AFST

Annales de la Faculté des Sciences de Toulouse

MATHÉMATIQUES

Pierre Monmarché

A note on a Vlasov–Fokker–Planck equation with non-symmetric interaction

Tome XXXIV, nº 2 (2025), p. 243–255.

https://doi.org/10.5802/afst.1812

Publication membre du centre
Mersenne pour l'édition scientifique ouverte
http://www.centre-mersenne.org/
e-ISSN : 2258-7519

Annales de la faculté des sciences de Toulouse Volume XXXIV, n° 2, 2025
pp. 243-255

A note on a Vlasov–Fokker–Planck equation with non-symmetric interaction (*)

Pierre Monmarché (1)

Abstract. — In the recent [3], Cesbron and Herda study a Vlasov–Fokker–Planck (VFP) equation with non-symmetric interaction, introduced in physics to model the distribution of electrons in a synchrotron particle accelerator. We make four remarks in view of their work: first, it is noticed in [3] that the free energy classically considered for the (symmetric) VFP equation is *not* a Lyapunov function in the non-symmetric case, and we will show however that this is still the case for a suitable definition of the free energy (with no explicit expression in general). Second, when the interaction is sufficiently small (in $W^{1,\infty}$), it is proven in [3] that the equation has a unique stationary solution which is locally attractive; in this spirit, we will see that, when the interaction force is Lipschitz with a sufficiently small constant, the convergence is global. Third, we also briefly discuss the mean-field interacting particle system corresponding to the VFP equation which, interestingly, is a non-equilibrium Langevin process. Finally, we will see that, in the small interaction regime, a suitable (explicit) non-linear Fisher information is contracted at constant rate, similarly to the situation of Wasserstein gradient flows for convex functionals (although here the dynamics is not a gradient flow).

Résumé. — Dans le récent [3], Cesbron et Herda étudient une équation de Vlasov–Fokker–Planck (VFP) avec interaction non symétrique, introduite en physique pour modéliser la distribution d'électrons dans un synchrotron. Nous faisons ici quatre remarques au vu de leur étude : d'abord, il est noté dans [3] que l'énergie libre classiquement considérée pour l'équation de VFP (symétrique) n'est *pas* une fonction de Lyapunov dans le cas non symétrique, et nous allons voir que c'est néanmoins le cas pour une définition ad hoc de l'énergie libre (sans expression explicite en général). Ensuite, quand l'interaction est assez faible (dans $W^{1,\infty}$), il est démontré dans [3] que l'équation admet une unique solution stationnaire qui est localement

(*) Reçu le 20 novembre 2023, accepté le 7 février 2024.

Keywords: mean-field interaction, Vlasov–Fokker–Planck, free energy, Fisher information.

2020 *Mathematics Subject Classification:* 35Q83, 35Q84, 35B35, 82C22.

(1) Sorbonne Université, Laboratoire Jacques-Louis Lions, 4 place Jussieu, 75005 Paris, France — pierre.monmarche@sorbonne-universite.fr

This work has been supported by the projects CONVIVIALITY (ANR-23-CE40-0003) and SWIDIMS (ANR-20-CE40-0022) of the French National Research Agency.

Article proposé par Pascal Maillard.

attractive ; dans cet esprit, nous verrons que, pour une force d'interaction Lipschitz pour une constante suffisamment petite, la convergence est globale. Nous discuterons également brièvement le système de particules en interaction champ moyen qui, de façon notable, est un processus de Langevin hors équilibre. Finalement, nous verrons que, dans le régime de faible interaction, une information de Fisher non-linéaire bien adaptée (explicite) est contractée à taux constant, comme dans le cas des flots de gradients Wasserstein pour les fonctionnelles convexes (bien qu'ici la dynamique ne soit pas un flot de gradient).

---

1. Introduction

The Vlasov–Fokker–Planck (VFP) equation studied in [3] reads

$$\begin{cases} \partial_t f + v\cdot\partial_x f + (F_f(t,x) - x)\partial_v f = \gamma\partial_v\cdot(vf+\partial_v f) \\ F_f(t,x) = -\lambda\int_{\mathbb{R}^2}\partial_x K(x-y) f(t,y,w)\mathrm{d}y\mathrm{d}w \\ f(0,x,v) = f^{\mathrm{in}}(x,v) \end{cases} \tag{1.1}$$

where $f(t,x,v)$ stands for the density of electrons at time $t\geqslant 0$, position $x\in\mathbb{R}$ and velocity $v\in\mathbb{R}$, γ,λ are positive parameters and $K\in\mathcal{C}^2(\mathbb{R},\mathbb{R})$ is an interacting potential, with $\|\partial_x^2 K\|_\infty<\infty$. We assume that the initial density f_{in} has finite second moment. In fact, in [3], the equation involves additional physical parameters θ,α to describe the temperature and external potential intensity of the system, but we normalized them to 1, which amounts to rescale time and position. Similarly, in the present work, by contrast to [3], we assume that the mass of the initial condition is normalized to $\int_{\mathbb{R}^2} f^{\mathrm{in}} = 1$ (which amounts to a rescaling of λ, since the mass is preserved by the equation). We refer to [3] for physical motivations, well-posedness results and further references. This equation has been extensively studied in various context but, classically, motivated by the reciprocity principle of forces in classical physics, the potential K is assumed to be even, while [3] focuses on the case where it is not (motivated by models of electrons in a synchrotron particle accelerator, cf. [3]). From a mathematical perspective, this leads to interesting consequences. Classically (see e.g. [2, 3, 7]), the free energy associated to (1.1) is defined as

