Relaxation to equilibrium for the kinetic Fokker-Planck equation

Davide Piazzoli
Cambridge Centre for Analysis
and Downing College
University of Cambridge

September 2017

This dissertation is submitted for
the degree of Master of Science.

RELAXATION TO EQUILIBRIUM FOR THE KINETIC FOKKER-PLANCK EQUATION

DAVIDE PIAZZOLI

ABSTRACT. We want to study long-time behaviour of solutions f_t of kinetic Fokker-Planck equation in $\mathbb{R}^d$, namely their convergence towards equilibrium f_∞ in the form

$$\mathrm{d}(f_t, f_\infty) \le C_1 e^{-C_2 t} \mathrm{d}(f_0, \mu)$$

for appropriate distances d and constants $C_1 \ge 1$, $C_2 > 0$.

In Section 1 we provide an introduction and motivation for the equation, together with the setting of [9] which will be useful in Section 2.

In Section 2 we will review the monograph [9], where such convergence is proved, for $h = f/\mu$, in $H^1(\mu)$ and $H_\mu + I_\mu$, that is, the sum of relative entropy and Fisher information. Here results are stated in terms of general operators $\partial_t + A^*A + B = 0$, and commutation conditions on A and B are to be imposed.

In Section 3 we shall take into consideration the work by Monmarché [5] in which such convergence is established by rephrasing some concepts in term of Γ-calculus: with respect to [9] there is no need for regularization along the semigroup since the functional taken into account is a modified $H + I$ that at initial time only takes entropy into account, and the argument turns out to be shorter. Also, the convergence rate is $e^{-Ct(1-e^{-t})^2}$ instead of $C_1 e^{-C_2 t}$. However it turns out, as in [9], that for this case it is strictly needed to have a pointwise bound on $D^2 U$, where U is the confinement potential. A drawback of this method with respect to [9] is that, in a more general setting than kinetic Fokker-Planck equation, stronger commutation assumptions are required, which imply that the diffusion matrix is basically required to be constant. On this work a specific analysis was carried out, simplifying the proof for our Fokker-Planck case and finding explicit and improved expressions for convergence constants.

The same author in [6], which is the subject of Section 4, addresses a Vlasov-Fokker-Planck equation with a potential that generalizes U and the related particle system. Chaos propagation in W_2, the 2-Wasserstein distance, is proved, namely $W_2(f_t^{(1,N)}, f_t) \le CN^{-\epsilon}$. This leads to both Wasserstein and L^1 hypocoercivity, however dependence of the right hand side from the initial data is not linear as wished.

This dissertation is the result of my own work and includes nothing which is the outcome of work done in collaboration;

it is not substantially the same as any that I have submitted, or is being concurrently submitted for a degree or diploma or other qualification at the University of Cambridge or any other University or similar institution. I further state that no substantial part of my dissertation has already been submitted, or, is being concurrently submitted for any such degree, diploma or other qualification at the University of Cambridge or any other University or similar institution;

it does not exceed the prescribed word limit for the relevant Degree Committee.

Contents

1. A general approach for Sobolev and entropical convergence

In this Section we shall present basic facts on kinetic Fokker-Planck equation, together with the main abstract tools to be used in Section 1. Here we follow [9] where we expand some classical computations which are often left as an exercise.

Here we will study the equation

$$\partial_t h + Lh = 0 \tag{1.1}$$

where h is a real function defined on $\mathbb{R}^N$ – $N \geq 1$ – and belongs to a Hilbert space $\mathcal{H}$ – in our case it will be $L^2(\mu)$, with μ equilibrium measure for (1.1).

We will consider operators of the form

$$L = A^*A + B \ , \tag{1.2}$$

where $A : \operatorname{Dom} A \subset \mathcal{H} \to \mathcal{H}^m$ and $B : \operatorname{Dom} B \subset \mathcal{H} \to \mathcal{H}$ is antisymmetric, that is, we suppose some positivity of the symmetric part of L. Let us stress that A^* is to be meant as the adjoint of A according to the product of $\mathcal{H}^m$, that is the product of $\mathcal{H}$ component by component, so that $A^* : \operatorname{Dom} A^* \subset \mathcal{H}^m \to \mathcal{H}$ thanks to the identification of $\mathcal{H}^*$ with $\mathcal{H}$. It is then easy to prove that for $g \in \mathcal{H}^m$

$$A^*g = \sum_{i=1}^m A_i^* g_i \in \mathcal{H}$$

so that (1.2) is to be read as

$$L = \sum_{i=1}^m A_i^* A_i + B \ .$$

We do not require any boundedness of A or B, as in our case they will be derivation operators.

We will indifferently denote with $e^{-tL}h$, $S_t h$ or h_t the semigroup associated to L: if h is a function $\mathbb{R}^N \to \mathbb{R}$, $e^{-tL}h$ solves $\partial_t g + Lg = 0$ with $g_0 = h$.

Remark 1. A first, simple effect of the structure of our generator L is that we can readily compute

$$\operatorname{Dom} L \cap \operatorname{Ker}(A^*A + B) = \operatorname{Ker} A \cap \operatorname{Ker} B \ .$$

In order to see this, notice first that obviously for all h antisymmetricity of B gives

$$\langle Bh, h\rangle = -\langle h, Bh\rangle = -\overline{\langle Bh, h\rangle} \ ,$$

that is, concerning the real part

$$\operatorname{Re}\langle Bh, h\rangle = 0 \ ,$$

so that

$$\operatorname{Re}\langle Lh, h\rangle = \Re\langle A^*Ah, h\rangle = \|Ah\|^2 \ ,$$

giving that $Lh = 0$ forces $Ah = Bh = 0$.

Let us now see a motivation for the study of such equations: take the stochastic differential equation in $\mathbb{R}^N$

$$dZ_t = \xi(Z_t)\,\mathrm{d}t + \Xi(Z_t)\,\mathrm{d}B_t \ ,$$

where $\xi : \mathbb{R}^N \to \mathbb{R}^N$ and $\Xi : \mathbb{R}^N \to \mathbb{R}^{N\times N}$ are smooth enough vector and matrix fields on $\mathbb{R}^N$. By considering the law $f_t(x)$ of Z_t, we want to study its evolution: pick a test function $\phi \in C_c^\infty(\mathbb{R}^N,\mathbb{R})$, then by Itô formula

$$\mathrm{d}[\phi(Z_t)] = \sum_i \partial_i\phi(Z_t)\,\mathrm{d}Z_t^i + \frac{1}{2}\sum_{i,j,k}\partial^2_{ij}\phi(Z_t)\Xi_{ik}(Z_t)\Xi_{jk}(Z_t)\,\mathrm{d}t$$

so that

$$\begin{aligned}\phi(Z_t) = &\int_0^t \Big[\sum_i \partial_i\phi(Z_s)\xi^i(Z_s) + \frac{1}{2}\sum_{i,j,k}\partial^2_{ij}\phi(Z_s)\Xi_{ik}(Z_s)\Xi_{jk}(Z_s)\Big]\,\mathrm{d}s\\ &+ \int_0^t \sum_i \partial_i\phi(Z_s)\Xi_{ik}(Z_s)\,\mathrm{d}B_s^k + \phi(Z_0)\end{aligned}$$

