7, 2009. Revised January 18, 2010.
Published online August 26, 2010.

## 1. 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

$$
\begin{equation*}
\frac{\partial f_{t}}{\partial t}+v \cdot \nabla_{x} f_{t}-C *_{x} \rho\left[f_{t}\right](x) \cdot \nabla_{v} f_{t}=\Delta_{v} f_{t}+\nabla_{v} \cdot\left((A(v)+B(x)) f_{t}\right), \quad t>0, x, v \in \mathbb{R}^{d} \tag{1.1}
\end{equation*}
$$

where

$$
\rho\left[f_{t}\right](x)=\int_{\mathbb{R}^{d}} f_{t}(x, v) \mathrm{d} v
$$

[^0]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 $x \in \mathbb{R}^{d}$ :
$$
C *_{x} \rho\left[f_{t}\right](x)=\int_{\mathbb{R}^{d}} C(x-y) \rho\left[f_{t}\right](y) \mathrm{d} y=\int_{\mathbb{R}^{2 d}} C(x-y) f_{t}(y, v) \mathrm{d} y \mathrm{~d} v
$$

Moreover $\nabla_{x}$ stands for the gradient with respect to the position variable $x \in \mathbb{R}^{d}$ whereas $\nabla_{v}, \nabla_{v}$. and $\Delta_{v}$ respectively stand for the gradient, divergence and Laplace operators with respect to the velocity variable $v \in \mathbb{R}^{d}$.

The three maps $A, B$ and $C$ are $\mathbb{R}^{d}$ to $\mathbb{R}^{d}$ maps. 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 symmetric 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 [9] for instance). We shall consider measure solutions, which have the following natural probabilistic interpretation: if $f_{0}$ is a density function, the solution $f_{t}$ of (1.1) is the density of the law at time $t$ of the $\mathbb{R}^{2 d}$-valued process $\left(X_{t}, V_{t}\right)_{t \geqslant 0}$ evolving according to the mean field stochastic differential equation (diffusive Newton's equations)

$$
\left\{\begin{align*}
\mathrm{d} X_{t} & =V_{t} \mathrm{~d} t  \tag{1.2}\\
\mathrm{~d} V_{t} & =-A\left(V_{t}\right) \mathrm{d} t-B\left(X_{t}\right) \mathrm{d} t-C *_{x} \rho\left[f_{t}\right]\left(X_{t}\right) \mathrm{d} t+\sqrt{2} \mathrm{~d} W_{t} .
\end{align*}\right.
$$

Here $\left(W_{t}\right)_{t \geqslant 0}$ is a Brownian motion in the velocity space $\mathbb{R}^{d}$ and $f_{t}$ is the law of $\left(X_{t}, V_{t}\right)$ in $\mathbb{R}^{2 d}$, so that $\rho\left[f_{t}\right]$ 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 Cattiaux et al. in particular [13], [24,25] (see also [30] for non-gradient equations), by means of a stochastic interpretation analogous to (1.2) and a particle approximation analogous to (1.4) below. These models were interpreted as gradient flows in the space of probability measures by Carrillo et al. [11,12] (see also [32]), both approaches leading to explicit exponential (or algebraic for non uniformly convex potentials) rates of convergence to equilibrium. Time-uniform propagation of chaos was also proven for the associated particle system.

Obtaining rates of convergence to equilibrium for (1.1) 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 (1.1), when $C=0$, also known as the kinetic Fokker-Planck equation. On the one hand a probabilistic approach based on Lyapunov functionals, and thus easy to check conditions, has led Talay [29], Wu [33] or Bakry et al. [3] to exponential or subexponential convergence to equilibrium in total variation distance. On the other hand the case when $A(v)=v$ and $B(x)=\nabla \Psi(x)$, and when the equilibrium solution is explicitly given by $f_{\infty}(x, v)=\mathrm{e}^{-\Psi(x)-|v|^{2} / 2}$ was studied in [18,21,22] and [31], 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 $\mathrm{e}^{-\Psi}$, to ensure $L^{2}$ or entropic convergence with an explicit exponential rate. We also refer to [20] for the evolution of two species, modelled by two coupled Vlasov-Fokker-Planck equations.

Villani's approach extends to the selfconsistent situation when $C$ derives from a nonzero potential $U$ (see [31], Chap. 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 $\mathrm{e}^{-|v|^{2} / 2}$.

In this work we consider the case when the equation is set on the whole $\mathbb{R}^{d}$, with linear-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 fulfilled if there exist nonnegative constants $\alpha, \alpha^{\prime}, \beta, \gamma$ and $\delta$ such that

$$
\begin{aligned}
|A(v)-A(w)| & \leqslant \alpha|v-w|, \quad(v-w) \cdot(A(v)-A(w)) \geqslant \alpha^{\prime}|v-w|^{2}, \\
B(x) & =\beta x+D(x) \quad \text { where }|D(x)-D(y)| \leqslant \delta|x-y|
\end{aligned}
$$

and

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

for all $x, y, v, w$ in $\mathbb{R}^{d}$.
One may also consider the case of non-constant diffusion coefficients: then our method would require smallness assumptions on their Lipschitz seminorms with respect to $\alpha, \alpha^{\prime}$ and $\beta$.

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

$$
d(\mu, \nu)^{2}=\inf _{(X, V),(Y, W)} \mathbb{E}\left(|X-Y|^{2}+|V-W|^{2}\right)
$$

where the infimum runs over all the couples $(X, V)$ and $(Y, W)$ of random variables on $\mathbb{R}^{2 d}$ with respective laws $\mu$ and $\nu$. Convergence in this metric is equivalent to narrow convergence plus convergence of the second moment (see [32], Chap. 6, for instance).

The coefficients $A, B$ and $C$ being Lipschitz, existence and uniqueness for equation (1.2) with squareintegrable initial data are ensured by [26]. It follows that, for all initial data $f_{0}$ in $\mathcal{P}_{2}$, equation (1.1) admits a unique measure solution in $\mathcal{P}_{2}$, that is, continuous on $\left[0,+\infty\left[\right.\right.$ with values in $\mathcal{P}_{2}$.

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

Theorem 1.1. Under Assumption ( $\mathcal{A}$ ), for all positive $\alpha, \alpha^{\prime}$ and $\beta$ there exists a positive constant $c$ such that, if $0 \leqslant \gamma, \delta<c$, then there exist positive constants $C$ and $C^{\prime}$ such that

$$
\begin{equation*}
d\left(f_{t}, \bar{f}_{t}\right) \leqslant C^{\prime} \mathrm{e}^{-C t} d\left(f_{0}, \bar{f}_{0}\right), \quad t \geqslant 0 \tag{1.3}
\end{equation*}
$$

for all measure solutions $\left(f_{t}\right)_{t \geqslant 0}$ and $\left(\bar{f}_{t}\right)_{t \geqslant 0}$ to (1.1) with respective initial data $f_{0}$ and $\bar{f}_{0}$ in $\mathcal{P}_{2}$.
Moreover (1.1) admits a unique stationary solution $\mu_{\infty}$ and all solutions $\left(f_{t}\right)_{t \geqslant 0}$ converge towards it, with

$$
d\left(f_{t}, \mu_{\infty}\right) \leqslant C^{\prime} \mathrm{e}^{-C t} d\left(f_{0}, \mu_{\infty}\right), \quad t \geqslant 0
$$

For instance, for $\alpha=\alpha^{\prime}=\beta=1$, the general proof below shows that the nonnegative $\gamma$ and $\delta$ 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 $\gamma$ and $\delta$ 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 - the latter being a stronger convergence since, in this specific situation, the equilibrium measure $\mathrm{e}^{-|v|^{2} / 2}$ satisfies a logarithmic Sobolev inequality, hence a transportation inequality. We refer to $[1,23]$ or [32], 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 cases. 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). Theorem 1.1 provides the convergence to equilibrium for a large class of initial data, but also a stability result of all solutions. Let us finally note that it is based on the natural stochastic interpretation (1.2) and a simple
coupling argument, and does not need any regularity property of the solutions that could be deduced from the hypoellipticity of the model.

The particle approximation of solutions to (1.1) consists in the introduction of a large number $N$ of $\mathbb{R}^{2 d}$-valued processes $\left(X_{t}^{i, N}, V_{t}^{i, N}\right)_{t \geqslant 0}$ with $1 \leqslant i \leqslant N$, evolving according to the force field $C *_{x} \rho\left[\hat{\mu}_{t}^{N}\right]$ generated by the empirical measure $\hat{\mu}_{t}^{N}=\frac{1}{N} \sum_{i=1}^{N} \delta_{\left(X_{t}^{i, N}, V_{t}^{i, N}\right)}$ of the system: if $\left(W^{i}\right)_{i \geqslant 1}$ with $i \geqslant 1$ are independent standard Brownian motions on $\mathbb{R}^{d}$ and $\left(X_{0}^{i}, V_{0}^{i}\right)$ with $i \geqslant 1$ are independent random vectors on $\mathbb{R}^{2 d}$ with law $f_{0}$ in $\mathcal{P}_{2}$ and independent of $\left(W_{.}^{i}\right)_{i \geqslant 1}$, we let $\left(X_{t}^{(N)}, V_{t}^{(N)}\right)_{t \geqslant 0}=\left(X_{t}^{1, N}, \ldots, X_{t}^{N, N}, V_{t}^{1, N}, \ldots, V_{t}^{N, N}\right)_{t \geqslant 0}$ be the solution of the following stochastic differential equation in $\left(\mathbb{R}^{2 d}\right)^{N}$ :

$$
\left\{\begin{array}{l}
\mathrm{d} X_{t}^{i, N}=V_{t}^{i, N} \mathrm{~d} t  \tag{1.4}\\
\mathrm{~d} V_{t}^{i, N}=-A\left(V_{t}^{i, N}\right) \mathrm{d} t-B\left(X_{t}^{i, N}\right) \mathrm{d} t-\frac{1}{N} \sum_{j=1}^{N} C\left(X_{t}^{i, N}-X_{t}^{j, N}\right) \mathrm{d} t+\sqrt{2} \mathrm{~d} W_{t}^{i}, \quad 1 \leqslant i \leqslant N \\
\left(X_{0}^{i, N}, V_{0}^{i, N}\right)=\left(X_{0}^{i}, V_{0}^{i}\right)
\end{array}\right.
$$