$$\mathcal{E}(f) = \int_{\mathbb{R}^2} f(x,v)\ln f(x,v)\mathrm{d}x\mathrm{d}v + \int_{\mathbb{R}^2}\frac{v^2+x^2}{2} f(x,v)\mathrm{d}x\mathrm{d}v \\ + \lambda\int_{\mathbb{R}^4} K(x-y) f(x,v) f(y,w)\mathrm{d}x\mathrm{d}v\mathrm{d}y\mathrm{d}w\,,$$

and, when K is even, it is known to decay along the flow of (1.1) (i.e. the free energy is a Lyapunov function for (1.1)). However, [3, Corollary 3.6] shows that, in fact, as soon as K is not even then, whatever the intensity λ of the interaction, there exist initial conditions such that $t \mapsto \mathcal{E}(f(t,\cdot))$ initially increases. Notice that $\mathcal{E}(f)$ is unchanged if we replace the potential K by the symmetrized $x \mapsto (K(x)+K(-x))/2$. This problem is reminiscent of other McKean–Vlasov models where the construction of suitable non-linear free energy functions is unclear, see e.g. [12], which gives a general motivation for the present note. Following Large Deviation theory (e.g. [6, 11]), the free energy should be obtained as the limit as $N \to \infty$ of the scaled relative entropy with respect to equilibrium for the system of N interacting particles corresponding to (1.1). The latter is the Markov process $(X,V) = (X^1,\dots,X^N,V^1,\dots V^N) \in \mathbb{R}^{2N}$ solving, for all $i \in [\![1,N]\!]$,

$$\begin{cases} \mathrm{d}X_t^i = V_t^i \mathrm{d}t \\ \mathrm{d}V_t^i = -\left(X_t^i + \gamma V_t^i + \dfrac{\lambda}{N-1}\displaystyle\sum_{j\neq i} \partial_x K\left(X_t^i - X_t^j\right)\right)\mathrm{d}t + \sqrt{2\gamma}\,\mathrm{d}B_t^i\,, \end{cases} \tag{1.2}$$

where $B = (B^1,\dots,B^N)$ is a standard N-dimensional Brownian motion. Equivalently, this reads

$$\begin{cases} \mathrm{d}X_t = V_t \mathrm{d}t \\ \mathrm{d}V_t = -(X_t + \gamma V_t + \lambda F(X_t))\mathrm{d}t + \sqrt{2\gamma}\,\mathrm{d}B_t\,, \end{cases} \tag{1.3}$$

where, for all $i \in [\![1,N]\!]$,

$$F_i(x_1,\dots,x_N) = \frac{1}{N-1}\sum_{j\,\in\,[\![1,n]\!]\setminus\{i\}} \partial_x K(x_i - x_j)\,.$$

We write $(P_t^N)_{t\geqslant 0}$ the associated Markov semi-group, so that νP_t^N is the law of (X_t,V_t) when (X_0,V_0) is distributed according to ν.

When K is even, $F = \nabla U_N$ where

$$U_N(x_1,\dots,x_n) = \frac{1}{2(N-1)}\sum_{i,j\,\in\,[\![1,N]\!]} K(x_i - x_j)\,,$$

in which case (1.3) is a classical Langevin diffusion on $\mathbb{R}^{2N}$ with invariant density proportional to $\exp(-|x|^2/2 - U_N(x) - |v|^2/2)$ (under suitable conditions on K), but in general if K is not even then F may not be a gradient, in which case the invariant measure of (1.3) doesn't have an explicit form (see the discussion in [16]). In that case, (1.3) is referred to as a *non-equilibrium* Langevin process [10]. Notice that in the non-linear equation (1.1) however the force is always the gradient of $U_f(x) = \int_{\mathbb{R}^2} K(x-y)f(t,y,w)\mathrm{d}y\mathrm{d}w$, in particular the stationary solutions of (1.1) are solutions of an explicit fixed-point problem, see [3, Theorem 1.2].

The long-time convergence of (1.1) and (1.3) (for λ small enough) is studied in [1]. More precisely, since it is based on coupling methods, which do not rely on explicit densities, gradient structures or explicit dual operators, [1] considers a general case where the forces are not necessarily of gradient forms. However, it does assume that the interaction force is odd (which, in our settings, corresponds to the case where K is even), but this has little impact in the analysis (see e.g. [15]). We briefly present below how the coupling argument of [1] still works in the non-symmetric case (in fact, this is also done in the recent [17, Corollary 4.2]). This gives a convergence to equilibrium in Wasserstein distance starting from any distribution with a finite second moment, provided $\|\partial_x^2 K\|_\infty$ is small enough, to look at in relation with the result of [3] which requires the solution to be close to stationarity (assuming however that $\|\partial_x K\|_\infty$ is small enough instead of $\|\partial_x^2 K\|_\infty$, so that the results are not directly comparable). This global weak convergence can then be improved by regularization thanks to [17] to a convergence in terms of the relative entropy, hence to a strong L^1 convergence. As a conclusion, writing respectively $\mathcal{W}_2$ and $\mathcal{H}$ the L^2 Wasserstein distance and relative entropy, based on known arguments, the following holds.

Proposition 1.1. — *Assume that*

$$\lambda\left\|\partial_x^2 K\right\|_\infty \leqslant \frac{\min\left(\gamma,\gamma^{-1}\right)}{8}\,. \tag{1.4}$$

Then there exist $C, a > 0$ such that the following hold.