from which, by martingale property,

$$\begin{aligned}\int_{\mathbb{R}^N}\frac{\mathrm{d}}{\mathrm{d}t}f_t(x)\phi(x)\,\mathrm{d}x = \frac{\mathrm{d}}{\mathrm{d}t}\int_{\mathbb{R}^N}f_t(x)\phi(x)\,\mathrm{d}x = \frac{\mathrm{d}}{\mathrm{d}t}\mathbb{E}[\phi(Z_t)]\\ = \mathbb{E}\Big[\sum_i \partial_i\phi(Z_t)\xi^i(Z_t) + \frac{1}{2}\sum_{i,j,k}\partial^2_{ij}\phi(Z_t)\Xi_{ik}(Z_t)\Xi_{jk}(Z_t)\Big]\\ = \int_{\mathbb{R}^N}\Big[\sum_i \partial_i\phi(x)\xi^i(x) + \frac{1}{2}\sum_{i,j,k}\partial^2_{ij}\phi(x)\Xi_{ik}(x)\Xi_{jk}(x)\Big]f_t(x)\,\mathrm{d}x\ ,\end{aligned}$$

where in the first equality we may interchange derivation and integration by supposing some regularity of f – for instance $\partial_t f$ locally bounded in $\mathbb{R}_+\times\mathbb{R}^N$ – since $\phi \in C_c^\infty$. Here and throughout the Section we will use indifferently the notations $\int_{\mathbb{R}^p} g$ and $\int_{\mathbb{R}^p} g(x)\,\mathrm{d}x$. Since, formally,

$$\int_{\mathbb{R}^N} f_t\nabla\phi\cdot\xi = -\int_{\mathbb{R}^N}\phi\nabla\cdot(f_t\xi)\ ,$$

meaning we are done with the first part, let us focus on the second term and, by defining

$$D_{ij} := \frac{1}{2}\sum_k \Xi_{ik}\Xi_{jk} = \frac{1}{2}(\Xi\Xi^T)_{ij}\ ,$$

there holds, still formally and up to regularity issues of f_t and Ξ,

$$\begin{aligned}\int_{\mathbb{R}^N}\sum_{i,j}\partial^2_{ij}\phi D_{ij}f_t = \int_{\mathbb{R}^N}\sum_{i,j}\partial^2_{ij}(f_tD_{ij})\phi =\\ = \int_{\mathbb{R}^N}\sum_i \phi\partial_i\big((D\nabla f_t)_i\big) + \sum_{i,j}\phi\partial_i\big(f_t\partial_jD_{ij}\big)\\ = \int_{\mathbb{R}^N}\phi\nabla\cdot\big(D\nabla f_t + f_t\nabla(D\mathbf{1})\big)\ ,\end{aligned}$$

which means that, by calling

$$d_i(x) := \sum_j \partial_j D_{ij}(x)\ ,$$

we have reached the weak formulation for the evolution of f_t,

$$\partial_t f_t = \nabla\cdot\big(D\nabla f_t + f_t(d-\xi)\big)\ . \tag{1.3}$$

We will refer to this as *natural* equation since it arises from a physically natural setting.

Suppose now that there exists a stationary solution f_∞ for (1.3), smooth and positive, possibly of infinite mass. Then the ratio $h_t := \frac{f_t}{f_\infty}$ satisfies

(1.4)
$$\begin{aligned}
\frac{\mathrm{d}h_t}{\mathrm{d}t} &= \frac{1}{f_\infty}\nabla\cdot\big(D\nabla f_t + f_t(d-\xi)\big) = \frac{1}{f_\infty}\nabla\cdot\big(f_\infty D\nabla h_t + f_\infty h_t(d-\xi) + h_t D\nabla f_\infty\big)\\
&= \nabla\cdot(D\nabla h_t) + \frac{1}{f_\infty}\nabla f_\infty D\nabla h_t + \frac{h_t}{f_\infty}\nabla\cdot\big(f_\infty(d-\xi)\big) + (d-\xi)\cdot\nabla h_t\\
&\quad + \frac{1}{f_\infty}\nabla\cdot(D\nabla f_\infty)h_t + \frac{1}{f_\infty}\nabla h_t D\nabla f_\infty\\
&= \nabla\cdot(D\nabla h_t) + \left(\frac{2D\nabla f_\infty}{f_\infty} + (d-\xi)\right)\cdot\nabla h_t
\end{aligned}$$

where we used that f_∞ has the right hand side of (1.3) vanishing.
We wish to write this under the form of (1.1), so let us choose the Hilbert setting of $\mathcal{H} = L^2(f_\infty)$. Let us also consider $\Xi\nabla$ and compute its adjoint in $\mathcal{H}$: pick smooth $g : \mathbb{R}^m \to \mathbb{R}^m$ and $h : \mathbb{R}^m \to \mathbb{R}$; then

$$\begin{aligned}
\langle(\Xi\nabla)^* g, h\rangle_{\mathcal{H}} = \langle g, \Xi\nabla h\rangle_{\mathcal{H}^m} &= \int_{\mathbb{R}^N} \nabla h \Xi^T g f_\infty = -\int_{\mathbb{R}^N} h\nabla\cdot(\Xi^T g f_\infty)\\
&= -\int_{\mathbb{R}^N} h\Big[\nabla\cdot(\Xi^T g) + \frac{\nabla f_\infty}{f_\infty}\Xi^T g\Big] f_\infty \ ,
\end{aligned}$$

that is,

$$(\Xi\nabla)^* g = -\nabla\cdot(\Xi^T g) - \frac{\nabla f_\infty}{f_\infty}\Xi^T g$$

and

$$(\Xi\nabla)^*(\Xi\nabla)h = -\nabla\cdot(2D\nabla h) - \frac{\nabla f_\infty}{f_\infty}2D\nabla h \ ,$$

giving that we shall choose

$$A = \frac{1}{\sqrt{2}}\Xi\nabla$$

which gives

$$\frac{dh_t}{dt} = -A^* A h_t + \left(\frac{D\nabla f_\infty}{f_\infty} + (d-\xi)\right)\cdot\nabla h_t \ .$$

We just need to prove that the last term is antisymmetric: by calling it

$$B = b\cdot\nabla \ ,$$

we shall prove that for all h

$$\langle Bh, h\rangle = 0$$

which is equivalent to antisymmetricity of B by writing

$$\langle Bh, g\rangle + \langle h, Bg\rangle = \langle B(g+h), g+h\rangle - \langle Bh, h\rangle - \langle Bg, g\rangle \ .$$

Indeed

$$\begin{aligned}
\langle Bh, h\rangle = \int_{\mathbb{R}^N}(b\cdot\nabla h)hf_\infty &= \frac{1}{2}\int_{\mathbb{R}^N} b\cdot\nabla(h^2)f_\infty = -\frac{1}{2}\int_{\mathbb{R}^N} h^2\nabla\cdot(bf_\infty)\\
&= -\frac{1}{2}\int_{\mathbb{R}^N} h^2\nabla\cdot\big(D\nabla f_\infty + (d-\xi)f_\infty\big) = 0
\end{aligned}$$

for all h.
This means that h_t solves an equation as (1.1), with B antisymmetric derivation field, in the Hilbertian setting of $L^2(f_\infty)$. In view of this fact, we shall call (1.4) *formal* equation.