The mean field force $C *_{x} \rho\left[f_{t}\right]$ in (1.2) is replaced by the sum of the pairwise actions $\frac{1}{N} C\left(X_{t}^{i, N}-X_{t}^{j, N}\right)$ 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 $i \geqslant 1,\left(\bar{X}_{t}^{i}, \bar{V}_{t}^{i}\right)_{t \geqslant 0}$ be the solution of the kinetic McKean-Vlasov type equation on $\mathbb{R}^{2 d}$

$$
\left\{\begin{array}{l}
\mathrm{d} \bar{X}_{t}^{i}=\bar{V}_{t}^{i} \mathrm{~d} t  \tag{1.5}\\
\mathrm{~d} \bar{V}_{t}^{i}=-A\left(\bar{V}_{t}^{i}\right) \mathrm{d} t-B\left(\bar{X}_{t}^{i}\right) \mathrm{d} t-C *_{x} \rho\left[\nu_{t}\right]\left(\bar{X}_{t}^{i}\right) \mathrm{d} t+\sqrt{2} \mathrm{~d} W_{t}^{i} \\
\left(\bar{X}_{0}^{i}, \bar{V}_{0}^{i}\right)=\left(X_{0}^{i}, V_{0}^{i}\right)
\end{array}\right.
$$

where $\nu_{t}$ is the distribution of $\left(\bar{X}_{t}^{i}, \bar{V}_{t}^{i}\right)$. The processes $\left(\bar{X}_{t}^{i}, \bar{V}_{t}^{i}\right)_{t \geqslant 0}$ 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 evolves according to (1.1), so at time $t$ is the solution $f_{t}$ of (1.1) with initial datum $f_{0}$. In this notation, and as $N$ gets large, the $N$ processes $\left(X_{t}^{i, N}, V_{t}^{i, N}\right)_{t \geqslant 0}$ look more and more like the $N$ independent processes $\left(\bar{X}_{t}^{i}, \bar{V}_{t}^{i}\right)_{t \geqslant 0}$ :
Theorem 1.2 (time-uniform propagation of chaos). Let $\left(X_{0}^{i}, V_{0}^{i}\right)$ for $1 \leqslant i \leqslant N$ be $N$ independent $\mathbb{R}^{2 d}$-valued random variables with law $f_{0}$ in $\mathcal{P}_{2}\left(\mathbb{R}^{2 d}\right)$. Let also $\left(X_{t}^{i, N}, V_{t}^{i, N}\right)_{t \geqslant 0,1 \leqslant i \leqslant N}$ be the solution to (1.2) and $\left(\bar{X}_{t}^{i}, \bar{V}_{t}^{i}\right)_{t \geqslant 0}$ the solution to (1.5) with initial datum $\left(X_{0}^{i}, V_{0}^{i}\right)$ for $1 \leqslant i \leqslant N$. Under Assumption $(\mathcal{A})$, for all positive $\alpha, \alpha^{\prime}$ and $\beta$ 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 $i=1, \ldots, N$

$$
\sup _{t \geqslant 0} \mathbb{E}\left(\left|X_{t}^{i, N}-\bar{X}_{t}^{i}\right|^{2}+\left|V_{t}^{i, N}-\bar{V}_{t}^{i}\right|^{2}\right) \leqslant \frac{C}{N}
$$

Here the constant $C$ depends only on the coefficients of the equation and the second moment of $f_{0}$.
Remark 1.3. In particular the law $f_{t}^{(1, N)}$ at time $t$ of any $\left(X_{t}^{i, N}, V_{t}^{i, N}\right)$ (by symmetry) converges to $f_{t}$ as $N$ goes to infinity, according to

$$
d\left(f_{t}^{(1, N)}, f_{t}\right)^{2} \leqslant \mathbb{E}\left(\left|X_{t}^{i, N}-\bar{X}_{t}^{i}\right|^{2}+\left|V_{t}^{i, N}-\bar{V}_{t}^{i}\right|^{2}\right) \leqslant \frac{C}{N}
$$

Theorem 1.2 also ensures propagation of chaos for the particle system as it is defined in the famous review by Sznitman [28]: let $k \geqslant 1$ be fixed, and let $\varphi:\left(\mathbb{R}^{2 d}\right)^{k} \rightarrow \mathbb{R}$ be $l$-Lipschitz with respect to each argument. Then, for all $k$-uple $\left(i_{1}, \ldots, i_{k}\right)$ of $[1, N]$, and letting $Z_{t}^{i, N}=\left(X_{t}^{i, N}, V_{t}^{i, N}\right)$ and $\bar{Z}_{t}^{i}=\left(\bar{X}_{t}^{i}, \bar{V}_{t}^{i}\right)$, the quantity

$$
\left|\mathbb{E} \varphi\left(Z_{t}^{i_{1}, N}, \ldots, Z_{t}^{i_{k}, N}\right)-\int_{\mathbb{R}^{2 k d}} \varphi \mathrm{~d} f_{t}^{\otimes k}\right|=\left|\mathbb{E} \varphi\left(Z_{t}^{i_{1}, N}, \ldots, Z_{t}^{i_{k}, N}\right)-\mathbb{E} \varphi\left(\bar{Z}_{t}^{i_{1}}, \ldots, \bar{Z}_{t}^{i_{k}}\right)\right|
$$

is bounded by

$$
\begin{aligned}
& \mathbb{E}\left|\varphi\left(Z_{t}^{i_{1}, N}, Z_{t}^{i_{2}, N}, \ldots, Z_{t}^{i_{k}, N}\right)-\varphi\left(\bar{Z}_{t}^{i_{1}}, Z_{t}^{i_{2}, N}, \ldots, Z_{t}^{i_{k}, N}\right)\right|+\ldots+\mathbb{E} \mid \varphi\left(\bar{Z}_{t}^{i_{1}}, \ldots, \bar{Z}_{t}^{i_{k-1}}, Z_{t}^{i_{k}, N}\right) \\
&-\varphi\left(\bar{Z}_{t}^{i_{1}}, \ldots, \bar{Z}_{t}^{i_{k-1}}, \bar{Z}_{t}^{i_{k}}\right) \mid
\end{aligned}
$$

whence by

$$
l \sum_{j=1}^{k} \mathbb{E}\left|Z_{t}^{i_{j}, N}-\bar{Z}_{t}^{i_{j}}\right| \leqslant C l \frac{k}{\sqrt{N}}
$$

where $C$ is independent of $N$ : hence it tends to 0 as $N$ tends to infinity.
Propagation of chaos at the level of the trajectories, and not only of the time-marginals, is studied in $[7,26]$ for a broad class of equations, but with time dependent constants.

We finally turn to the approximation of the equilibrium solution of the Vlasov-Fokker-Planck equation (as given by Thm. 1.1) by the particle system at a given time $T$.

Since all solutions $\left(f_{t}\right)_{t \geqslant 0}$ to (1.1) with initial data in $\mathcal{P}_{2}$ converge to the equilibrium solution $\mu_{\infty}$, we let $\left(x_{0}, v_{0}\right)$ in $\mathbb{R}^{2 d}$ be given and we consider the Dirac mass $\delta_{\left(x_{0}, v_{0}\right)}$ at $\left(x_{0}, v_{0}\right)$ as the initial datum $f_{0}$. We shall give precise bounds on the approximation of $\mu_{\infty}$ by the empirical measure of the particles $\left(X_{t}^{i, N}, V_{t}^{i, N}\right)$ for $1 \leqslant i \leqslant N$, all of them initially at $\left(x_{0}, v_{0}\right)$.

In the space homogeneous case of the granular media equation, this was performed by the third author $[24,25]$ 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 [2]). 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 1.4. 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 $\left|\nabla_{x} f\right|^{2}+\left|\nabla_{v} f\right|^{2}$; however it does not satisfy a logarithmic Sobolev inequality with carré du champ $\left|\nabla_{v} f\right|^{2}$, corresponding to the diffusion term $\Delta_{v} f$, which would at once lead to exponential entropic convergence to equilibrium for the particle system (see [2]).

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

$$
d(\mu, \nu)^{2} \leq D H(\nu \mid \mu)
$$

for all probability measures $\nu$; here

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

if $\nu \ll \mu$, and $+\infty$ otherwise, is the relative entropy of $\nu$ with respect to $\mu$.
Theorem 1.5. Under Assumption $(\mathcal{A})$, for all positive $\alpha, \alpha^{\prime}$ and $\beta$ there exists a positive constant $c$ such that if $0 \leqslant \gamma, \delta<c$, then the joint law of the $N$ particles $\left(X_{T}^{(i, N)}, V_{T}^{(i, N)}\right)$ at given time $T$, all with deterministic starting
points $\left(x_{0}, v_{0}\right) \in \mathbb{R}^{2 d}$, satisfies a $T_{2}$ inequality with a constant $D$ independent of the number $N$ of particles, of time $T$ and of the point $\left(x_{0}, v_{0}\right)$.

It follows that there exists a constant $D^{\prime}$ such that

$$
\mathbb{P}\left(\frac{1}{N} \sum_{i=1}^{N} h\left(X_{T}^{i, N}, V_{T}^{i, N}\right)-\int_{\mathbb{R}^{2 d}} h \mathrm{~d} \mu_{\infty}(h) \geq r+D^{\prime}\left(\frac{1}{\sqrt{N}}+\mathrm{e}^{-C T}\right)\right) \leq \exp \left(-\frac{N r^{2}}{2 D}\right)
$$

for all $N, T, r \geqslant 0$ and all 1-Lipschitz observables $h$ on $\mathbb{R}^{2 d}$.
Here the constant $C$ has been obtained in Theorem 1.1 and the constant $D^{\prime}$ depends only on the point $\left(x_{0}, v_{0}\right)$ and the coefficients of the equation.

Remark 1.6. Such single observable deviation inequalities were obtained in [24] for the space homogeneous granular media equation; they were upgraded in [8] 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 1.7. 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 1.1, 1.2 and 1.5 still hold but with constants growing exponentially fast with time $T$.

As a conclusion to this introduction, let us point out that our model belongs to the huge class of mean field interacting particle models. Another important family of such processes consists in the Nambu and Moran type interacting jump processes and their associated Feynmac-Kac semigroups, which play a role in filtering problems or in the study of Schrödinger operators for example (see [14]). For this kind of models, asymptotic stability, uniform propagation of chaos and sharp concentration bounds have been obtained in [14-17] or [27] for instance.

Sections 2-4 are respectively devoted to the proofs of Theorems 1.1, 1.2 and 1.5.