(1) *For all $N \geqslant 2$, the process* (1.3) *admits a unique invariant measure μ_N and, for all initial distributions ν on $\mathbb{R}^{2N}$ and all $t \geqslant 1$,*

$$\mathcal{H}\left(\nu P_t^N \,\middle|\, \mu_N\right) + \mathcal{W}_2^2\left(\nu P_t^N, \mu_N\right) \leqslant C e^{-at}\mathcal{W}_2^2(\nu, \mu_N)\,. \tag{1.5}$$

(2) *There exists a unique stationary solution f_* of* (1.1) *and, for all solution f of* (1.1) *and all $t \geqslant 1$,*

$$\mathcal{H}(f(t,\cdot)\,|\,f_*) + \mathcal{W}_2^2(f(t,\cdot), f_*) \leqslant C e^{-at}\mathcal{W}_2^2(f(0,\cdot), f_*)\,. \tag{1.6}$$

Remark 1.2. — In view of the deterministic contraction established along the proof of Proposition 1.1, from [15, Proposition 3], under (1.4), we also get that μ_N satisfies a so-called log-Sobolev inequality with a constant uniform in N. See [5] on that topic.

Moreover, thanks to Pinsker's inequality $\|\nu-\mu\|^2_{L^1(\mathbb{R}^{2N})} \leqslant 2\mathcal{H}(\nu,\mu)$ for all ν,μ, (1.5) and (1.6) both imply the corresponding convergence in total variation norm.

In the remaining we assume that (1.3) admits an invariant measure μ_N (which is thus the case under the condition (1.4) or, for instance, using classical Lyapunov arguments [13], if $\partial_x K$ is bounded or if K is convex outside some compact interval with bounded Hessian – see [8, Lemma 10] for similar computations), and we focus on a free energy defined for any probability measure ν on $\mathbb{R}^2$ as

$$\mathcal{F}(\nu) = \liminf_{N\to\infty} \frac{1}{N}\mathcal{H}\left(\nu^{\otimes N} \mid \mu_N\right).$$

In the non-interacting case were $K = 0$, $\mu_N = \mu_1^{\otimes N}$ and we simply get $\frac{1}{N}\mathcal{H}(\nu^{\otimes N}|\mu_N) = \mathcal{H}(\nu|\mu_1) = \mathcal{F}(\nu)$. More generally, according to [7], when K is even, $\mathcal{F}(\nu) = \mathcal{E}(\nu) + C$ for some constant C (the free energy is in fact always defined up to an additive constant). This is not the case when K is not even and, by contrast to [3, Corollary 3.6], $\mathcal{F}$ is a Lyapunov functional for (1.1):

Proposition 1.3. — *For any f solution to* (1.1), *$t \mapsto \mathcal{F}(f(t,\cdot))$ is non-increasing.*

It is interesting to consider the very simple case where $K(x) = ax^2 + bx$ for some $a, b \in \mathbb{R}$, and $\lambda = 1$. Indeed, in that case, (1.3) is an equilibrium Langevin diffusions with equilibrium density proportional to $\exp(-|\mathbf{x}|^2/2 - W_N(\mathbf{x}) - |\mathbf{v}|^2/2)$ with

$$W_N(x_1,\ldots,x_N) = \frac{a}{2(N-1)}\sum_{i,j\,\in\,[\![1,N]\!]}(x_i-x_j)^2 + b\sum_{i=1}^N x_i\,.$$

In fact, the corresponding non-linear equation (1.1) can be as a Vlasov–Fokker–Planck equation with even interaction potential ax^2 and with the external potential $x^2/2 + bx$ (instead of simply $x^2/2$ as in (1.1)). From [7] we get that, in that case,

$$\begin{aligned}\mathcal{F}(f) = \int_{\mathbb{R}^2} f(x,v)\ln f(x,v)\mathrm{d}x\mathrm{d}v + \int_{\mathbb{R}^2}\left(\frac{v^2+x^2}{2} + bx\right)f(x,v)\mathrm{d}x\mathrm{d}v \\ + a\int_{\mathbb{R}^4}(x-y)^2 f(x,v)f(y,w)\mathrm{d}x\mathrm{d}v\mathrm{d}y\mathrm{d}w + C\,, \quad (1.7)\end{aligned}$$

for some constant $C \in \mathbb{R}$. In particular, when $b \neq 0$, we do not recover $\mathcal{E}$. When $a = 0$, the particles (1.2) are independent, so that $\mathcal{F}(f) = \mathcal{H}(\nu|\mu_1)$. However, due to the non-linear term in (1.7), this is not the case when $a \neq 0$. As a conclusion, when both a and b are non-zero, the free energy is neither the usual free energy $\mathcal{E}$ nor the relative entropy with respect to the stationary solution $\bar{\mu}$, but something else. In the general non-Gaussian non-symmetric case, a priori we do not have an explicit formula for $\mathcal{F}$.