Here arises our favourite example, linear kinetic Fokker-Planck equation – actually, the particular case from which we decide to build a general theory: let the physical dimension be $d \geq 1$ and take two space and velocity processes X_t and V_t in $\mathbb{R}^d$ defined by

$$\begin{cases} \mathrm{d}X_t = V_t \,\mathrm{d}t \\ \mathrm{d}V_t = -\nabla_x U(X_t)\,\mathrm{d}t - \gamma V_t \,\mathrm{d}t + \sigma \,\mathrm{d}B_t \end{cases} ,$$

where U is a C^1 potential $\mathbb{R}^d \to \mathbb{R}$ – it corresponds to space confinement, and from time to time we shall require growth conditions at infinity in order to prevent particles from being scattered away – B_t is a d-dimensional Brownian motion , γ and σ are positive constants. Then it is easy to see that the induced natural equation (1.3) reads

$$\partial_t f + v \cdot \nabla_x f - \nabla_x U(x) \cdot \nabla_v f = \frac{\sigma^2}{2} \Delta_v f + \gamma \nabla_v \cdot (vf) \ ,$$

that is,

$$D(x,v) = \frac{\sigma^2}{2} \begin{bmatrix} 0_x & 0_v \\ 0_x & \mathrm{Id}_v \end{bmatrix}$$

and

$$\xi(x,v) = \begin{bmatrix} v \\ -\nabla_x U(x) - \gamma v \end{bmatrix}$$

and $f_\infty(x,v) = e^{-\frac{2\gamma}{\sigma^2}\left(U(x) + \frac{|v|^2}{2}\right)}$, up to a normalizing constant. The matrix D being constant, the vector $d(x,v)$ from (1.4) is null, allowing us to refer to d as the dimension only, with no notation ambiguity.

Let us us now provide the direct computation of the classical change of the equation from *natural* to *formal*. This is just an easy computation, often left as an exercise, that we wish to provide.

Proposition 1. *Suppose that $f = f(t,x,v)$ satisfies*

$$\partial_t f + v \cdot \nabla_x f - \nabla_x U \cdot \nabla_v f = \frac{\sigma^2}{2} \Delta_v f + \gamma \nabla_v \cdot (vf) \ , \quad t \geq 0, \ (x,v) \in \mathbb{R}^d \times \mathbb{R}^d$$

for some γ, $\sigma > 0$, which admits as equilibrium, up to a normalizing constant,

$$f_\infty(x,v) = e^{-\frac{2\gamma}{\sigma^2}\left(U(x) + \frac{|v|^2}{2}\right)} .$$

Then the density with respect to equilibrium $\frac{f}{f_\infty}$ satisfies

$$\partial_t h + v \cdot \nabla_x h - \nabla_x U \cdot \nabla_v h = \frac{\sigma^2}{2} \Delta_v h - \gamma v \cdot \nabla_v h$$

Proof. Let us compute all differential items when applied to $\frac{f}{f_\infty}$. First

$$\partial_t \left(\frac{f}{f_\infty} \right) = -\frac{v \cdot \nabla_x f}{f_\infty} + \frac{\nabla_x U \cdot \nabla_v f}{f_\infty} + \frac{\sigma^2}{2} \frac{\Delta_v f}{f_\infty} + \gamma \frac{\nabla_v \cdot (vf)}{f_\infty} .$$

Next, we may compute, for all differential operators involved G, $G(f/f_\infty)$ and look for $G(f)/f_\infty$ on the right hand side. Indeed

$$\nabla_x\left(\frac{f}{f_\infty}\right) = \nabla_x(e^{\frac{2\gamma}{\sigma^2}(U(x)+\frac{|v|^2}{2})})f + \frac{\nabla_x f}{f_\infty} = \frac{2\gamma}{\sigma^2}e^{\frac{2\gamma}{\sigma^2}(U(x)+\frac{|v|^2}{2})}f\nabla_x U + \frac{\nabla_x f}{f_\infty}$$

so that

$$-\frac{v\cdot\nabla_x f}{f_\infty} = \frac{2\gamma}{\sigma^2}v\cdot\nabla_x U\frac{f}{f_\infty} - \nabla_x\left(\frac{f}{f_\infty}\right).$$

In the same fashion

$$\nabla_v\left(\frac{f}{f_\infty}\right) = \nabla_v(e^{\frac{2\gamma}{\sigma^2}(U(x)+\frac{|v|^2}{2})})f + \frac{\nabla_v f}{f_\infty} = \frac{2\gamma}{\sigma^2}e^{\frac{2\gamma}{\sigma^2}(U(x)+\frac{|v|^2}{2})}fv + \frac{\nabla_v f}{f_\infty}$$

leads to

$$\frac{\nabla_x U\cdot\nabla_v f}{f_\infty} = \nabla_x U\cdot\nabla_v\left(\frac{f}{f_\infty}\right) - \frac{2\gamma}{\sigma^2}\frac{f}{f_\infty}\nabla_x U\cdot v\ .$$

Summing the two expressions gives

$$-\frac{v\cdot\nabla_x f}{f_\infty} + \frac{\nabla_x U\cdot\nabla_v f}{f_\infty} = -v\cdot\nabla_x\left(\frac{f}{f_\infty}\right) + \nabla_x U\cdot\nabla_v\left(\frac{f}{f_\infty}\right)\quad .$$

Now let us take into consideration the Laplacian: by recalling that, for all smooth f and g, the formula $\Delta(fg) = g\Delta f + 2\nabla f\cdot\nabla g + f\Delta g$ holds,

$$\begin{aligned}\frac{\sigma^2}{2}\Delta_v\left(\frac{f}{f_\infty}\right) &= \frac{\sigma^2}{2}\frac{\Delta_v f}{f_\infty} + \sigma^2\nabla_v f\cdot\nabla_v(e^{\frac{2\gamma}{\sigma^2}(U(x)+\frac{|v|^2}{2})}) + \frac{\sigma^2}{2}f\Delta_v\left(\frac{1}{f_\infty}\right)\\ &= \frac{\sigma^2}{2}\frac{\Delta_v f}{f_\infty} + 2\gamma\frac{\nabla_v f}{f_\infty}\cdot v + \frac{\sigma^2}{2}f\Delta_v\left(\frac{1}{f_\infty}\right)\ .\end{aligned}$$