## 2. Long time behaviour for the Vlasov-Fokker-Planck equation

This section is devoted to the proof of Theorem 1.1, which is based on the stochastic interpretation (1.2) of (1.1) and a coupling argument. The main idea, also present in [29,31], is to perturb the Euclidean metric on $\mathbb{R}^{2 d}$ in such a way that (1.2) is dissipative for this metric.

If $Q$ is a positive definite quadratic form on $\mathbb{R}^{2 d}$ and $\mu$ and $\nu$ 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 $\mathbb{R}^{2 d}$ with respective laws $\mu$ and $\nu$; so that $d_{Q}=d$ if $Q$ is the squared Euclidean norm on $\mathbb{R}^{2 d}$. The key step in the proof is the following:

Proposition 2.1. Under the assumptions of Theorem 1.1, there exist a positive constant $C$ and a positive definite quadratic form $Q$ on $\mathbb{R}^{2 d}$ such that

$$
d_{Q}\left(f_{t}, \bar{f}_{t}\right) \leqslant \mathrm{e}^{-C t} d_{Q}\left(f_{0}, \bar{f}_{0}\right), \quad t \geqslant 0
$$

for all solutions $\left(f_{t}\right)_{t \geqslant 0}$ and $\left(\bar{f}_{t}\right)_{t \geqslant 0}$ to (1.1) with respective initial data $f_{0}$ and $\bar{f}_{0}$ in $\mathcal{P}_{2}$.
Proof of Proposition 2.1. Let $\left(f_{t}\right)_{t \geqslant 0}$ and $\left(\bar{f}_{t}\right)_{t \geqslant 0}$ be two solutions to (1.1) with initial data $f_{0}$ and $\bar{f}_{0}$ in $\mathcal{P}_{2}$. Let also $\left(X_{0}, V_{0}\right)$ and $\left(\bar{X}_{0}, \bar{V}_{0}\right)$ with respectively law $f_{0}$ and $\bar{f}_{0}$, evolving into $\left(X_{t}, V_{t}\right)$ and ( $\left.\bar{X}_{t}, \bar{V}_{t}\right)$ according to (1.2),
both with the same Brownian motion $\left(W_{t}\right)_{t \geqslant 0}$ in $\mathbb{R}^{d}$. Then, by difference, $\left(x_{t}, v_{t}\right)=\left(X_{t}-\bar{X}_{t}, V_{t}-\bar{V}_{t}\right)$ evolves according to

$$
\left\{\begin{array}{l}
\mathrm{d} x_{t}=v_{t} \mathrm{~d} t \\
\mathrm{~d} v_{t}=-\left(A\left(V_{t}\right)-A\left(\bar{V}_{t}\right)+\beta x_{t}+D\left(X_{t}\right)-D\left(\bar{X}_{t}\right)\right) \mathrm{d} t-\left(C *_{x} \rho\left[f_{t}\right]\left(X_{t}\right)-C *_{x} \rho\left[\bar{f}_{t}\right]\left(\bar{X}_{t}\right)\right) \mathrm{d} t .
\end{array}\right.
$$

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

$$
\begin{aligned}
\frac{\mathrm{d}}{\mathrm{~d} t}\left(a\left|x_{t}\right|^{2}+2 x_{t} \cdot v_{t}+b\left|v_{t}\right|^{2}\right)= & 2 a x_{t} \cdot v_{t}+2\left|v_{t}\right|^{2}-2 x_{t} \cdot\left(A\left(V_{t}\right)-A\left(\bar{V}_{t}\right)+\beta x_{t}+D\left(X_{t}\right)-D\left(\bar{X}_{t}\right)\right) \\
& -2 b v_{t} \cdot\left(A\left(V_{t}\right)-A\left(\bar{V}_{t}\right)+\beta x_{t}+D\left(X_{t}\right)-D\left(\bar{X}_{t}\right)\right) \\
& -2\left(x_{t}+b v_{t}\right) \cdot\left(C *_{x} \rho\left[f_{t}\right]\left(X_{t}\right)-C *_{x} \rho\left[\bar{f}_{t}\right]\left(\bar{X}_{t}\right)\right) .
\end{aligned}
$$

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

$$
2(a-b \beta) x_{t} \cdot v_{t}+2(\alpha+b \delta)\left|x_{t}\right|\left|v_{t}\right|-2\left(b \alpha^{\prime}-1\right)\left|v_{t}\right|^{2}-2(\beta-\delta)\left|x_{t}\right|^{2} .
$$

Let now $\pi_{t}$ be the law of $\left(X_{t}, V_{t} ; \bar{X}_{t}, \bar{V}_{t}\right)$ on $\mathbb{R}^{2 d} \times \mathbb{R}^{2 d}$ : then its marginals on $\mathbb{R}^{2 d}$ are the respective distributions $f_{t}$ and $\bar{f}_{t}$ of $\left(X_{t}, V_{t}\right)$ and $\left(\bar{X}_{t}, \bar{V}_{t}\right)$, so that, since moreover $C$ is odd:

$$
\begin{aligned}
& -2 \mathbb{E} x_{t} \cdot\left(C *_{x} \rho\left[f_{t}\right]\left(X_{t}\right)-C *_{x} \rho\left[\bar{f}_{t}\right]\left(\bar{X}_{t}\right)\right) \\
= & -2 \int_{\mathbb{R}^{8 d}}(Y-\bar{Y}) \cdot(C(Y-y)-C(\bar{Y}-\bar{y})) \times \mathrm{d} \pi_{t}(y, w ; \bar{y}, \bar{w}) \mathrm{d} \pi_{t}(Y, W ; \bar{Y}, \bar{W}) \\
= & -\int_{\mathbb{R}^{8 d}}((Y-y)-(\bar{Y}-\bar{y})) \cdot(C(Y-y)-C(\bar{Y}-\bar{y})) \mathrm{d} \pi_{t}(y, w ; \bar{y}, \bar{w}) \mathrm{d} \pi_{t}(Y, W ; \bar{Y}, \bar{W}) \\
\leqslant & \gamma \int_{\mathbb{R}^{8 d}}|(Y-y)-(\bar{Y}-\bar{y})|^{2} \mathrm{~d} \pi_{t}(y, w ; \bar{y}, \bar{w}) \mathrm{d} \pi_{t}(Y, W ; \bar{Y}, \bar{W}) \\
= & 2 \gamma\left[\int_{\mathbb{R}^{4 \mathrm{~d}}}|y-\bar{y}|^{2} \mathrm{~d} \pi_{t}(y, w ; \bar{y}, \bar{w})-\left|\int_{\mathbb{R}^{4 \mathrm{~d}}}(y-\bar{y}) \mathrm{d} \pi_{t}(y, w ; \bar{y}, \bar{w})\right|^{2}\right] \\
\leqslant & 2 \gamma \mathbb{E}\left|x_{t}\right|^{2} .
\end{aligned}
$$

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

$$
\begin{aligned}
& -2 \mathbb{E} v_{t} \cdot\left(C *_{x} \rho\left[f_{t}\right]\left(X_{t}\right)-C *_{x} \rho\left[\bar{f}_{t}\right]\left(\bar{X}_{t}\right)\right) \\
= & -2 \int_{\mathbb{R}^{8 d}}(W-\bar{W}) \cdot(C(Y-y)-C(\bar{Y}-\bar{y})) \times \mathrm{d} \pi_{t}(y, w ; \bar{y}, \bar{w}) \mathrm{d} \pi_{t}(Y, W ; \bar{Y}, \bar{W}) \\
= & -\int_{\mathbb{R}^{8 d}}((W-\bar{W})-(w-\bar{w})) \cdot(C(Y-y)-C(\bar{Y}-\bar{y})) \mathrm{d} \pi_{t}(y, w ; \bar{y}, \bar{w}) \mathrm{d} \pi_{t}(Y, W ; \bar{Y}, \bar{W}) \\
\leqslant & \frac{\gamma}{2} \int_{\mathbb{R}^{8 d}}|(W-\bar{W})-(w-\bar{w})|^{2}+|(Y-\bar{Y})-(y-\bar{y})|^{2} \mathrm{~d} \pi_{t}(y, w ; \bar{y}, \bar{w}) \mathrm{d} \pi_{t}(Y, W ; \bar{Y}, \bar{W}) \\
\leqslant & \gamma \mathbb{E}\left[\left|x_{t}\right|^{2}+\left|v_{t}\right|^{2}\right] .
\end{aligned}
$$

Collecting all terms leads to the bound

$$
\begin{aligned}
\frac{\mathrm{d}}{\mathrm{~d} t} \mathbb{E}\left(a\left|x_{t}\right|^{2}+2 x_{t} \cdot v_{t}+b\left|v_{t}\right|^{2}\right) \leqslant & 2(a-b \beta) \mathbb{E} x_{t} \cdot v_{t}+2(\alpha+b \delta) \mathbb{E}\left|x_{t}\right|\left|v_{t}\right| \\
& -2\left(\beta-\delta-\gamma-\frac{\gamma b}{2}\right) \mathbb{E}\left|x_{t}\right|^{2}-2\left(\alpha^{\prime} b-1-\frac{\gamma b}{2}\right) \mathbb{E}\left|v_{t}\right|^{2}
\end{aligned}
$$

and then (with $a=b \beta$ ) to

$$
\frac{\mathrm{d}}{\mathrm{~d} t} \mathbb{E}\left(b \beta\left|x_{t}\right|^{2}+2 x_{t} \cdot v_{t}+b\left|v_{t}\right|^{2}\right) \leqslant-(2 \beta-2 \eta-\varepsilon-\eta b) \mathbb{E}\left|x_{t}\right|^{2}-\left(\left(2 \alpha^{\prime}-\eta\right) b-2-\frac{\alpha^{2}}{\varepsilon}\right) \mathbb{E}\left|v_{t}\right|^{2}
$$

for every positive $\varepsilon$ : here we have used by Young's inequality and let $\eta=\gamma+\delta$.
If $4-4 \beta b^{2}<0$, that is, if $b>1 / \sqrt{\beta}$, then $Q:(x, v) \mapsto b \beta|x|^{2}+2 x \cdot v+b|v|^{2}$ is a positive definite quadratic form on $\mathbb{R}^{2 d}$. Then we look for $b$ and $\varepsilon$ such that

$$
\begin{equation*}
2 \beta-2 \eta-\varepsilon-\eta b>0 \quad \text { and } \quad\left(2 \alpha^{\prime}-\eta\right) b-2-\frac{\alpha^{2}}{\varepsilon}>0 \tag{2.1}
\end{equation*}
$$

in such a way that

$$
\frac{\mathrm{d}}{\mathrm{~d} t} \mathbb{E} Q\left(x_{t}, v_{t}\right) \leqslant-C \mathbb{E}\left[\left|x_{t}\right|^{2}+\left|v_{t}\right|^{2}\right]
$$

holds for a positive constant $C$.
Necessarily $\eta<2 \alpha^{\prime}$, which is assumed in the sequel. Then, for instance for $\varepsilon=\beta$, the conditions (2.1) are equivalent to

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

We look for $\eta$ such that $\frac{2+\alpha^{2} / \beta}{2 \alpha^{\prime}-\eta}<\frac{\beta-2 \eta}{\eta}$, that is, $2 \eta^{2}-\eta\left(2+\frac{\alpha^{2}}{\beta}+\beta+4 \alpha^{\prime}\right)+2 \alpha^{\prime} \beta>0$. This polynomial takes a positive value at $\eta=0$ and a negative value at $\eta=2 \alpha^{\prime}$, so it is positive on an interval $\left[0, \eta_{0}\right.$ [ for some $\eta_{0}<2 \alpha^{\prime}$. We further notice that $\eta_{0}<\frac{\beta \sqrt{\beta}}{1+2 \sqrt{\beta}}$, so that there exists $b$ with all the above conditions for any $0 \leqslant \eta<\eta_{0}$.