Finally, if, instead of the free energy, we focus on the non-linear Fisher information which classically appears as the free energy dissipation in the symmetric interaction case (at least in the elliptic case, i.e. the granular media equation, which is the overdamped limit of (1.1) as $\gamma \to \infty$ when accelerating time by a factor γ and averaging out the velocities ; in the kinetic case, the free energy dissipation is given by a partial Fisher Information that involves only the gradient in velocity, as in the proof of Proposition 1.3), we see that we can obtain an exponential decay in the spirit of [2] for an explicit quantity of the same nature. More precisely, for a 2×2 matrix A and a probability density f, let

$$\mathcal{I}_A(f) = \int_{\mathbb{R}^2} \left| A\nabla \ln \frac{f}{\widehat{f}} \right|^2 f$$

where

$$\begin{aligned} \widehat{f}(x,y) &= \frac{e^{-\frac{1}{2}|x|^2 - \lambda K \star f(x) - \frac{1}{2}|v|^2}}{Z_f}, \\ Z_f &= \int_{\mathbb{R}^2} e^{-\frac{1}{2}|x|^2 - \lambda K \star f(x) - \frac{1}{2}|v|^2} \mathrm{d}x\mathrm{d}v \end{aligned} \tag{1.8}$$

stands for the invariant measure of the Markov generator

$$L_f = v\partial_x - \big(x + \partial_x K \star f(x) + \gamma v\big)\partial_v + \gamma \partial_v^2$$

(here we write $K \star f(x,v) = \int_{\mathbb{R}^2} K(x-y) f(y,w) \mathrm{d}y\mathrm{d}w$).

Proposition 1.4. — *There exist $c > 0$ and an invertible 2×2 matrix A (depending only on γ) such that, under the condition* (1.4)*, for any solution f of* (1.1) *and any $t \geqslant 0$,*

$$\mathcal{I}_A(f(t,\cdot)) \leqslant e^{-ct} \mathcal{I}_A(f(0,\cdot)) .$$

Using the specific form of A, c given in the proof of this result, we can get for the full non-linear Fisher information (i.e. the case $A = I$) the explicit exponential decay

$$\mathcal{I}_I(f(t,\cdot)) \leqslant 4 \exp\left(-\min\big(\gamma, \gamma^{-1}\big) \frac{t}{8} \right) \mathcal{I}_I(f(0,\cdot)) .$$

2. Proofs

Proof of Proposition 1.1. — We start by the analysis of the particle system (1.3). Denoting by ∇F the Jacobian matrix of F and $|\nabla F|$ its operator norm, we see that, for all $x, u \in \mathbb{R}^n$,

$$|\nabla F(x) u|^2 = \frac{1}{(N-1)^2} \sum_{i=1}^{n} \left| \sum_{j \neq i} \partial_x^2 K(x_i - x_j)(u_i - u_j) \right|^2$$

$$\leqslant \frac{\left\|\partial_x^2 K\right\|_\infty^2}{N-1}\sum_{i=1}^n\sum_{j\neq i}|u_i-u_j|^2$$

$$\leqslant 4\left\|\partial_x^2 K\right\|_\infty^2 \sum_{i=1}^n |u_i|^2\,,$$

namely $|\nabla F(x)| \leqslant 2\|\partial_x^2 K\|_\infty$. Let $(Z_t)_{t\geqslant 0} = (X_t,V_t)_{t\geqslant 0}$ and $(\widetilde{Z}_t)_{t\geqslant 0} = (\widetilde{X}_t,\widetilde{V}_t)_{t\geqslant 0}$ be two solutions of (1.3) with different initial conditions but driven by the same Brownian motion. For a parameter $a\in\mathbb{R}$ to be chosen later on, we set

$$P_t = X_t - \widetilde{X}_t + a\left(V_t - \widetilde{V}_t\right), \qquad Q_t = V_t - \widetilde{V}_t\,.$$

Straightforward computations give, for any $b>0$,

$$\frac{1}{2}\partial_t\left(|P_t|^2 + b|Q_t|^2\right) = -a|P_t|^2 + b(a-\gamma)|Q_t|^2 + \left(1+a^2-a\gamma-b\right)P_t\cdot Q_t - \lambda(aP+bQ)\cdot\left(F(X_t)-F(\widetilde{X}_t)\right). \quad (2.1)$$

Bounding $|F(X_t)-F(\widetilde{X}_t)| \leqslant 2\|\partial_x^2 K\|_\infty |P-aQ|$, we take for simplicity $a=\min(\gamma,\gamma^{-1})/2$ and $b = 1+a^2-a\gamma$ to get, writing $\lambda' = 2\|\partial_x^2 K\|_\infty\lambda$,

$$\begin{aligned}
&\frac{1}{2}\partial_t\left(|P_t|^2 + b|Q_t|^2\right)\\
&\leqslant -a|P_t|^2 - \frac{b\gamma}{2}|Q_t|^2 + \lambda'|aP+bQ|\,|P-aQ|\\
&\leqslant -a(1-\lambda')|P_t|^2 - \left(\frac{b\gamma}{2} - \lambda' ab\right)|Q_t|^2 + \lambda'\left(a^2+b\right)|P_t|\,|Q_t|\\
&\leqslant -\frac{3}{4}a|P_t|^2 - \frac{3b\gamma}{8}|Q_t|^2 + \frac{3}{4}\lambda'\left(|P_t|^2+|Q_t|^2\right),
\end{aligned}$$

where we used that $a=\min(\gamma,\gamma^{-1})/2\leqslant 1/2$, that $b+a^2\leqslant 3/2$ and that $\lambda'\leqslant 1/4$. Using that $b\in[1/2,5/4]$ and $\lambda'\leqslant a/4\leqslant b\gamma/4$ gives

$$\frac{1}{2}\partial_t\left(|P_t|^2+b|Q_t|^2\right) \leqslant -\frac{a}{2}|P_t|^2 - \frac{b\gamma}{16}|Q_t|^2 \leqslant -\frac{a}{8}\left(|P_t|^2+b|Q_t|^2\right).$$