By writing explicitly the last term

$$\frac{\sigma^2}{2}f\Delta_v\left(\frac{1}{f_\infty}\right) = \gamma f\nabla_v\cdot(ve^{\frac{2\gamma}{\sigma^2}(U(x)+\frac{|v|^2}{2})}) = \gamma f\nabla_v\cdot\left(\frac{v}{f_\infty}\right)$$

which gives

$$\frac{\sigma^2}{2}\frac{\Delta_v f}{f_\infty} = \frac{\sigma^2}{2}\Delta_v\left(\frac{f}{f_\infty}\right) - 2\gamma\frac{\nabla_v f}{f_\infty}\cdot v - \gamma f\nabla_v\cdot\left(\frac{v}{f_\infty}\right)$$

and, summing up everything we have computed,

$$\begin{aligned}&\partial_t\left(\frac{f}{f_\infty}\right) + v\cdot\nabla_x\left(\frac{f}{f_\infty}\right) - \nabla_x U\cdot\nabla_v\left(\frac{f}{f_\infty}\right)\\ &\qquad = \frac{\sigma^2}{2}\Delta_v\left(\frac{f}{f_\infty}\right) - 2\gamma\frac{v\cdot\nabla_v f}{f_\infty} - \gamma f\nabla_v\cdot\left(\frac{v}{f_\infty}\right) + \gamma\frac{\nabla_v\cdot(vf)}{f_\infty}\ .\end{aligned}$$

It is now easy to find the coefficient of γ as a function of $\frac{f}{f_\infty}$ by applying several times Leibniz rule:

$$\begin{aligned}\frac{\nabla_v\cdot(vf)}{f_\infty} - \frac{2v\cdot\nabla_v f}{f_\infty} - f\nabla_v\cdot\left(\frac{v}{f_\infty}\right) &= \frac{\nabla_v\cdot(vf)}{f_\infty} - \frac{v\cdot\nabla_v f}{f_\infty} - \nabla_v\cdot\left(v\frac{f}{f_\infty}\right)\\ &= -\nabla_v\left(\frac{1}{f_\infty}\right)\cdot vf - \frac{v\cdot\nabla_v f}{f_\infty} = -v\cdot\nabla_v\left(\frac{f}{f_\infty}\right)\end{aligned}$$

which gives

$$\partial_t\left(\frac{f}{f_\infty}\right) + v\cdot\nabla_x\left(\frac{f}{f_\infty}\right) - \nabla_x U\cdot\nabla_v\left(\frac{f}{f_\infty}\right) = \frac{\sigma^2}{2}\Delta_v\left(\frac{f}{f_\infty}\right) - \gamma v\cdot\nabla_v\left(\frac{f}{f_\infty}\right)\ .\ \square$$

2. General approach to hypocoercivity: Villani's method

In this Section we are going to briefly expose some features of the monograph [9], with a stress on kinetic Fokker-Planck equation, which is the main application therein - not the only one, though. We have tried to be to some extent complementary to its author, by sketching some parts which are there in deep detail and, conversely, expanding some computations which are glossed over.

In Subsection 2.1 we are going to provide results in a general setting for Sobolev H^1 norm – or better, an abstract version of it – with Theorem 2 and entropy – or rather the sum of entropy and Fisher information – in the way more detailed Theorem 3. In Subsection 2.2 we are going to apply such results to kinetic Fokker-Planck equation. We will also provide some regularity results for solutions of the equation, which, in addition to being of their own interest, allow for convergence results starting from a way less regular data, at the price of a worse dependence from data on the right hand side. In Subsection 2.3 we provide the outline of the theory developed to tackle a nonlinear variant of kinetic Fokker-Planck equation.

2.1. **Linear setting: L^2 and entropic convergence.**

Now let us turn to the abstract study of our operator: take a Hilbert space $\tilde{\mathcal{H}} \hookrightarrow (\operatorname{Ker} L)^{\perp} \cap \operatorname{Dom} L$ densely, where orthogonality is to be meant according to the structure of $\mathcal{H}$ – typically we will consider H^1 embedded in 0-mean functions. We will say that the generator $L : \mathcal{H} \to \mathcal{H}$ is λ-*coercive* if for all $h \in \operatorname{Dom} L \cap (\operatorname{Ker} L)^{\perp}$

$$\operatorname{Re}\langle Lh, h\rangle_{\tilde{\mathcal{H}}} \geq \lambda \left\|h\right\|^2_{\tilde{\mathcal{H}}} \ ,$$

that is, if the symmetric part of L, when restricted, admits a spectral gap in $\tilde{\mathcal{H}}$. An equivalent formulation, closer to our purposes, comes from noticing that the previous definition is nothing but Gronwall inequality on the squared norm along L flow: for every $h_0 \in \tilde{\mathcal{H}}$ and h_t satisfying (1.1), coercivity is equivalent to

$$\left\|h_t\right\|_{\tilde{\mathcal{H}}} \leq e^{-\lambda t} \left\|h_0\right\|_{\tilde{\mathcal{H}}} \ .$$

Concerning this last formulation, what is always true is that

$$\frac{\mathrm{d}}{\mathrm{d}t} \left\|h_t\right\|^2_{\mathcal{H}} = -2 \left\|Ah_t\right\|^2_{\mathcal{H}} \ ,$$

which gives that $\left\|\cdot\right\|_{\mathcal{H}}$ does not increase along the evolution, and therefore that the semigroup operator S_t is $\mathcal{H}$-non expansive for all t. This will not be useful for coercivity though, since it will never hold

$$A^* A \gtrsim \mathrm{Id} \ ,$$

as we will have $A = \nabla_v$. This is a symptom of the usual difficulty that we encounter when dealing with relaxation and kinetic equations: the full Laplacian that helps in spatial Fokker-Planck equation is now present in half of its form only.

However our hope to have a coercive L is going to be frustrated, so we will say that L is *hypocoercive* if

$$\left\|h_t\right\|_{\tilde{\mathcal{H}}} \leq C e^{-\lambda t} \left\|h_0\right\|_{\tilde{\mathcal{H}}} \ , \quad h_0 \in \tilde{\mathcal{H}},$$

for some constants $C \geq 1$ and $\lambda > 0$, and for an appropriate and nontrivial $\tilde{\mathcal{H}}$. This is what we will always aim to prove, in an appropriate setting.

We shall see that a important role is played by commutation. Indeed, suppose A and B commute, in that B commutes with all of A_i; then so do their exponentials, and $e^{-tL} = e^{-tA^*A} e^{-tB}$. Since e^{-tB} is norm-preserving, hypocoercivity of L

is equivalent to hypocoercivity of A^*A, however in our case this does not occur. Indeed, by defining

$$[A,B] := AB - B^{\otimes m}A : \mathcal{H} \to \mathcal{H}^m \ ,$$

where by $B^{\otimes m} : \mathcal{H}^m \to \mathcal{H}^m$ we denote B applied component by component, in our Fokker-Planck setting we shall have $[A,B] = \nabla_x$. Since in this case it will be non-zero, the strategy will be to impose magnitude bounds on several commutators in order to make its effect negligible.