Hence there exists a constant $\eta_{0}$, depending only on $\alpha, \alpha^{\prime}$ and $\beta$, such that, if $\gamma+\delta<\eta_{0}$, then there exist a positive definite quadratic form $Q$ on $\mathbb{R}^{2 d}$ and a constant $C$, depending only on $\alpha, \alpha^{\prime}, \beta, \gamma$ and $\delta$ such that

$$
\frac{\mathrm{d}}{\mathrm{~d} t} \mathbb{E} Q\left(x_{t}, v_{t}\right) \leqslant-C \mathbb{E}\left[\left|x_{t}\right|^{2}+\left|v_{t}\right|^{2}\right]
$$

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

$$
\mathbb{E} Q\left(\left(X_{t}, V_{t}\right)-\left(\bar{X}_{t}, \bar{V}_{t}\right)\right) \leqslant \mathrm{e}^{-C t} \mathbb{E} Q\left(\left(X_{0}, V_{0}\right)-\left(\bar{X}_{0}, \bar{V}_{0}\right)\right)
$$

for all $t \geqslant 0$ by integration. We finally optimize over $\left(X_{0}, V_{0}\right)$ and $\left(\bar{X}_{0}, \bar{V}_{0}\right)$ with respective laws $f_{0}$ and $\bar{f}_{0}$ and use the relation $d_{Q}\left(f_{t}, \bar{f}_{t}\right) \leqslant \mathbb{E} Q\left(\left(X_{t}, V_{t}\right)-\left(\bar{X}_{t}, \bar{V}_{t}\right)\right)$ to deduce

$$
d_{Q}\left(f_{t}, \bar{f}_{t}\right) \leqslant \mathrm{e}^{-C t} d_{Q}\left(f_{0}, \bar{f}_{0}\right)
$$

This concludes the argument.
Remark 2.2. This coupling argument can also be performed for the (space homogeneous) granular media equation; in this case it exactly recovers the contraction property in Wasserstein distance given in [12], Theorem 5, and then the statements which follow on the trend to equilibrium.

We now turn to the:
Proof of Theorem 1.1. First of all, the positive definite quadratic form $Q(x, v)$ on $\mathbb{R}^{2 d}$ given by Proposition 2.1 is equivalent to $|x|^{2}+|v|^{2}$, so there exist positive constants $C^{\prime \prime}$ and $C^{\prime}$ such that

$$
d\left(f_{t}, \bar{f}_{t}\right) \leqslant C^{\prime \prime} d_{Q}\left(f_{t}, \bar{f}_{t}\right) \leqslant C^{\prime \prime} \mathrm{e}^{-C t} d_{Q}\left(f_{0}, \bar{f}_{0}\right) \leqslant C^{\prime} \mathrm{e}^{-C t} d\left(f_{0}, \bar{f}_{0}\right), \quad t \geqslant 0
$$

for all solutions $\left(f_{t}\right)_{t \geqslant 0}$ and $\left(\bar{f}_{t}\right)_{t \geqslant 0}$ to (1.1) by the contraction property of Proposition 2.1: this proves the first assertion (1.3) of Theorem 1.1.

Now, if $Q$ is the positive definite quadratic form on $\mathbb{R}^{2 d}$ given by Proposition 2.1, then $\sqrt{Q}$ is a norm on $\mathbb{R}^{2 d}$ so that the space $\left(\mathcal{P}_{2}, d_{Q}\right)$ is a complete metric space (see [6] or [32], Chap. 6 , for instance).

Then Lemma 2.3 below (see [10], Lem. 7.3, for instance) and the contraction property of Proposition 2.1 ensure the existence of a unique stationary solution $\mu_{\infty}$ in $\mathcal{P}_{2}$ to (1.1):

Lemma 2.3. 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

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

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

Finally, with $\bar{f}_{0}=\mu_{\infty},(1.3)$ specifies into

$$
d\left(f_{t}, \mu_{\infty}\right) \leqslant C^{\prime} \mathrm{e}^{-C t} d\left(f_{0}, \mu_{\infty}\right), \quad t \geqslant 0
$$

which concludes the proof of Theorem 1.1.

## 3. Particle approximation

The time-uniform propagation of chaos in Theorem 1.2 requires a time-uniform bound of the second moment of the solutions to (1.1).

Lemma 3.1. Under Assumption $(\mathcal{A})$, for all positive $\alpha, \alpha^{\prime}$ and $\beta$ there exists a positive constant $c$ such that $\sup _{t \geqslant 0} \int_{\mathbb{R}^{2 d}}\left(|x|^{2}+|v|^{2}\right) \mathrm{d} f_{t}(x, v)$ is finite for all solutions $\left(f_{t}\right)_{t \geqslant 0}$ to (1.1) with initial datum $f_{0}$ in $\mathcal{P}_{2}$ and $\gamma, \delta$ in $[0, c)$.

Lemma 3.1, as suggested by a referee, can be seen as consequence of Theorem 1.1 since, for all solutions $\left(f_{t}\right)_{t \geqslant 0}$ to (1.1) with initial datum $f_{0}$ in $\mathcal{P}_{2}$,

$$
\int_{\mathbb{R}^{2 d}}\left(|x|^{2}+|v|^{2}\right) \mathrm{d} f_{t}(x, v)=d\left(f_{t}, \delta_{0}\right)^{2} \leqslant 2 d\left(f_{t}, \mu_{\infty}\right)^{2}+2 d\left(\mu_{\infty}, \delta_{0}\right)^{2} \leqslant 2 C^{\prime} \mathrm{e}^{-C t} d\left(f_{0}, \mu_{\infty}\right)^{2}+2 d\left(\mu_{\infty}, \delta_{0}\right)^{2}
$$

is bounded in $t$.
However we prefer to give an elementary proof of it, to stress on the fact that our particle approximation estimate is independent of the long time behaviour of the solutions.

Proof of Lemma 3.1. Let $\left(f_{t}\right)_{t \geqslant 0}$ be a solution to (1.1) with initial datum $f_{0}$ in $\mathcal{P}_{2}$, and let $a$ and $b$ be positive numbers to be chosen later on. Then

$$
\begin{aligned}
\frac{\mathrm{d}}{\mathrm{~d} t} \int_{\mathbb{R}^{2 d}}\left(a|x|^{2}+2 x \cdot v+b|v|^{2}\right) \mathrm{d} f_{t}(x, v)=2 b d+2 \int_{\mathbb{R}^{2 d}}(v \cdot(a x+v)-(x+b v) \cdot(A(v) & +B(x) \\
& \left.\left.+C *_{x} \rho\left[f_{t}\right](x)\right)\right) \mathrm{d} f_{t}(x, v)
\end{aligned}
$$

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

$$
\begin{aligned}
-2 x \cdot A(v) & =-2 x \cdot(A(v)-A(0))-2 x \cdot A(0) \leqslant 2 \alpha|x||v|-2 x \cdot A(0), \\
-2 x \cdot B(x) & =-2 x \cdot(\beta x+D(x)-D(0)+D(0)) \leqslant-(2 \beta-2 \delta)|x|^{2}-2 x \cdot D(0), \\
-2 b v \cdot A(v) & =-2 b v \cdot(A(v)-A(0)+A(0)) \leqslant-2 b \alpha^{\prime}|v|^{2}-2 b v \cdot A(0), \\
-2 b v \cdot B(x) & =-2 b v \cdot(\beta x+D(x)-D(0)+D(0)) \leqslant-2 b \beta v \cdot x+2 b \delta|v||x|-2 b v \cdot D(0), \\
-2 \int_{\mathbb{R}^{2 d}} x \cdot C *_{x} \rho\left[f_{t}\right](x) \mathrm{d} f_{t}(x, v) & =-\int_{\mathbb{R}^{4 d}}(x-y) \cdot C(x-y) \mathrm{d} f_{t}(x, v) \mathrm{d} f_{t}(y, w) \\
& \leqslant \gamma \int_{\mathbb{R}^{4 d}}|x-y|^{2} \mathrm{~d} f_{t}(x, v) \mathrm{d} f_{t}(y, w) \\
& \leqslant 2 \gamma \int_{\mathbb{R}^{2 d}}|x|^{2} \mathrm{~d} f_{t}(x, v)
\end{aligned}
$$

and

$$
\begin{aligned}
-2 b \int_{\mathbb{R}^{2 d}} v \cdot C *_{x} \rho\left[f_{t}\right](x) \mathrm{d} f_{t}(x, v) & =-b \int_{\mathbb{R}^{4 d}}(v-w) \cdot C(x-y) \mathrm{d} f_{t}(x, v) \mathrm{d} f_{t}(y, w) \\
& \leqslant b \gamma \int_{\mathbb{R}^{4 d}}|v-w||x-y| \mathrm{d} f_{t}(x, v) \mathrm{d} f_{t}(y, w) \\
& \leqslant \frac{b \gamma}{2} \int_{\mathbb{R}^{4 d}}\left(|v-w|^{2}+|x-y|^{2}\right) \mathrm{d} f_{t}(x, v) \mathrm{d} f_{t}(y, w) \\
& \leqslant b \gamma \int_{\mathbb{R}^{2 d}}\left(|x|^{2}+|v|^{2}\right) \mathrm{d} f_{t}(x, v)
\end{aligned}
$$