Going back to the Euclidean coordinates using the equivalence between norms leads to

$$|Z_t-\widetilde{Z}_t|^2 \leqslant \frac{\max(2,b+2a^2)}{\min\left(\frac{1}{2},\frac{b}{1+2a^2}\right)}e^{-\frac{a}{4}t}|Z_0-\widetilde{Z}_0|^2 = 4e^{-\frac{a}{4}t}|Z_0-\widetilde{Z}_0|^2\,.$$

This classically implies

$$\mathcal{W}_2\left(\nu P_t^N,\widetilde{\nu}P_t^N\right) \leqslant 2e^{-\frac{a}{8}t}\mathcal{W}_2(\nu,\widetilde{\nu}) \quad (2.2)$$

for all initial distributions $\nu,\widetilde{\nu}$ on $\mathbb{R}^{2N}$, and then the existence of a unique invariant measure μ_N by the Banach fixed-point theorem and a semi-group argument. Moreover, thanks to [9, Corollary 4.7], there exists a constant C such that

$$\mathcal{H}\big(\nu P_t^N \,\big|\, \mu_N\big) \leqslant C\mathcal{W}_2^2(\nu,\mu_N)\,,$$

for all distribution ν on $\mathbb{R}^{2N}$ and all $t \geqslant 1$ (Moreover the constant C does not depend on N here, as mentionned after [8, Proposition 15]). Combining this with (2.2) concludes the proof of (1.5).

Using a synchronous coupling as in [8, Proposition 11] (which doesn't rely on the fact the force is a gradient and thus applies readily in the present case) we get that, for any initial density $f(0,\cdot) \in \mathcal{P}_2(\mathbb{R}^2)$ (the set of probability measures with finite second moment) and any $t > 0$, there exists $C_t = C_t(f(0,\cdot)) > 0$ such that , for all $N \in \mathbb{N}$,

$$\mathcal{W}_2\big(f^{\otimes N}(t,\cdot), f^{\otimes N}(0,\cdot)P_t^N\big) \leqslant C_t(f(0,\cdot))\,.$$

Hence, for any $N \in \mathbb{N}$, writing $\nu = f^{\otimes N}(0,\cdot)$ and $\widetilde{\nu} = \widetilde{f}^{\otimes N}(0,\cdot)$, using (2.2),

$$\begin{aligned}
&\mathcal{W}_2\big(f(t,\cdot),\widetilde{f}(t,\cdot)\big)\\
&= \frac{1}{\sqrt{N}}\mathcal{W}_2\Big(f^{\otimes N}(t,\cdot),\widetilde{f}^{\otimes N}(t,\cdot)\Big)\\
&\leqslant \frac{1}{\sqrt{N}}\Big(\mathcal{W}_2\big(f^{\otimes N}(t,\cdot),\nu P_t^N\big) + \mathcal{W}_2\big(\nu P_t^N,\widetilde{\nu}P_t^N\big) + \mathcal{W}_2\Big(\widetilde{\nu}P_t^N,\widetilde{f}^{\otimes N}(t,\cdot)\Big)\Big)\\
&\leqslant \frac{1}{\sqrt{N}}\Big(C_t(f(0,\cdot)) + 2e^{-\frac{a}{8}t}\mathcal{W}_2(\nu,\widetilde{\nu}) + C_t\big(\widetilde{f}(0,\cdot)\big)\Big)\\
&= 2e^{-\frac{a}{8}t}\mathcal{W}_2\Big(f(0,\cdot),\widetilde{f}(0,\cdot)\Big) + \frac{C_t(f(0,\cdot)) + C_t(\widetilde{f}(0,\cdot))}{\sqrt{N}}\,.
\end{aligned}$$

Sending $N \to \infty$ gives

$$\mathcal{W}_2\big(f(t,\cdot),\widetilde{f}(t,\cdot)\big) \leqslant 2e^{-\frac{a}{8}t}\mathcal{W}_2\big(f(0,\cdot),\widetilde{f}(0,\cdot)\big)$$

for any solutions $f,\widetilde{f}$ of (1.1) with initial condition in $\mathcal{P}_2(\mathbb{R}^2)$. Again, this implies the existence of a unique stationary solution for (1.1) by a fixed-point argument, and the proof of (1.5) is concluded by applying the Wasserstein/entropy regularization for (1.1) stated in [17, Theorem 4.1]. □

Proof of Proposition 1.3. — Fix a solution f to (1.1). For $N \in \mathbb{N}$, $t \geqslant 0$, let $L_{t,N}$ be the time-inhomogeneous Markov generator on $\mathbb{R}^{2N}$ given by

$$L_{t,N} = x\cdot\nabla_v - (x+\gamma v)\cdot\nabla_v - \sum_{i=1}^N F_f(t,x_i)\cdot\nabla_{v_i} + \gamma\Delta_v\,,$$

which corresponds to N independent copies of the solution on $\mathbb{R}^2$ of

$$\begin{cases} \mathrm{d}X_t = V_t\mathrm{d}t\\ \mathrm{d}V_t = -(X_t+\gamma V_t + \lambda F_f(t,X_t))\mathrm{d}t + \sqrt{2}\mathrm{d}B_t\,. \end{cases}$$

Since f solves (1.1), $(X_t, V_t) \sim f(t,\cdot)$ if $(X_0, V_0) \sim f(0,\cdot)$. In particular, for any nice function g on $\mathbb{R}^{2N}$,

$$\partial_t \int_{\mathbb{R}^{2N}} g f^{\otimes N}(t,\cdot) = \int_{\mathbb{R}^{2N}} L_{t,N} g f^{\otimes N}(t,\cdot)\,.$$