Now, define recursively the operators

$$C_0 := A$$

and

$$C_{k+1} := [C_k, B] : \mathcal{H} \to \mathcal{H}^m$$

to be meant as above. Then

- $\operatorname{Ker} L \subset \bigcap_{k\geq 0} \operatorname{Ker} C_k$: if $Lh = 0$, Remark 1 gives that $Ah = 0 \in \mathcal{H}^m$ and $Bh = 0 \in \mathcal{H}$, so that $C_0 h = 0$. Inductively, $C_{k+1}h = C_k Bh - B^{\otimes m} C_k h = 0$
- $\bigcap_{k\geq 0} \operatorname{Ker} C_k$ is invariant under e^{-tL}.

Since we wish to have $\operatorname{Ker} L = \bigcap_{k\geq 0} \operatorname{Ker} C_k$, or better $\bigcap_{k=0}^K \operatorname{Ker} C_k$, we will ask for $\sum_{k=0}^K C_k^* C_k \neq 0$ on $\{L \neq 0\}$. More strongly, we will ask for $\sum_{k=0}^K C_k^* C_k$ coercive on $(\operatorname{Ker} L)^\perp$.

Now we can state the first big result, which corresponds to Theorem 18 in [9]. Its proof will be only sketched by synthetically pointing at main ideas, since a close analogue will be proved in greater generality for the entropic case. We will write

$$\|h\|^2_{\mathcal{H}^1} := \|h\|^2_{\mathcal{H}} + \sum_{k=0}^K \|C_k h\|^2_{\mathcal{H}}$$

and we will say that an operator T is α-bounded with respect to $\{S_j\}_j$ if for all h it holds $\|Th\| \leq \alpha \sum_j \|S_j h\|$. Also, we will write $\mathcal{K} = \operatorname{Ker} L$ and write $[A, A^*]$ for the matrix $\{[A_j, A_k^*]\}_{jk}$.

Theorem 2 ($\mathcal{H}^1$ convergence)**.** *Here we will deal with $K = 1$. We will suppose that:*

- *A commutes with C, A^* commutes with C and A_i commutes with A_j for all i and j*
- *$[A, A^*]$ is α-bounded with respect to I and A, and $[B, C]$ is β-bounded by A, A^2, C and AC*
- *$A^*A + C^*C$ is coercive in $\mathcal{H}$-norm on $(\mathcal{K})^\perp$.*

Then there exists a scalar product $((\cdot,\cdot))$ on $(\mathcal{K}^\perp)_{\mathcal{H}^1}$ of the form $\langle\cdot,\cdot\rangle + a\,\|A\cdot\|^2 + b\langle A\cdot, C\cdot\rangle + c\,\|C\cdot\|^2$, where products and norms are to be meant in $\mathcal{H}$, such that L is $((\cdot,\cdot))$-coercive on $(\mathcal{K}^\perp)_{\mathcal{H}^1}$. This inner product is equivalent to $\mathcal{H}^1$, so that L is $\mathcal{H}^1$-hypocoercive.

Proof. Take $h \in \mathcal{H}^1 \cap \mathcal{K}^\perp$, write down $((h, Lh))$ and separate it between various terms of $((\cdot,\cdot))$; commute and estimate various terms, either by domination assumptions, or by antisymmetricity or simply with Cauchy-Schwarz; at this point, we reached an expression of the form

$$((h, Lh)) \geq \big(\, \|Ah\|\,, \left\|A^2 h\right\|, \|Ch\|\,, \|CAh\| \,\big) \cdot M \cdot \big(\, \|Ah\|\,, \left\|A^2 h\right\|, \|Ch\|\,, \|CAh\| \,\big) \ ,$$

where M is a 4 by 4 matrix, whose entries are just numbers depending on constants and parameters a, b and c which are still to be chosen. We should just prove that the symmetric part of M is positive definite, so that we will yield, by denoting with c a generic constant,

$$((h, Lh)) \geq c(\|Ah\|^2 + \|Ch\|^2) \geq \frac{c}{2}(\|Ah\|^2 + \|Ch\|^2) + \frac{c\lambda}{2}\|h\|^2 \geq c((h))$$

thanks to coercivity of commutators. In order to get positivity of the matrix, and to make sure that our product is positive, we need an upper bound on the linear growth and a lower bound on the geometric growth of a, b and c, which is done by an arithmetical argument. □

This argument may be generalized, to $K > 1$ and by writing

$$[C_k, B] = Z_{k+1}C_{k+1} + R_{k+1} \ ,$$

where Z_k are fixed fields $\mathcal{H}^m \to \mathcal{H}^m$ bounded from above and below, and R_k are remainder terms with a magnitude condition.

After dealing with a generalization of Sobolev convergence – this comparison will be clearer in Subsection 2.2 – let us focus on the entropy case, by giving definitions in a more general framework than the one we need; recall that, given μ measure on $\mathbb{R}^N$ and a μ-measurable function $h : \mathbb{R}^N \to [0, +\infty)$, we define the relative entropy of $h\mu$ with respect to μ as

$$H_\mu(h\mu) := \int_{\mathbb{R}^N} h \log h \, \mathrm{d}\mu \ ,$$

and the Fisher information of $h\mu$ with respect to μ as

$$I_\mu(h\mu) := \int_{\mathbb{R}^N} \frac{|\nabla h|^2}{h} \, \mathrm{d}\mu \ .$$

We will write $H_\mu(\nu) = I_\mu(\nu) = \infty$ if ν is not absolutely continuous with respect to μ, and we will even write $I_\mu(h) \equiv I_\mu(h\mu)$ in case there is no possibility of confusion. Here we abandon the Hilbertian setting of $L^2(\mu)$, and we will just consider functions on $\mathbb{R}^N$; we will consider a reference measure ν rapidly decreasing at infinity and with sufficiently well behaved semigroup, in that it maps a set of smooth and fast decaying enough functions into itself. Also, we will consider operators of the form $A = \Xi \cdot \nabla$ and $B = b \cdot \nabla$, where Ξ and b are a matrix and a vector field, and we will refer to such operators as *derivations*. Here we will deal with relative boundedness of derivations in terms of pointwise boundedness of the representing fields when viewed as matrices. In other words, if an array of derivations is represented by the matrix field $\{\xi_{ij}\}_{ij}$, *pointwise boundedness* with respect to A means that there exists $c > 0$ such that

$$w_i \xi_{ij}(x) w_j \leq c \ w_i \Xi_{ij}(x) w_j \ , \quad x \in \mathbb{R}^N \ , \quad w \in \mathbb{R}^m \ .$$

We are now ready for the following result, which amounts to Theorem 28 in [9]; the proof here shown also contains the important Lemma 32 in the monograph, together with the expansion of therein glossed over computations.