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

$$
\begin{aligned}
& \quad \frac{\mathrm{d}}{\mathrm{~d} t} \int_{\mathbb{R}^{2 d}}\left(\beta b|x|^{2}+2 x \cdot v+b|v|^{2}\right) \mathrm{d} f_{t}(x, v) \leqslant 2 b d+(2 \alpha+2 b \delta) \int|x||v| \mathrm{d} f_{t}(x, v)+[b \gamma+2 \gamma-2 \beta+2 \delta] \\
& \quad \times \int|x|^{2} \mathrm{~d} f_{t}(x, v)+\left[2+\gamma b-2 \alpha^{\prime} b\right] \int|v|^{2} \mathrm{~d} f_{t}(x, v)-2(A(0)+D(0)) \cdot\left(\int x \mathrm{~d} f_{t}(x, v)+b \int v \mathrm{~d} f_{t}(x, v)\right) \\
& \leqslant \\
& \quad 2 b d+\left(\frac{2}{\varepsilon}+\frac{b^{2} \varepsilon}{2 \alpha^{2}}\right)|A(0)+D(0)|^{2}-[2 \beta-2 \eta-\varepsilon-\eta b] \int|x|^{2} \mathrm{~d} f_{t}(x, v)-\left[\left(2 \alpha^{\prime}-\eta\right) b-2-\frac{4 \alpha^{2}}{\varepsilon}\right] \\
& \quad \times \int|v|^{2} \mathrm{~d} f_{t}(x, v)
\end{aligned}
$$

for all positive $\varepsilon$.
Now, as in the proof of Proposition 2.1, with $\alpha$ replaced by $2 \alpha$, we get the existence of a positive constant $\eta_{0}$, depending only on $\alpha, \alpha^{\prime}$ and $\beta$, such that for all $0 \leqslant \gamma+\delta<\eta_{0}$ there exist $b$ (and $\varepsilon$ ) such that $Q(x, v)=$ $\beta b|x|^{2}+2 x \cdot v+b|v|^{2}$ be a positive definite quadratic form on $\mathbb{R}^{2 d}$ and such that

$$
\frac{\mathrm{d}}{\mathrm{~d} t} \int_{\mathbb{R}^{2 d}} Q(x, v) \mathrm{d} f_{t}(x, v) \leqslant C_{1}-C_{2} \int_{\mathbb{R}^{2 d}}\left(|x|^{2}+|v|^{2}\right) \mathrm{d} f_{t}(x, v) \leqslant C_{1}-C_{3} \int_{\mathbb{R}^{2 d}} Q(x, v) \mathrm{d} f_{t}(x, v)
$$

for positive constants $C_{i}$. It follows that

$$
\sup _{t \geqslant 0} \int_{\mathbb{R}^{2 d}} Q(x, v) \mathrm{d} f_{t}(x, v)<+\infty
$$

if initially $\int_{\mathbb{R}^{2 d}} Q(x, v) \mathrm{d} f_{0}(x, v)<+\infty$, that is,

$$
\sup _{t \geqslant 0} \int_{\mathbb{R}^{2 d}}\left(|x|^{2}+|v|^{2}\right) \mathrm{d} f_{t}(x, v)<+\infty
$$

if initially $f_{0}$ belongs to $\mathcal{P}_{2}$. This concludes the argument.
We now turn to the:

Proof of Theorem 1.2. For each $1 \leqslant i \leqslant N$ the law $f_{t}$ of $\left(\bar{X}_{t}^{i}, \bar{V}_{t}^{i}\right)$ is the solution to (1.1) with $f_{0}$ as initial datum and the processes $\left(\bar{X}_{t}^{i}, \bar{V}_{t}^{i}\right)_{t \geqslant 0}$ and $\left(X_{t}^{i, N}, V_{t}^{i, N}\right)_{t \geqslant 0}$ 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

$$
\left\{\begin{array}{l}
\mathrm{d} x_{t}^{i}=v_{t}^{i} \mathrm{~d} t \\
\mathrm{~d} v_{t}^{i}=-\left(A\left(V_{t}^{i, N}\right)-A\left(\bar{V}_{t}^{i}\right)+\beta x_{t}^{i}+D\left(X_{t}^{i, N}\right)-D\left(\bar{X}_{t}^{i}\right)\right) \mathrm{d} t-\frac{1}{N} \sum_{j=1}^{N}\left(C\left(X_{t}^{i, N}-X_{t}^{j, N}\right)-C *_{x} \rho\left[f_{t}\right]\left(\bar{X}_{t}^{i}\right)\right) \mathrm{d} t
\end{array}\right.
$$

with $\left(x_{0}^{i}, v_{0}^{i}\right)=(0,0)$.
Then, if $a$ and $b$ are positive constants to be chosen later on,

$$
\begin{aligned}
\frac{\mathrm{d}}{\mathrm{~d} t}\left(a\left|x_{t}^{i}\right|^{2}+2 x_{t}^{i} \cdot v_{t}^{i}+b\left|v_{t}^{i}\right|^{2}\right)= & 2 a x_{t}^{i} \cdot v_{t}^{i}+2\left|v_{t}^{i}\right|^{2}-2 x_{t}^{i} \cdot\left(A\left(V_{t}^{i, N}\right)-A\left(\bar{V}_{t}^{i}\right)+\beta x_{t}^{i}+D\left(X_{t}^{i, N}\right)-D\left(\bar{X}_{t}^{i}\right)\right) \\
& -2 b v_{t}^{i} \cdot\left(A\left(V_{t}^{i, N}\right)-A\left(\bar{V}_{t}^{i}\right)+\beta x_{t}^{i}+D\left(X_{t}^{i, N}\right)-D\left(\bar{X}_{t}^{i}\right)\right) \\
& -\frac{2}{N} \sum_{j=1}^{N}\left(x_{t}^{i}+b v_{t}^{i}\right) \cdot\left(C\left(X_{t}^{i, N}-X_{t}^{j, N}\right)-C *_{x} \rho\left[f_{t}\right]\left(\bar{X}_{t}^{i}\right)\right)
\end{aligned}
$$

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

$$
\begin{aligned}
& -2 x_{t}^{i} \cdot\left(A\left(V_{t}^{i, N}\right)-A\left(\bar{V}_{t}^{i}\right)\right) \leqslant 2\left|x_{t}^{i}\right|\left|A\left(V_{t}^{i, N}\right)-A\left(\bar{V}_{t}^{i}\right)\right| \leqslant 2 \alpha\left|x_{t}^{i}\right|\left|v_{t}^{i}\right| \leqslant \frac{\varepsilon}{2}\left|x_{t}^{i}\right|^{2}+\frac{2 \alpha^{2}}{\varepsilon}\left|v_{t}^{i}\right|^{2}, \\
& -2 x_{t}^{i} \cdot\left(D\left(X_{t}^{i, N}\right)-D\left(\bar{X}_{t}^{i}\right)\right)=-2\left(X_{t}^{i, N}-\bar{X}_{t}^{i}\right) \cdot\left(D\left(X_{t}^{i, N}\right)-D\left(\bar{X}_{t}^{i}\right)\right) \leqslant 2 \delta\left|x_{t}^{i}\right|^{2} \\
& -2 b v_{t}^{i} \cdot\left(A\left(V_{t}^{i, N}\right)-A\left(\bar{V}_{t}^{i}\right)\right)=-2 b\left(V_{t}^{i, N}-\bar{V}_{t}^{i}\right) \cdot\left(A\left(V_{t}^{i, N}\right)-A\left(\bar{V}_{t}^{i}\right)\right) \leqslant-2 b \alpha^{\prime}\left|v_{t}^{i}\right|^{2}
\end{aligned}
$$

and

$$
-2 b v_{t}^{i} \cdot\left(D\left(X_{t}^{i, N}\right)-D\left(\bar{X}_{t}^{i}\right)\right) \leqslant 2 b \delta\left|v_{t}^{i}\right|\left|x_{t}^{i}\right| \leqslant b \delta\left(\left|x_{t}^{i}\right|^{2}+\left|v_{t}^{i}\right|^{2}\right)
$$

Hence, with $a=\beta b$,

$$
\begin{aligned}
\frac{\mathrm{d}}{\mathrm{~d} t}\left(\beta b\left|x_{t}^{i}\right|^{2}+2 x_{t}^{i} \cdot v_{t}^{i}+b\left|v_{t}^{i}\right|^{2}\right) \leqslant & \left(\frac{\varepsilon}{2}-2 \beta+2 \delta+\delta b\right)\left|x_{t}^{i}\right|^{2}+\left(2+\frac{2 \alpha^{2}}{\varepsilon}-2 \alpha^{\prime} b+\delta b\right)\left|v_{t}^{i}\right|^{2} \\
& -\frac{2}{N} \sum_{j=1}^{N}\left(x_{t}^{i}+b v_{t}^{i}\right) \cdot\left(C\left(X_{t}^{i, N}-X_{t}^{j, N}\right)-C *_{x} \rho\left[f_{t}\right]\left(\bar{X}_{t}^{i}\right)\right)
\end{aligned}
$$

Moreover, by symmetry, $\mathbb{E}\left|x_{t}^{i}\right|^{2}, \mathbb{E}\left(x_{t}^{i} \cdot v_{t}^{i}\right), \ldots$ are independent of $i=1, \ldots, N$, so that, by averaging on $i$,

$$
\begin{align*}
\frac{\mathrm{d}}{\mathrm{~d} t} \mathbb{E}\left[\beta b\left|x_{t}^{1}\right|^{2}+2 x_{t}^{1} \cdot v_{t}^{1}+b\left|v_{t}^{1}\right|^{2}\right] \leqslant & -\left(2 \beta-2 \delta-\frac{\varepsilon}{2}-\delta b\right) \mathbb{E}\left|x_{t}^{1}\right|^{2}-\left(\left(2 \alpha^{\prime}-\delta\right) b-2-\frac{2 \alpha^{2}}{\varepsilon}\right) \mathbb{E}\left|v_{t}^{1}\right|^{2} \\
& -\frac{2}{N^{2}} \sum_{i, j=1}^{N} \mathbb{E}\left[\left(x_{t}^{i}+b v_{t}^{i}\right) \cdot\left(C\left(X_{t}^{i, N}-X_{t}^{j, N}\right)-C *_{x} \rho\left[f_{t}\right]\left(\bar{X}_{t}^{i}\right)\right)\right] \tag{3.1}
\end{align*}
$$