Similarly, let L_N be the generator associated with the interacting particle system (1.3), i.e.

$$L_N = x\cdot\nabla_v - (x+\gamma v)\cdot\nabla_v - \frac{\lambda}{N-1}\sum_{\substack{i,j=1\\ i\neq j}}^{N} \partial_x K(x_i - x_j)\cdot\nabla_{v_i} + \gamma\Delta_v\,.$$

Since μ_N is invariant for L_N, $\int L_N g \mu_N = 0$ for all suitable g. The computations in the remaining of the proof are justified when $f(0,\cdot)$ is $\mathcal{C}^\infty$ with compact support and all the derivatives of $\partial_x K$ are bounded, the general case being then obtained by an approximation step (which we omit), see for instance [4] for details). Then, using that mass is preserved along (1.1) so that $\int \partial_t f = 0$,

$$\begin{aligned}
\partial_t \int_{\mathbb{R}^{2N}} f^{\otimes N} \ln\frac{f^{\otimes N}}{\mu_N}&\\
&= \int_{\mathbb{R}^{2N}} f^{\otimes N} L_{f,N}\left(\ln\frac{f^{\otimes N}}{\mu_N}\right) + \partial_t\big(f^{\otimes N}\big)\\
&= \int_{\mathbb{R}^{2N}} f^{\otimes N}(L_{f,N} - L_N)\left(\ln\frac{f^{\otimes N}}{\mu_N}\right)\\
&\quad + \int_{\mathbb{R}^{2N}}\left[\frac{f^{\otimes N}}{\mu_N} L_N\left(\ln\frac{f^{\otimes N}}{\mu_N}\right) - L_N\left(\frac{f^{\otimes N}}{\mu_N}\ln\frac{f^{\otimes N}}{\mu_N}\right)\right]\mu_N\,.
\end{aligned}$$

Denoting by $(*)$ and $(**)$ respectively the integrals in the last two lines of this equation, we recognize for the latter the classical entropy dissipation for (1.3), for which standard computations give

$$(**) = -\int_{\mathbb{R}^{2N}}\left|\nabla_v \ln\frac{f_t^{\otimes N}}{\mu_N}\right|^2 f_t^{\otimes N}\,.$$

For the former, denoting $b = (b_1,\dots,b_N)$ with

$$b_i(x) = \frac{1}{N-1}\sum_{j\,\in\,[\![1,N]\!]\setminus\{i\}} \big(\partial_x K(x_i - x_j) - \partial_x K \star f(x_i)\big)\,,$$

we get

$$(*) = \lambda\int_{\mathbb{R}^{2N}} f^{\otimes N} b\cdot\nabla_v \ln\frac{f^{\otimes N}}{\mu_N}\,.$$

Using the Cauchy–Schwarz inequality,

$$\partial_t \int_{\mathbb{R}^{2N}} f^{\otimes N} \ln \frac{f^{\otimes N}}{\mu_N} \leqslant \frac{\lambda^2}{4} \int_{\mathbb{R}^{2N}} |b|^2 f^{\otimes N} .$$

For $i \in [\![1,N]\!]$, using the independence of particles and that $x_j \mapsto \partial_x K(x_i - x_j) - \partial_x K \star f(x_i)$ is centered under the law f,

$$\begin{aligned}\int_{\mathbb{R}^{2N}} |b_i|^2 f^{\otimes N} & \\ &= \frac{1}{(N-1)^2} \sum_{j \in [\![1,N]\!] \setminus \{i\}} \int_{\mathbb{R}^{2N}} \left(\partial_x K(x_i - x_j) - \partial_x K \star f(x_i)\right)^2 \\ &\leqslant \frac{2}{N-1} \left\|\partial_x^2 K\right\|_\infty \int_{\mathbb{R}^2} x^2 f(t,x,v) \mathrm{d}x \mathrm{d}v .\end{aligned}$$

It is easily seen that $\partial_t \int_{\mathbb{R}^2} |z|^2 f(t,z)\mathrm{d}z \leqslant C \int_{\mathbb{R}^2} |z|^2 f(t,z)\mathrm{d}z$ for some C. As a conclusion, for $t \geqslant s \geqslant 0$,

$$\int_{\mathbb{R}^{2N}} f_t^{\otimes N} \ln \frac{f_t^{\otimes N}}{\mu_N} \leqslant \int_{\mathbb{R}^{2N}} f_s^{\otimes N} \ln \frac{f_s^{\otimes N}}{\mu_N} + C_t$$

for some $C_t \geqslant 0$. Dividing by N and taking the lim inf concludes. □

Proof of Proposition 1.4. — Throughout the proof we keep computations formal, assuming sufficient regularity and integrability. See the approximation argument of [4] for a rigorous justification of similar computations.