Theorem 3 (Entropy convergence)**.** *Let us set ourselves in the situation where $C_0 = A$ and, for $0 \le k \le K$, $[C_k, B] = Z_{k+1}C_{k+1} + R_{k+1}$, where $Z_k : \mathbb{R}^N \to \mathbb{R}^N$, $1 \le k \le K$, are fields such that for real constants λ_k and Λ_k*

$$0 < \lambda_k |x|^2 \le Z_k(x) \cdot x \le \Lambda_k |x|^2 \ ,$$

and C_k, R_k are derivations. We will suppose:

(i) $[C_k, A]$ pointwise bounded with respect to A with constant c_1
(ii) $[C_k, A]^*$ pointwise bounded with respect to I and A with constant c_2
(iii) $[C_k, A^*]$ pointwise bounded with respect to I and $\{C_j\}_{j=0}^k$ with constant c_3
(iv) R_{k+1} pointwise bounded with respect to $\{C_j\}_{j=0}^k$ with constant c_R
(v) The matrices of coefficients C_k satisfy

$$\sum_{k=0}^{K} C_k^*(x)C_k(x) \ge cI_N \ , \quad x \in \mathbb{R}^N$$

(vi) μ satisfies a Logarithmic Sobolev Inequality, that is,

$$H_\mu(\nu) \le \frac{1}{2C} I_\mu(\nu)$$

for all ν such that $\int \, d\nu = \int \, d\mu$.

Then we have entropic hypocoercivity, in the weaker sense that for h_t satisfying $(\partial_t + L)h = 0$

$$(H_\mu + I_\mu)(h_t) \le Ce^{-\lambda t}(H_\mu + I_\mu)(h_0)$$

for a constant λ which depends on all previous constants.

Proof. We define the target functional as

$$\mathcal{E}(h) \equiv \int_{\mathbb{R}^N} h \log h \ \mathrm{d}\mu + \sum_{k=0}^{K} \left(a_k \int_{\mathbb{R}^N} \frac{|C_k h|^2}{h} \, \mathrm{d}\mu + 2b_k \int_{\mathbb{R}^N} \frac{C_k h \cdot C_{k+1} h}{h} \, \mathrm{d}\mu \right) \ ,$$

that is, the sum of relative entropy and Fisher information twisted via the tridiagonal quadratic form $M = M(x)$, which is a $N \times N$ matrix field of the following fashion: by identifying the derivation operator C_k with its $m \times N$ matrix field of coefficients,

$$\langle M(x)\xi, \xi\rangle_N = \sum_{k=0}^{K} a_k \, |C_k(x)\xi|_m^2 + 2\sum_{k=0}^{K-1} b_k \langle C_k(x)\xi, C_{k+1}(x)\xi\rangle_m \ , \quad \xi \in \mathbb{R}^N \ ,$$

where m is the amplitude of derivation operators, that is, $A = (A_i)_{i=1}^m$ We will sometimes be denoting inner products by $\langle v, w\rangle_p$ just to make clear, if ambiguous, that $v, w \in \mathbb{R}^p$.

From time to time we will be imposing conditions on a_k and b_k that will eventually be chosen accordingly. For instance, for $\mathcal{E}$ to be equivalent to $H + I$, we need uniform positivity of M with respect to I, for which we ask for

$$b_k^2 \le \delta a_k a_{k+1} \ ,$$

for some $\delta > 0$ to be fixed later, since it gives

$$\langle M(x)\xi, \xi\rangle_N \ge (1 - 2\delta) \sum_{k=0}^{K} a_k \, \|C_k(x)\xi\|^2 \ , \quad \xi \in \mathbb{R}^N \ .$$

We can now study evolution of $\mathcal{E}$ in itself:

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}H_\mu(h\mu) &= -\int_{\mathbb{R}^N}(\log h+1)(A^*Ah+Bh)\,\mathrm{d}\mu\\ &= -\int_{\mathbb{R}^N}A(\log h+1)\cdot Ah-\int_{\mathbb{R}^N}B(h\log h)\,\mathrm{d}\mu=-\int_{\mathbb{R}^N}\frac{|Ah|^2}{h}\,\mathrm{d}\mu\ .\end{aligned}$$

It is easy to notice that the effect of B is not a pure chance, but it holds with any $F(h)$ as long as the generator of the evolution is antisymmetric:

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\int_{\mathbb{R}^N}F(h)\,\mathrm{d}\mu &= -\int_{\mathbb{R}^N}F'(h)Bh\,\mathrm{d}\mu\\ &= -\int_{\mathbb{R}^N}B(F(h))\,\mathrm{d}\mu=\int_{\mathbb{R}^N}F(h)B^*(1)\,\mathrm{d}\mu=0\ ,\end{aligned}$$

thanks to chain rule and to B^* being a derivation.

Now let us turn to the dissipation of the modified Fisher information: first we shall deal with the general term in the sum by relabelling it as

$$\int_{\mathbb{R}^N}\frac{Ch\cdot\tilde{C}h}{h}\,\mathrm{d}\mu$$

and noticing that we can write it, for instance, as

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\int_{\mathbb{R}^N}\frac{Ch\cdot\tilde{C}h}{h}\,\mathrm{d}\mu &= 4\int C\sqrt{h}\cdot\tilde{C}\partial_t\sqrt{h}\,\mathrm{d}\mu+4\int_{\mathbb{R}^N}\tilde{C}\sqrt{h}\cdot C\partial_t\sqrt{h}\ \mathrm{d}\mu\\ &= -2\int_{\mathbb{R}^N}C\sqrt{h}\cdot\tilde{C}\Big(\frac{Lh}{\sqrt{h}}\Big)\,\mathrm{d}\mu-2\int_{\mathbb{R}^N}\tilde{C}\sqrt{h}\cdot C\Big(\frac{Lh}{\sqrt{h}}\Big)\,\mathrm{d}\mu\end{aligned}$$

highlighting thus L-additivity so that we can analyse separately the effect of A^*A and B, by writing

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\sum_{k=0}^{K}\Big(a_k\int_{\mathbb{R}^N}\frac{|C_kh|^2}{h}\,\mathrm{d}\mu+2b_k\int_{\mathbb{R}^N}\frac{C_kh\cdot C_{k+1}h}{h}\,\mathrm{d}\mu\Big)\\ =\sum_{k=0}^{K}a_k((I)_k^A+(I)_k^B)+b_k((II)_k^A+(II)_k^B)\end{aligned}\tag{2.1}$$