We decompose the last term in (3.1) according to

$$
C\left(X_{t}^{i, N}-X_{t}^{j, N}\right)-C *_{x} \rho\left[f_{t}\right]\left(\bar{X}_{t}^{i}\right)=C\left(X_{t}^{i, N}-X_{t}^{j, N}\right)-C\left(\bar{X}_{t}^{i}-\bar{X}_{t}^{j}\right)+C\left(\bar{X}_{t}^{i}-\bar{X}_{t}^{j}\right)-C *_{x} \rho\left[f_{t}\right]\left(\bar{X}_{t}^{i}\right)
$$

which leads to estimating four terms:
(1) By symmetry the quantity

$$
-\sum_{i, j=1}^{N} \mathbb{E}\left[x_{t}^{i} \cdot\left(C\left(X_{t}^{i, N}-X_{t}^{j, N}\right)-C\left(\bar{X}_{t}^{i}-\bar{X}_{t}^{j}\right)\right)\right]
$$

is equal to

$$
\begin{aligned}
-\sum_{i, j=1}^{N} \mathbb{E}\left[\left(X_{t}^{i, N}-\bar{X}_{t}^{i}\right) \cdot\left(C\left(X_{t}^{i, N}-X_{t}^{j, N}\right)-C\left(\bar{X}_{t}^{i}-\bar{X}_{t}^{j}\right)\right)\right] & =-\frac{1}{2} \sum_{i, j=1}^{N} \mathbb{E}\left[\left(\left(X_{t}^{i, N}-X_{t}^{j, N}\right)\right.\right. \\
& \left.\left.-\left(\bar{X}_{t}^{i}-\bar{X}_{t}^{j}\right)\right) \cdot\left(C\left(X_{t}^{i, N}-X_{t}^{j, N}\right)-C\left(\bar{X}_{t}^{i}-\bar{X}_{t}^{j}\right)\right)\right]
\end{aligned}
$$

Hence, by assumption on $C$, it is bounded by

$$
\begin{aligned}
\frac{\gamma}{2} \sum_{i, j=1}^{N} \mathbb{E}\left|\left(X_{t}^{i, N}-X_{t}^{j, N}\right)-\left(\bar{X}_{t}^{i}-\bar{X}_{t}^{j}\right)\right|^{2} & =\frac{\gamma}{2} \sum_{i, j=1}^{N} \mathbb{E}\left|\left(X_{t}^{i, N}-\bar{X}_{t}^{i}\right)-\left(\bar{X}_{t}^{j, N}-\bar{X}_{t}^{j}\right)\right|^{2} \\
& =\gamma \sum_{i, j=1}^{N} \mathbb{E}\left|x_{t}^{i}\right|^{2}-\gamma \mathbb{E}\left|\sum_{i=1}^{N}\left(X_{t}^{i, N}-\bar{X}_{t}^{i}\right)\right|^{2} \\
& \leqslant \gamma N^{2} \mathbb{E}\left|x_{t}^{1}\right|^{2} .
\end{aligned}
$$

(2) By assumption on $C$ and the Young inequality,

$$
\begin{aligned}
-\sum_{i, j=1}^{N} \mathbb{E} v_{t}^{i} \cdot\left(C\left(X_{t}^{i, N}-X_{t}^{j, N}\right)-C\left(\bar{X}_{t}^{i}-\bar{X}_{t}^{j}\right)\right) & =-\frac{1}{2} \sum_{i, j=1}^{N} \mathbb{E}\left[\left(v_{t}^{i}-v_{t}^{j}\right) \cdot\left(C\left(X_{t}^{i, N}-X_{t}^{j, N}\right)-C\left(\bar{X}_{t}^{i}-\bar{X}_{t}^{j}\right)\right)\right] \\
& \leqslant \frac{\gamma}{2} \sum_{i, j=1}^{N} \mathbb{E}\left[\left|v_{t}^{i}-v_{t}^{j}\right|\left|\left(X_{t}^{i, N}-X_{t}^{j, N}\right)-\left(\bar{X}_{t}^{i}-\bar{X}_{t}^{j}\right)\right|\right] \\
& \leqslant \frac{\gamma}{2} \sum_{i, j=1}^{N} \mathbb{E}\left[\frac{1}{2}\left|v_{t}^{i}-v_{t}^{j}\right|^{2}+\frac{1}{2}\left|\left(X_{t}^{i, N}-\bar{X}_{t}^{i}\right)-\left(X_{t}^{j, N}-\bar{X}_{t}^{j}\right)\right|^{2}\right] \\
& \leqslant \frac{\gamma}{2} N^{2} \mathbb{E}\left(\left|v_{t}^{1}\right|^{2}+\left|x_{t}^{1}\right|^{2}\right)
\end{aligned}
$$

(3) For each $i=1, \ldots, N$, and again by the Young inequality

$$
-2 \mathbb{E}\left[x_{t}^{i} \cdot \sum_{j=1}^{N}\left(C\left(\bar{X}_{t}^{i}-\bar{X}_{t}^{j}\right)-C *_{x} \rho\left[f_{t}\right]\left(\bar{X}_{t}^{i}\right)\right)\right] \leqslant L N \mathbb{E}\left|x_{t}^{i}\right|^{2}+\frac{1}{L N} \mathbb{E}\left|\sum_{j=1}^{N}\left(C\left(\bar{X}_{t}^{i}-\bar{X}_{t}^{j}\right)-C *_{x} \rho\left[f_{t}\right]\left(\bar{X}_{t}^{i}\right)\right)\right|^{2}
$$

for any positive constant $L$, where the last expectation is

$$
\sum_{j=1}^{N} \mathbb{E}\left|C\left(\bar{X}_{t}^{i}-\bar{X}_{t}^{j}\right)-C *_{x} \rho\left[f_{t}\right]\left(\bar{X}_{t}^{i}\right)\right|^{2}+\sum_{j \neq k} \mathbb{E}\left[\left(C\left(\bar{X}_{t}^{i}-\bar{X}_{t}^{j}\right)-C *_{x} \rho\left[f_{t}\right]\left(\bar{X}_{t}^{i}\right)\right) \cdot\left(C\left(\bar{X}_{t}^{i}-\bar{X}_{t}^{k}\right)-C *_{x} \rho\left[f_{t}\right]\left(\bar{X}_{t}^{i}\right)\right)\right] .
$$

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

$$
\begin{aligned}
\mathbb{E}\left|C\left(\bar{X}_{t}^{i}-\bar{X}_{t}^{j}\right)-C *_{x} \rho\left[f_{t}\right]\left(\bar{X}_{t}^{i}\right)\right|^{2} & \leqslant 2 \mathbb{E}\left|C\left(\bar{X}_{t}^{i}-\bar{X}_{t}^{j}\right)\right|^{2}+2 \mathbb{E}\left|C *_{x} \rho\left[f_{t}\right]\left(\bar{X}_{t}^{i}\right)\right|^{2} \\
& \leqslant 2 \gamma^{2}\left[\mathbb{E}\left|\bar{X}_{t}^{i}-\bar{X}_{t}^{j}\right|^{2}+\int_{\mathbb{R}^{4 d}}|y-x|^{2} f_{t}(x, v) f_{t}(y, w) \mathrm{d} x \mathrm{~d} v \mathrm{~d} y \mathrm{~d} w\right] \\
& \leqslant 8 \gamma^{2} \int_{\mathbb{R}^{2 d}}|x|^{2} f_{t}(x, v) \mathrm{d} x \mathrm{~d} v \\
& \leqslant M
\end{aligned}
$$

for a constant $M$, provided $\gamma$ and $\delta$ are small enough for the conclusion of Lemma 3.1 to hold. The constant $M$ depends on the initial moment $\int_{\mathbb{R}^{2 d}}\left(|x|^{2}+|v|^{2}\right) \mathrm{d} f_{0}(x, v)$ and the coefficients of the equation, but not on $t$ or $N$. Then, for all $j \neq k$,

$$
\begin{aligned}
& \mathbb{E}\left[\left(C\left(\bar{X}_{t}^{i}-\bar{X}_{t}^{j}\right)-C *_{x} \rho\left[f_{t}\right]\left(\bar{X}_{t}^{i}\right)\right) \cdot\left(C\left(\bar{X}_{t}^{i}-\bar{X}_{t}^{k}\right)-C *_{x} \rho\left[f_{t}\right]\left(\bar{X}_{t}^{i}\right)\right)\right] \\
& \quad=\mathbb{E}_{\bar{X}_{t}^{i}}\left[\left(\mathbb{E}_{\bar{X}_{t}^{j}}\left[C\left(\bar{X}_{t}^{i}-\bar{X}_{t}^{j}\right)-C *_{x} \rho\left[f_{t}\right]\left(\bar{X}_{t}^{i}\right)\right]\right) \cdot\left(\mathbb{E}_{\bar{X}_{t}^{k}}\left[C\left(\bar{X}_{t}^{i}-\bar{X}_{t}^{k}\right)-C *_{x} \rho\left[f_{t}\right]\left(\bar{X}_{t}^{i}\right)\right]\right)\right]=\mathbb{E}_{\bar{X}_{t}^{i}}[0]=0
\end{aligned}
$$

since $\bar{X}_{t}^{j}$ and $\bar{X}_{t}^{k}$ are independent and have law $\rho\left[f_{t}\right]$.