We decompose

$$\begin{aligned}\partial_t \int_{\mathbb{R}^2} \left| A \nabla \ln \frac{f_t}{\widehat{f}_t} \right|^2 f_t & \\ &= \partial_s \int_{\mathbb{R}^2} \left| A \nabla \ln \frac{f_{t+s}}{\widehat{f}_t} \right|^2 f_{t+s} + \partial_s \int_{\mathbb{R}^2} \left| A \nabla \ln \frac{f_t}{\widehat{f}_{t+s}} \right|^2 f_t = (\star) + (\star\star) .\end{aligned}$$

Since $\widehat{f}_t$ is the invariant measure of L_{f_t}, the first part leads to classical computations (e.g. [14, Lemma 8]) which yield

$$(\star) \leqslant 2 \int_{\mathbb{R}^2} \nabla \ln \frac{f_t}{\widehat{f}_t} \cdot A^T A J \nabla \ln \frac{f_t}{\widehat{f}_t} f_t ,$$

where

$$J(x) = \begin{pmatrix} 0 & 1 + \lambda \partial_x F_{f_t}(x) \\ -1 & -\gamma \end{pmatrix} =: J_0 + \lambda \partial_x F_{f_t}(x) \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$$

is the Jacobian matrix of the drift of

$$L_{f_t}^* = -v \cdot \nabla_x + \left(x + \partial_x K \star f(x) - \gamma v\right) \cdot \nabla_v + \gamma \Delta_v ,$$

the dual of L_{f_t} in $L^2(\widehat{f}_t)$. We bound

$$(\star) \leqslant 2\int_{\mathbb{R}^2}\left[\nabla \ln \frac{f_t}{\widehat{f}_t}\cdot A^T A J_0 \nabla \ln \frac{f_t}{\widehat{f}_t} + \lambda' \left| A\nabla \ln \frac{f_t}{\widehat{f}_t}\right| \left|\nabla_v \ln \frac{f_t}{\widehat{f}_t}\right|\right] f_t$$

with $\lambda' = \lambda \|\partial_x^2 K\|_\infty \sqrt{A_{11}^2 + A_{21}^2}$.

We now deal with $(\star\star)$. Note that this question is addressed in the proof of [4, Theorem 2.1]. We give slightly different computations. First, from the explicit expression (1.8) of $\widehat{f}_t$, $\partial_t \nabla_v \ln \widehat{f}_t(x,v) = 0$ and

$$\begin{aligned}
\partial_t \nabla_x \ln \widehat{f}_t(x,v) &= -\lambda \int_{\mathbb{R}^2} \partial_x K(x-y)\partial_t f_t(y,w)\mathrm{d}y\mathrm{d}w \\
&= -\lambda \int_{\mathbb{R}^2} L_{f_t}\big((y,w)\mapsto \partial_x K(x-y)\big) f_t(y,w)\mathrm{d}y\mathrm{d}w \\
&= \lambda \int_{\mathbb{R}^2} w\partial_x^2 K(x-y) f_t(y,w)\mathrm{d}y\mathrm{d}w \\
&= \lambda \int_{\mathbb{R}^2} w\partial_x^2 K(x-y) \frac{f_t(y,w)}{\widehat{f}_t(y,w)}\widehat{f}_t(y,w)\mathrm{d}y\mathrm{d}w \\
&= -\lambda \int_{\mathbb{R}^2} \partial_x^2 K(x-y) \frac{f_t(y,w)}{\widehat{f}_t(y,w)}\nabla_w \widehat{f}_t(y,w)\mathrm{d}y\mathrm{d}w \\
&= \lambda \int_{\mathbb{R}^2} \partial_x^2 K(x-y) \nabla_w \frac{f_t(y,w)}{\widehat{f}_t(y,w)}\widehat{f}_t(y,w)\mathrm{d}y\mathrm{d}w \\
&= \lambda \int_{\mathbb{R}^2} \partial_x^2 K(x-y) \nabla_w \ln\left(\frac{f_t(y,w)}{\widehat{f}_t(y,w)}\right) f_t(y,w)\mathrm{d}y\mathrm{d}w\,.
\end{aligned}$$

As a consequence, for all $(x,v)\in\mathbb{R}^2$,

$$|\partial_t A\nabla \ln \widehat{f}_t(x,v)|^2 \leqslant (\lambda')^2 \int_{\mathbb{R}^2} \left|\nabla_w \ln \frac{f_t}{\widehat{f}_t}\right|^2 f_t\,,$$

and thus

$$\begin{aligned}
(\star\star) &\leqslant 2\lambda' \left[\int_{\mathbb{R}^2}\left|A\nabla \ln \frac{f_t}{\widehat{f}_t}\right|^2 f_t \int_{\mathbb{R}^2}\left|\nabla_w \ln \frac{f_t}{\widehat{f}_t}\right|^2 f_t\right]^{1/2} \\
&\leqslant \lambda' \int_{\mathbb{R}^2} \nabla \ln \frac{f_t}{\widehat{f}_t}\cdot \left[\theta A^T A + \theta^{-1}\begin{pmatrix}0 & 0\\ 0 & 1\end{pmatrix}\right]\nabla \ln \frac{f_t}{\widehat{f}_t} f_t
\end{aligned}$$

for any $\theta > 0$. We end up with

$$\begin{aligned}\partial_t \int_{\mathbb{R}^2} \left| A\nabla \ln \frac{f_t}{\widehat{f}_t}\right|^2 f_t \\ \leqslant 2\int_{\mathbb{R}^2} \nabla \ln \frac{f_t}{\widehat{f}_t} \cdot \left[A^T A (J_0 + \lambda' \theta I) + \lambda' \theta^{-1} \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}\right] \nabla \ln \frac{f_t}{\widehat{f}_t} f_t \,.\end{aligned}$$