Let us study thoroughly the form of each of the four terms on the right hand side, by first dealing with B: consider h solving $(\partial_t+B)h=0$. Then, since it is a first-order derivation, we also have

$$\partial_t\sqrt{h}+B\sqrt{h}=\frac{1}{2\sqrt{h}}(\partial_t+B)h=0\ ,$$

so that it is useful to rewrite the ratio as the better-behaving quadratic formulation

$$
\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}\int_{\mathbb{R}^N}\frac{Ch\cdot\tilde{C}h}{h}\,\mathrm{d}\mu =& 4\frac{\mathrm{d}}{\mathrm{d}t}\int_{\mathbb{R}^N} C\sqrt{h}\cdot\tilde{C}\sqrt{h}\ \mathrm{d}\mu \\
=& -4\int_{\mathbb{R}^N} CB\sqrt{h}\cdot\tilde{C}\sqrt{h}\ \mathrm{d}\mu - 4\int_{\mathbb{R}^N}\tilde{C}B\sqrt{h}\cdot C\sqrt{h}\ \mathrm{d}\mu \\
=& -4\int_{\mathbb{R}^N}[C,B]\sqrt{h}\cdot\tilde{C}\sqrt{h}\ \mathrm{d}\mu - 4\int_{\mathbb{R}^N} BC\sqrt{h}\cdot\tilde{C}\sqrt{h}\ \mathrm{d}\mu \\
& -4\int_{\mathbb{R}^N}[\tilde{C},B]\sqrt{h}\cdot C\sqrt{h}\ \mathrm{d}\mu - 4\int_{\mathbb{R}^N} B\tilde{C}\sqrt{h}\cdot C\sqrt{h}\ \mathrm{d}\mu \\
=& -4\int_{\mathbb{R}^N}[C,B]\sqrt{h}\cdot\tilde{C}\sqrt{h}\ \mathrm{d}\mu - 4\int_{\mathbb{R}^N}[\tilde{C},B]\sqrt{h}\cdot C\sqrt{h}\ \mathrm{d}\mu \\
=& -\int_{\mathbb{R}^N}\frac{[C,B]h\cdot\tilde{C}h}{h}\ \mathrm{d}\mu - \int_{\mathbb{R}^N}\frac{[\tilde{C},B]h\cdot Ch}{h}\ \mathrm{d}\mu\ .
\end{aligned}
$$

We can now turn to A^*A, taking h such that $(\partial_t + A^*A)h = 0$.

$$
\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}\int_{\mathbb{R}^N}\frac{Ch\cdot\tilde{C}h}{h}\,\mathrm{d}\mu =& \frac{\mathrm{d}}{\mathrm{d}t}\int_{\mathbb{R}^N} hC\log h\cdot\tilde{C}\log h\ \mathrm{d}\mu = \frac{\mathrm{d}}{\mathrm{d}t}\int_{\mathbb{R}^N} fC\log h\cdot\tilde{C}\log h \\
=& \underbrace{\int_{\mathbb{R}^N}\partial_t f\ C\log h\cdot\tilde{C}\log h}_{(1)} \\
& + \underbrace{\int_{\mathbb{R}^N} fC\Big(\frac{\partial_t f}{f}\Big)\cdot\tilde{C}\log h}_{(2)} + \underbrace{\int_{\mathbb{R}^N} f\tilde{C}\Big(\frac{\partial_t f}{f}\Big)\cdot C\log h}_{(2')}\ .
\end{aligned}
$$

We shall deal separately with each of the terms, anyway $f_t = h_t e^{-E}$ evolves according to

$$
\partial_t f_t = \nabla\cdot(D\nabla f_t + f_t(\sqrt{2}d - \xi)) = \nabla\cdot(D(\nabla f_t + f_t\nabla E))
$$

since

$$
0 = -b = \frac{D\nabla e^{-E}}{e^{-E}} + (\sqrt{2}d - \xi)\ ,
$$

where $B = b\cdot\nabla$ and $D = \Xi^T\Xi$. Then

$$
\begin{aligned}
(1) =& -\int_{\mathbb{R}^N}(D(\nabla f + f\nabla E))\cdot\nabla(C\log h\cdot\tilde{C}\log h) \\
=& -\int_{\mathbb{R}^N}\Xi(e^{-E}\nabla h - h\nabla E e^{-E} + f\nabla E)\cdot\Xi\nabla(C\log h\cdot\tilde{C}\log h) \\
=& -\int_{\mathbb{R}^N}\Xi(f\nabla\log h)\cdot\Xi\nabla(C\log h\cdot\tilde{C}\log h) \\
=& -\int_{\mathbb{R}^N} fA\log h\cdot A(C\log h\cdot\tilde{C}\log h)
\end{aligned}
$$

and, since we now know the equation satisfied by f_t, let us highlight f

$$
\begin{aligned}
(2) &= \int_{\mathbb{R}^N} fC\Big(\frac{\nabla\cdot(D(\nabla f + f\nabla E))}{f}\Big)\cdot \tilde{C}\log h \\
&= \int_{\mathbb{R}^N} fC\Big(\frac{\nabla\cdot(fD\nabla\log h)}{f}\Big)\cdot \tilde{C}\log h \\
&= \int_{\mathbb{R}^N} f\Big\langle C(\nabla\cdot(D\nabla\log h)), \tilde{C}\log h\Big\rangle_m \\
&\quad + \int_{\mathbb{R}^N} f\Big\langle C\langle\nabla\log f, D\nabla\log h\rangle_N, \tilde{C}\log h\Big\rangle_m \\
&= \int_{\mathbb{R}^N} f\Big\langle C(\nabla\cdot(D\nabla\log h) - \langle D\nabla E, \nabla\log h\rangle_N), \tilde{C}\log h\Big\rangle_m \\
&\quad + \int_{\mathbb{R}^N} f\Big\langle C\langle\nabla\log h, D\nabla\log h\rangle_N, \tilde{C}\log h\Big\rangle_m \\
&= -\int_{\mathbb{R}^N} f\langle C(A^*A\log h), \tilde{C}\log h\rangle_m + \int_{\mathbb{R}^N} f\langle C\left|A\log h\right|^2, \tilde{C}\log h\rangle_m \\
&= \underbrace{-\int_{\mathbb{R}^N} f\langle [C,A^*]A\log h, \tilde{C}\log h\rangle_m}_{(2.1)} \underbrace{-\int_{\mathbb{R}^N} f\langle A^*CA\log h, \tilde{C}\log h\rangle_m}_{(2.2)} \\
&\quad + \underbrace{2\int_{\mathbb{R}^N} f\Big\langle (CA\log h)(A\log h), \tilde{C}\log h\Big\rangle_m}_{(2.3)}
\end{aligned}
$$

where we also used that $A^*A = -\nabla\cdot(D\nabla) + D\nabla E\cdot\nabla$ and wished to make dimensions clear: for instance, by calling $a \to b$ an operator on functions $\mathbb{R}^N \to \mathbb{R}^a$ taking values in functions $\mathbb{R}^N \to \mathbb{R}^b$, $[C, A^*]$ is to be meant as $m \to m$ and A^*CA is $1 \to m$. Now let us look at the second term:

$$
\begin{aligned}
(2.2) &= -\int_{\mathbb{R}^N} \langle A^*CA\log h, \tilde{C}h\rangle_m \,\mathrm{d}\mu \\
&= -\int_{\mathbb{R}^N} \langle CA\log h, [A,\tilde{C}]h\rangle_{m\times m}\,\mathrm{d}\mu - \int_{\mathbb{R}^N} \langle CA\log h, \tilde{C}Ah\rangle_{m\times m}\,\mathrm{d}\mu \\
&= -\int_{\mathbb{R}^N} \langle CA\log h, [A,\tilde{C}]h\rangle_{m\times m}\,\mathrm{d}\mu - \int_{\mathbb{R}^N} \langle CA\log h, \tilde{C}(hA\log h)\rangle_{m\times m}\,\mathrm{d}\mu \\
&= \underbrace{-\int_{\mathbb{R}^N} f\langle CA\log h, [A,\tilde{C}]\log h\rangle_{m\times m}}_{2.2.1} \\
&\quad \underbrace{-\int_{\mathbb{R}^N} f\langle CA\log h, \tilde{C}A\log h\rangle_{m\times m}}_{2.2.2} \underbrace{-\int_{\mathbb{R}^N} f\Big\langle (CA\log h)(A\log h), \tilde{C}\log h\Big\rangle_m}_{2.2.3}.
\end{aligned}
$$