To sum up,

$$
-2 \sum_{i, j=1}^{N} \mathbb{E}\left[x_{t}^{i} \cdot\left(C\left(\bar{X}_{t}^{i}-\bar{X}_{t}^{j}\right)-C *_{x} \rho\left[f_{t}\right]\left(\bar{X}_{t}^{i}\right)\right)\right] \leqslant L N^{2} \mathbb{E}\left|x_{t}^{1}\right|^{2}+\frac{M}{L} N
$$

(4) In the same way for any positive $L^{\prime}$ we obtain the bound

$$
-2 \sum_{i, j=1}^{N} \mathbb{E}\left[v_{t}^{i} \cdot\left(C\left(\bar{X}_{t}^{i}-\bar{X}_{t}^{j}\right)-C *_{x} \rho\left[f_{t}\right]\left(\bar{X}_{t}^{i}\right)\right)\right] \leqslant L^{\prime} N^{2} \mathbb{E}\left|v_{t}^{1}\right|^{2}+\frac{M}{L^{\prime}} N
$$

Collecting all terms and letting for instance $L=\frac{\varepsilon}{2}$ and $L^{\prime}=\frac{2 \alpha^{2}}{b \varepsilon}$, it follows from (3.1) that there exists a positive constant $c$ such that for all $\gamma, \delta$ in $[0, c)$ there exists a constant $M$ such that

$$
\begin{aligned}
\frac{\mathrm{d}}{\mathrm{~d} t} \mathbb{E}\left[\beta b\left|x_{t}^{1}\right|^{2}+2 x_{t}^{1} \cdot v_{t}^{1}+b\left|v_{t}^{1}\right|^{2}\right] & \\
& \leqslant-(2 \beta-\varepsilon-2 \eta-\eta b) \mathbb{E}\left|x_{t}^{1}\right|^{2}-\left(\left(2 \alpha^{\prime}-\eta\right) b-2-\frac{4 \alpha^{2}}{\varepsilon}\right) \mathbb{E}\left|v_{t}^{1}\right|^{2}+\frac{M}{N}\left(\frac{2}{\varepsilon}+\frac{\varepsilon b^{2}}{2 \alpha^{2}}\right)
\end{aligned}
$$

for all positive $t, b$ and $\varepsilon$, where $\eta=\gamma+\delta$.
Now, as in the proof of Proposition 2.1, with $\alpha$ replaced by $2 \alpha$, we get the existence of a positive constant $\eta_{0}$, depending only on $\alpha, \alpha^{\prime}$ and $\beta$, such that for all $0 \leqslant \gamma+\delta<\eta_{0}$ there exist $b$ (and $\varepsilon$ ) such that $Q(x, v)=$ $\beta b|x|^{2}+2 x \cdot v+b|v|^{2}$ be a positive definite quadratic form on $\mathbb{R}^{2 d}$ and such that

$$
\frac{\mathrm{d}}{\mathrm{~d} t} \mathbb{E} Q\left(x_{t}^{1}, v_{t}^{1}\right) \leqslant-C_{1} \mathbb{E}\left[\left|x_{t}^{1}\right|^{2}+\left|v_{t}^{1}\right|^{2}\right]+\frac{C_{2}}{N}
$$

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 $-C_{3} \mathbb{E} Q\left(x_{t}^{1}, v_{t}^{1}\right)+\frac{C_{2}}{N}$, so that

$$
\sup _{t \geqslant 0} \mathbb{E} Q\left(x_{t}^{1}, v_{t}^{1}\right) \leqslant \frac{C_{4}}{N}
$$

and finally

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

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 1.2.

Remark 3.2. One can prove a contraction property for the particle system, similar to Proposition 2.1 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 $\left(X_{t}^{i, N}, V_{t}^{i, N}\right)$. 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 definite quadratic form $Q$ on $\mathbb{R}^{2 d}$ such that

$$
d_{Q}\left(f_{t}^{(1, N)}, \tilde{f}_{t}^{(1, N)}\right) \leqslant \mathrm{e}^{-C t} d_{Q}\left(f_{0}^{(1, N)}, \tilde{f}_{0}^{(1, N)}\right)=\mathrm{e}^{-C t} d_{Q}\left(f_{0}, \tilde{f}_{0}\right)
$$

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 1.3, and following [13], one can recover the contraction property of Proposition 2.1, whence Theorem 1.1.

## 4. Transportation inequality and deviation result

This final section is devoted to the proof of Theorem 1.5. It is based on the idea, borrowed from [19], of proving a $T_{2}$ transportation inequality not only for the joint law $f_{T}^{(N)}$ of the $N$-particle system 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. The announced transportation inequality will follow since then the projection at time $T$ of the entropy is well identified, the metric used is the $L^{2}$ metric and the computation performed for this metric may be used again for the law of the trajectory at time $T$.

We only sketch the proof, emphasizing the main steps and referring to the previous sections and to [19] for further details.

We equip the space $\mathcal{C}$ of $\mathbb{R}^{2 d N}$-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 $\left\|\gamma_{1}-\gamma_{2}\right\|_{L^{2}}^{2}$ for $\gamma_{1}, \gamma_{2} \in \mathcal{C}$.

We write equation (1.4) on the particle system $\left(X_{t}^{(N)}, V_{t}^{(N)}\right)_{t \geqslant 0}$ in the form

$$
\left(\mathrm{d} X_{t}^{(N)}, \mathrm{d} V_{t}^{(N)}\right)=\sigma^{(N)}\left(X_{t}^{(N)}, V_{t}^{(N)}\right) \mathrm{d} W_{t}^{(N)}+b^{(N)}\left(X_{t}^{(N)}, V_{t}^{(N)}\right) \mathrm{d} t
$$

for some coefficients $\sigma^{(N)}$ and $b^{(N)}$.
Let $\mathbb{P} \in \mathcal{P}(\mathcal{C})$ be the law of the trajectory $\left(X^{(N)}, V^{(N)}\right)=\left(X_{t}^{(N)}, V_{t}^{(N)}\right)_{0 \leqslant t \leqslant T}$ of the particles, all of them starting from the deterministic point $\left(x_{0}, v_{0}\right) \in \mathbb{R}^{2 d}$.

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

Step 1. Following [19], Proof of Theorem 5.6, for every trajectory law $\mathbb{Q} \in \mathcal{P}(\mathcal{C})$, there exists (adapted to the Brownian filtration) $\left(\beta_{t}\right)_{t \in[0, T]} \in L^{2}\left([0, T], R^{2 d N}\right)$ such that $H(\mathbb{Q} \mid \mathbb{P})=\frac{1}{2} \mathbb{E}^{\mathbb{Q}} \int_{0}^{T}\left|\beta_{t}\right|^{2} \mathrm{~d} t$; moreover

$$
\left(\mathrm{d} X_{t}^{(N)}, \mathrm{d} V_{t}^{(N)}\right)=\sigma^{(N)}\left(X_{t}^{(N)}, V_{t}^{(N)}\right) \mathrm{d} \tilde{W}_{t}^{(N)}+b^{(N)}\left(X_{t}^{(N)}, V_{t}^{(N)}\right) \mathrm{d} t+\sigma^{(N)}\left(X_{t}^{(N)}, V_{t}^{(N)}\right) \beta_{t} \mathrm{~d} t
$$

under the law $\mathbb{Q}$, where $\tilde{W}_{t}^{(N)}=W_{t}^{(N)}-\int_{0}^{t} \beta_{s} \mathrm{~d} s$ is a Brownian motion under $\mathbb{Q}$. We now build a coupling between $\mathbb{Q}$ and $\mathbb{P}$ by letting $\left(\tilde{X}^{(N)}, \tilde{V}^{(N)}\right)=\left(\tilde{X}_{t}^{(N)}, \tilde{V}_{t}^{(N)}\right)_{0 \leqslant t \leqslant T}$ be the solution (under $\mathbb{Q}$ ) of

$$
\left(\mathrm{d} \tilde{X}_{t}^{(N)}, \mathrm{d} \tilde{V}_{t}^{(N)}\right)=\sigma^{(N)}\left(\tilde{X}_{t}^{(N)}, \tilde{V}_{t}^{(N)}\right) \mathrm{d} \tilde{W}_{t}^{(N)}+b^{(N)}\left(\tilde{X}_{t}^{(N)}, \tilde{V}_{t}^{(N)}\right) \mathrm{d} t
$$

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 definite quadratic form $Q(x, v)=a|x|^{2}+2 x \cdot v+b|v|^{2}$. We have to control the following quantity to get the transportation inequality for the law of the trajectory:

$$
\int_{0}^{T} \mathbb{E}^{\mathbb{Q}} Q\left(\left(X_{t}^{(N)}, V_{t}^{(N)}\right)-\left(\tilde{X}_{t}^{(N)}, \tilde{V}_{t}^{(N)}\right)\right) \mathrm{d} t
$$

which amounts to the control of

$$
\mathbb{E}^{\mathbb{Q}} Q\left(\left(X_{t}^{(N)}, V_{t}^{(N)}\right)-\left(\tilde{X}_{t}^{(N)}, \tilde{V}_{t}^{(N)}\right)\right),
$$

needed also for the transportation inequality for the law at time $t$. We then prove the existence of a positive constant $D$, independent of $N$ and $T$, such that

$$
\mathbb{E}^{\mathbb{Q}}\left[a\left|x_{t}\right|^{2}+2 x_{t} \cdot v_{t}+b\left|v_{t}\right|^{2}\right] \leq-D \int_{0}^{t} \mathbb{E}^{\mathbb{Q}}\left[\left|x_{s}\right|^{2}+\left|v_{s}\right|^{2}\right] \mathrm{d} s+\int_{0}^{t} \mathbb{E}^{\mathbb{Q}}\left[\nabla Q\left(x_{s}, v_{s}\right) \cdot \sigma^{(N)}\left(X_{s}^{(N)}, V_{s}^{(N)}\right) \beta_{s}\right] \mathrm{d} s
$$

in the notation $x_{t}=X_{t}^{(N)}-\tilde{X}_{t}^{(N)}$ and $v_{t}=V_{t}^{(N)}-\tilde{V}_{t}^{(N)}$. Then, since $\sigma^{(N)}$ is bounded and $\nabla Q$ is linear, we bound the last term by

$$
\varepsilon \int_{0}^{t} \mathbb{E}^{\mathbb{Q}}\left[\left|x_{s}\right|^{2}+\left|v_{s}\right|^{2}\right] \mathrm{d} s+\frac{1}{\varepsilon} \int_{0}^{t} \mathbb{E}^{\mathbb{Q}}\left|\beta_{s}\right|^{2} \mathrm{~d} s
$$