As discussed in [15, Section 2.1] the matrix $A^T A$ that we should take here to get a contraction is related to the modified Euclidean distance used along the coupling in the proof of Proposition 1.1. As a consequence, considering $a = \min(\gamma, \gamma^{-1})/2$ and $b = 1 + a^2 - a\gamma$ as in the latter, we take A to be the symmetric square root of $AA^T = M^{-1}$ where

$$M = \begin{pmatrix} 1 & -a \\ -a & b + a^2 \end{pmatrix}.$$

Notice that the norm used in the proof of Proposition 1.1 is the square-root of $(x, v) \mapsto (x, -v)^T M (x, -v)$ (The velocity flip is due to the fact we work here with $L^*_{f_t}$ instead of L_{f_t}, i.e. we work with the relative density $f_t/\widehat{f}_t$ instead of the Markov semi-group associated to L_{f_t}). Indeed, revisiting the computations of the proof of Proposition 1.1, we see that, for any $u \in \mathbb{R}^2$, writing $v = M^{-1} u$,

$$\begin{aligned} u^T AA^T J_0 u &= v^T J_0 M v \\ &= v^T \begin{pmatrix} -a & b + a^2 \\ -1 + \gamma a & a - \gamma(b + a^2) \end{pmatrix} v \\ &= v^T \left[-a \begin{pmatrix} 1 & -a \\ -a & a^2 \end{pmatrix} + b(a - \gamma) \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \right] v \\ &= -a|v_1 - a v_2|^2 + b(a - \gamma)|v_2|^2 \,, \end{aligned}$$

to be compared to (2.1). Following the computations of the proof of Proposition 1.1, under (1.4), we get that, for a suitable choice of θ,

$$u^T \left[A^T A (J_0 + \lambda' \theta I) + \lambda' \theta^{-1} \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \right] u \leqslant -\frac{a}{8} v^T M v = -\frac{a}{8} |Au|^2 \,.$$

In other words, with this choice for A,

$$\partial_t \int_{\mathbb{R}^2} \left| A\nabla \ln \frac{f_t}{\widehat{f}_t}\right|^2 f_t \leqslant -\frac{a}{4} \int_{\mathbb{R}^2} \left| A\nabla \ln \frac{f_t}{\widehat{f}_t}\right|^2 f_t \,,$$

which concludes the proof of Proposition 1.4. □

Bibliography

[1] F. Bolley, A. Guillin & F. Malrieu, "Trend to equilibrium and particle approximation for a weakly selfconsistent Vlasov–Fokker–Planck equation", *ESAIM, Math. Model. Numer. Anal.* **44** (2010), no. 5, p. 867-884.

[2] J. A. Carrillo, R. J. McCann & C. Villani, "Kinetic equilibration rates for granular media and related equations: entropy dissipation and mass transportation estimates", *Rev. Mat. Iberoam.* **19** (2003), no. 3, p. 971-1018.

[3] L. Cesbron & M. Herda, "On a Vlasov–Fokker–Planck equation for stored electron beams", 2023, `https://arxiv.org/abs/2307.15964v1`.

[4] F. Chen, Y. Lin, Z. Ren & S. Wang, "Uniform-in-time propagation of chaos for kinetic mean field Langevin dynamics", 2023, `https://arxiv.org/abs/2307.02168v1`.

[5] M. G. Delgadino, R. S. Gvalani, G. A. Pavliotis & S. A. Smith, "Phase transitions, logarithmic Sobolev inequalities, and uniform-in-time propagation of chaos for weakly interacting diffusions", *Commun. Math. Phys.* **401** (2023), no. 1, p. 275-323.

[6] R. S. Ellis, "Large deviations for a general class of random vectors", *Ann. Probab.* **12** (1984), p. 1-12.

[7] A. Guillin, W. Liu, L. Wu & C. Zhang, "Uniform Poincaré and logarithmic Sobolev inequalities for mean field particle systems", *Ann. Appl. Probab.* **32** (2022), no. 3, p. 1590-1614.

[8] A. Guillin & P. Monmarché, "Uniform long-time and propagation of chaos estimates for mean field kinetic particles in non-convex landscapes", *J. Stat. Phys.* **185** (2021), no. 2, article no. 15 (20 pages).

[9] A. Guillin & F.-Y. Wang, "Degenerate Fokker–Planck equations: Bismut formula, gradient estimate and Harnack inequality", *J. Differ. Equations* **253** (2012), no. 1, p. 20-40.

[10] A. Iacobucci, S. Olla & G. Stoltz, "Convergence rates for nonequilibrium Langevin dynamics", *Ann. Math. Qué.* **43** (2019), no. 1, p. 73-98.

[11] C. Léonard, "Large deviations and law of large numbers for a mean field type interacting particle system", *Stochastic Processes Appl.* **25** (1987), p. 215-235.

[12] Y. Lu, "Two-scale gradient descent ascent dynamics finds mixed Nash equilibria of continuous games: A mean-field perspective", in *ICML'23: Proceedings of the 40th International Conference on Machine Learning*, Proceedings of Machine Learning Research, vol. 202, PMLR, 2023, p. 22790-22811.

[13] S. P. Meyn & R. L. Tweedie, *Markov chains and stochastic stability*, Communications and Control Engineering, Springer, 1993.

[14] P. Monmarché, "Generalized Γ calculus and application to interacting particles on a graph", *Potential Anal.* **50** (2019), no. 3, p. 439-466.

[15] ———, "Almost sure contraction for diffusions on $\mathbb{R}^d$. Application to generalised Langevin diffusions", *Stochastic Processes Appl.* **161** (2023), p. 316-349.

[16] P. Monmarché & M. Ramil, "Overdamped limit at stationarity for non-equilibrium Langevin diffusion", *Electron. Commun. Probab.* **27** (2022), article no. 3 (8 pages).

[17] Z. Qian, P. Ren & F.-Y. Wang, "Entropy estimate for degenerate SDEs with applications to nonlinear kinetic Fokker–Planck equations", 2023, `https://arxiv.org/abs/2308.16373v1`.