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

The transportation inequality for the law $f_{T}^{(N)}$ at time $T$ derives from the identification of the projection of the entropy:

$$
H\left(\nu \mid f_{T}^{(N)}\right)=\inf \left\{H(\mathbb{Q} \mid \mathbb{P}) ; \mathbb{Q}\left(x_{T} \in \cdot\right)=\nu(\cdot)\right\}
$$

the bound

$$
d_{2}\left(\nu, f_{T}^{(N)}\right) \leq c \mathbb{E}^{\mathbb{Q}} Q\left(\left(X_{T}^{(N)}, V_{T}^{(N)}\right)-\left(\tilde{X}_{T}^{(N)}, \tilde{V}_{T}^{(N)}\right)\right)
$$

and the preceding control.
We now turn to the deviation inequality in Theorem 1.5. First of all, if $h$ is a 1 -Lipschitz function on $\mathbb{R}^{2 d}$, then

$$
\begin{aligned}
& \frac{1}{N} \sum_{i=1}^{N} h\left(X_{T}^{i, N}, V_{T}^{i, N}\right)-\int_{\mathbb{R}^{2 d N}} \frac{1}{N} \sum_{i=1}^{N} h\left(x_{i}, v_{i}\right) \mathrm{d} f_{T}^{(N)}\left(x_{1}, \ldots, v_{N}\right)=\frac{1}{N} \sum_{i=1}^{N} h\left(X_{T}^{i, N}, V_{T}^{i, N}\right) \\
& -\int h \mathrm{~d} \mu_{\infty}+\int h \mathrm{~d} \mu_{\infty}-\int h \mathrm{~d} f_{T}+\int h \mathrm{~d} f_{T}-\int h \mathrm{~d} f_{T}^{(1, N)} \\
& \geq \frac{1}{N} \sum_{i=1}^{N} h\left(X_{T}^{i, N}, V_{T}^{i, N}\right)-\int h \mathrm{~d} \mu_{\infty}-\mathrm{d}\left(f_{T}, \mu_{\infty}\right)-\mathrm{d}\left(f_{T}, f_{T}^{(1, N)}\right)
\end{aligned}
$$

by exchangeability and by Kantorovich-Rubinstein duality (see [32] for instance). But, by Theorem 1.1 with $f_{0}=\delta_{\left(x_{0}, v_{0}\right)}$, there exist two constants $C$ and $C^{\prime}$, depending only on the coefficients of the equation, such that

$$
d\left(f_{T}, \mu_{\infty}\right) \leqslant C^{\prime} \mathrm{e}^{-C T} d\left(f_{0}, \mu_{\infty}\right)
$$

where $f_{T}$ is the solution at time $T$ of Equation (1.1) with initial datum $f_{0}=\delta_{\left(x_{0}, v_{0}\right)}$. Moreover, by Remark 1.3, there exists a constant $C^{\prime \prime}$, depending only on the equation and on $\left(x_{0}, v_{0}\right)$, such that

$$
d\left(f_{T}, f_{T}^{(1, N)}\right) \leqslant \frac{C^{\prime \prime}}{\sqrt{N}}
$$

Hence

$$
\begin{aligned}
\mathbb{P}\left[\frac{1}{N} \sum_{i=1}^{N} h\left(X_{T}^{i, N}, V_{T}^{i, N}\right)-\int_{\mathbb{R}^{2 d}} h \mathrm{~d} \mu_{\infty}\right. & \left.\geq r+D^{\prime}\left(\frac{1}{\sqrt{N}}+\mathrm{e}^{-C T}\right)\right] \\
& \leq \mathbb{P}\left[\frac{1}{N} \sum_{i=1}^{N} h\left(X_{T}^{i, N}, V_{T}^{i, N}\right)-\int_{\mathbb{R}^{2 d N}} \frac{1}{N} \sum_{i=1}^{N} h\left(x_{i}, v_{i}\right) \mathrm{d} f_{T}^{(N)}\left(x_{1}, \ldots, v_{N}\right) \geqslant r\right]
\end{aligned}
$$

where $D^{\prime}=\max \left(C^{\prime} d\left(f_{0}, \mu_{\infty}\right), C^{\prime \prime}\right)$ depends on $\left(x_{0}, v_{0}\right)$.
Now the law $f_{T}^{(N)}$ satisfies a $T_{2}$ inequality on $\mathbb{R}^{2 d N}$ with constant $D$, hence a Gaussian deviation inequality for Lipschitz functions (see [5]); moreover the map $\left(x_{1}, \ldots, v_{N}\right) \mapsto \frac{1}{N} \sum_{i=1}^{N} h\left(x_{i}, v_{i}\right)$ is $\frac{1}{\sqrt{N}}$-Lipschitz on $\mathbb{R}^{2 d N}$,
so the probability on the right-hand side is bounded by

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

for all $r \geqslant 0$. This concludes the proof of Theorem 1.5.

Acknowledgements. The authors thank the referees for a careful reading of this work and for valuable comments which helped to improve the presentation.

## 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 Mathématique de France, Paris (2000).
[2] D. Bakry and M. Émery, Diffusions hypercontractives, in Séminaire de probabilités XIX, 1983/84, Lecture Notes in Math. 1123, Springer, Berlin (1985) 177-206.
[3] D. Bakry, P. Cattiaux and A. Guillin, Rate of convergence for ergodic continuous Markov processes: Lyapunov versus Poincaré. J. Funct. Anal. 254 (2008) 727-759.
[4] D. Benedetto, E. Caglioti, J.A. Carrillo and M. Pulvirenti, A non-Maxwellian steady distribution for one-dimensional granular media. J. Statist. Phys. 91 (1998) 979-990.
[5] S.G. Bobkov and F. Götze, Exponential integrability and transportation cost related to logarithmic Sobolev inequalities. J. Funct. Anal. 163 (1999) 1-28.
[6] F. Bolley, Separability and completeness for the Wasserstein distance, in Séminaire de probabilités XLI, Lecture Notes in Math. 1934, Springer, Berlin (2008) 371-377.
[7] F. Bolley, Quantitative concentration inequalities on sample path space for mean field interaction. ESAIM: PS (to appear).
[8] F. Bolley, C. Guillin and A. Villani, Quantitative concentration inequalities for empirical measures on non-compact spaces. Probab. Theor. Relat. Fields 137 (2007) 541-593.
[9] F. Bouchut and J. Dolbeault, On long time asymptotics of the Vlasov-Fokker-Planck equation and of the Vlasov-Poisson-Fokker-Planck system with Coulombic and Newtonian potentials. Diff. Int. Eq. 8 (1995) 487-514.
[10] J.A. Carrillo and G. Toscani, Contractive probability metrics and asymptotic behavior of dissipative kinetic equations. Riv. Mat. Univ. Parma 6 (2007) 75-198.
[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) 971-1018.
[12] J.A. Carrillo, R.J. McCann and C. Villani, Contractions in the 2-Wasserstein length space and thermalization of granular media. Arch. Rat. Mech. Anal. 179 (2006) 217-263.
[13] P. Cattiaux, A. Guillin and F. Malrieu, Probabilistic approach for granular media equations in the non uniformly convex case. Probab. Theor. Relat. Fields 140 (2008) 19-40.
[14] P. Del Moral, Feynman-Kac formulae - Genealogical and interacting particle systems with applications, Probability and its Applications. Springer-Verlag, New York (2004).
[15] P. Del Moral and A. Guionnet, On the stability of measure valued processes with applications to filtering. C. R. Acad. Sci. Paris Sér. I Math. 329 (1999) 429-434.
[16] P. Del Moral and L. Miclo, Branching and interacting particle systems approximations of Feynman-Kac formulae with applications to non-linear filtering, in Séminaire de Probabilités XXXIV, Lecture Notes in Math. 1729, Springer, Berlin (2000) 1-145.
[17] P. Del Moral and E. Rio, Concentration Inequalities for Mean Field Particle Models. Preprint, http://hal. archives-ouvertes.fr/inria-00375134/en/ (2009).
[18] L. Desvillettes and C. Villani, On the trend to global equilibrium in spatially inhomogeneous entropy-dissipating systems: the linear Fokker-Planck equation. Comm. Pure Appl. Math. 54 (2001) 1-42.
[19] H. Djellout, A. Guillin and L. Wu, Transportation cost-information inequalities and applications to random dynamical systems and diffusions. Ann. Probab. 32 (2004) 2702-2732.
[20] R. Esposito, Y. Guo and R. Marra, Stability of the front under a Vlasov-Fokker-Planck dynamics. Arch. Rat. Mech. Anal. (to appear).
[21] F. Hérau, Short and long time behavior of the Fokker-Planck equation in a confining potential and applications. J. Funct. Anal. 244 (2007) 95-118.
[22] 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) 151-218.
[23] M. Ledoux, The concentration of measure phenomenon, Mathematical Surveys and Monographs 89. American Mathematical Society, Providence (2001).
[24] F. Malrieu, Logarithmic Sobolev inequalities for some nonlinear PDE's. Stochastic Process. Appl. 95 (2001) 109-132.
[25] F. Malrieu, Convergence to equilibrium for granular media equations and their Euler schemes. Ann. Appl. Probab. 13 (2003) 540-560.
[26] S. Méléard, Asymptotic behaviour of some interacting particle systems; McKean-Vlasov and Boltzmann models, in Probabilistic models for nonlinear partial differential equations (Montecatini Terme, 1995), Lecture Notes in Math. 1627, Springer, Berlin (1996) 42-95.
[27] M. Rousset, On the control of an interacting particle estimation of Schrödinger ground states. SIAM J. Math. Anal. 38 (2006) 824-844.
[28] A. Sznitman, Topics in propagation of chaos, École d'été de Probabilités de Saint-Flour XIX-1989, Lecture Notes Math. 1464, Springer, Berlin (1991) 165-251.
[29] 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) 163-198.
[30] A. Veretennikov, On ergodic measures for McKean-Vlasov stochastic equations, in Monte Carlo and quasi-Monte Carlo methods 2004, Springer, Berlin (2006) 471-486.
[31] C. Villani, Hypocoercivity, Mem. Amer. Math. Soc. 202. AMS (2009).
[32] C. Villani, Optimal transport, old and new, Grund. der Math. Wissenschaften 338. Springer-Verlag, Berlin (2009).
[33] L. Wu, Large and moderate deviations and exponential convergence for stochastic damping Hamiltonian systems. Stoch. Proc. Appl. 91 (2001) 205-238.


[^0]:    Keywords and phrases. Vlasov-Fokker-Planck equation, particular approximation, concentration inequalities, transportation inequalities.
    ${ }^{1}$ Ceremade, UMR CNRS 7534, Université Paris-Dauphine, Place du Maréchal De Lattre De Tassigny, 75775 Paris Cedex 16, France. bolley@ceremade.dauphine.fr
    ${ }^{2}$ UMR CNRS 6620, Laboratoire de Mathématiques, Université Blaise Pascal, avenue des Landais, 63177 Aubière Cedex, France. guillin@math.univ-bpclermont.fr
    ${ }^{3}$ UMR CNRS 6625, Institut de Recherche Mathématique de Rennes (IRMAR), Université de Rennes I, Campus de Beaulieu, 35042 Rennes Cedex, France. florent.malrieu@univ-rennes1.fr