Here (2.2.3) may be summed with (2.3), giving that, together with the symmetric equivalent,

$$
\begin{aligned}
&(1)+(2.3)+(2.2.3)+(2'.3)+(2'.2.3) = -\int_{\mathbb{R}^N} fA\log h\cdot A(C\log h\cdot\tilde{C}\log h)\\
&\quad+\int_{\mathbb{R}^N} f\big\langle (CA\log h)(A\log h),\tilde{C}\log h\big\rangle_m+\int_{\mathbb{R}^N} f\big\langle (\tilde{C}A\log h)(A\log h),C\log h\big\rangle_m\\
&=\int_{\mathbb{R}^N} f\tilde{C}_j\log h\ [C_j,A_i]\log h\ A_i\log h+\int_{\mathbb{R}^N} fC_j\log h\ [\tilde{C}_j,A_i]\log h\ A_i\log h\\
&=\int_{\mathbb{R}^N} \tilde{C}_j\log h\ [C_j,A_i]h\ A_i\log h\ \ \mathrm{d}\mu+\int_{\mathbb{R}^N} C_j\log h\ [\tilde{C}_j,A_i]h\ A_i\log h\ \ \mathrm{d}\mu\\
&=\int_{\mathbb{R}^N} f[C_j,A_i]^*(\tilde{C}_j\log h\cdot A_i\log h)+\int_{\mathbb{R}^N} f[\tilde{C}_j,A_i]^*(C_j\log h\cdot A_i\log h)
\end{aligned}
$$

where the sum is implicit for $1\le i,j\le m$. All in all, for $\partial_t h+A^*Ah=0$, one can perform symmetric computations and yield

$$
\begin{aligned}
&\frac{\mathrm{d}}{\mathrm{d}t}\int_{\mathbb{R}^N}\frac{Ch\cdot\tilde{C}h}{h}\,\mathrm{d}\mu=[(2.1)+(2'.1)]+[(2.2.1)+(2'.2.1)]+[(2.2.2)+(2'.2.2)]\\
&\quad+[(1)+(2.3)+(2.2.3)+(2'.3)+(2'.2.3)]\\
&=-\int_{\mathbb{R}^N} f\langle [C,A^*]A\log h,\tilde{C}\log h\rangle_m-\int_{\mathbb{R}^N} f\langle [\tilde{C},A^*]A\log h,C\log h\rangle_m\\
&\quad-\int_{\mathbb{R}^N} f\langle CA\log h,[A,\tilde{C}]\log h\rangle_{m\times m}-\int_{\mathbb{R}^N} f\langle \tilde{C}A\log h,[A,C]\log h\rangle_{m\times m}\\
&\quad-2\int_{\mathbb{R}^N} f\langle CA\log h,\tilde{C}A\log h\rangle_{m\times m}\\
&\quad+\int_{\mathbb{R}^N} f[C_j,A_i]^*(\tilde{C}_j\log h\cdot A_i\log h)+\int_{\mathbb{R}^N} f[\tilde{C}_j,A_i]^*(C_j\log h\cdot A_i\log h)\ .
\end{aligned}
$$

We can now compute $(I)_k^A$, $(II)_k^A$, $(I)_k^B$, $(II)_k^B$ and use commutators' decomposition

$$
\begin{aligned}
(I)_k^B&=-2\int_{\mathbb{R}^N}\frac{[C_k,B]h\cdot C_kh}{h}\ \mathrm{d}\mu\\
&=-2\int_{\mathbb{R}^N}\frac{Z_{k+1}C_{k+1}h\cdot C_kh}{h}\ \mathrm{d}\mu-2\int_{\mathbb{R}^N}\frac{R_{k+1}h\cdot C_kh}{h}\ \mathrm{d}\mu\\
&\le 2\Lambda_{k+1}\sqrt{\int_{\mathbb{R}^N}\frac{|C_{k+1}h|^2}{h}\,\mathrm{d}\mu}\sqrt{\int_{\mathbb{R}^N}\frac{|C_kh|^2}{h}\,\mathrm{d}\mu}\\
&\quad+2c_R\sqrt{\sum_{j=0}^k\int_{\mathbb{R}^N}\frac{|C_jh|^2}{h}\,\mathrm{d}\mu}\sqrt{\int_{\mathbb{R}^N}\frac{|C_kh|^2}{h}\,\mathrm{d}\mu}
\end{aligned}
$$

where we used that $Z_{k+1}(x) \cdot x \geq \Lambda_{k+1} |x|^2$ for all $x \in \mathbb{R}^N$ and enforced condition (iv) on R_{k+1}. The mixed term in B gives

$$
\begin{aligned}
(II)_k^B = & - \int_{\mathbb{R}^N} \frac{Z_{k+1} C_{k+1} h \cdot C_{k+1} h}{h} \, \mathrm{d}\mu - \int_{\mathbb{R}^N} \frac{R_{k+1} h \cdot C_{k+1} h}{h} \, \mathrm{d}\mu \\
& - \int_{\mathbb{R}^N} \frac{Z_{k+2} C_{k+2} h \cdot C_k h}{h} \, \mathrm{d}\mu - \int_{\mathbb{R}^N} \frac{R_{k+2} h \cdot C_k h}{h} \, \mathrm{d}\mu \\
\leq & - \lambda_{k+1} \int_{\mathbb{R}^N} \frac{|C_{k+1} h|^2}{h} \, \mathrm{d}\mu + c_R \sqrt{\sum_{j=0}^{k} \int_{\mathbb{R}^N} \frac{|C_j h|^2}{h} \mathrm{d}\mu} \sqrt{\int_{\mathbb{R}^N} \frac{|C_{k+1} h|^2}{h} \mathrm{d}\mu} \\
& + \Lambda_{k+2} \sqrt{\int_{\mathbb{R}^N} \frac{|C_{k+2} h|^2}{h} \mathrm{d}\mu} \sqrt{\int_{\mathbb{R}^N} \frac{|C_k h|^2}{h} \mathrm{d}\mu} \\
& + c_R \sqrt{\sum_{j=0}^{k+1} \int_{\mathbb{R}^N} \frac{|C_j h|^2}{h} \mathrm{d}\mu} \sqrt{\int_{\mathbb{R}^N} \frac{|C_k h|^2}{h} \mathrm{d}\mu}
\end{aligned}
$$

again thanks to bounds on Z_{k+1}. All terms we have gained for the moment are positive and we wish to estimate them via the only negative one. Let us now work on A, by first focusing on the last term appearing in $(I)_k^A$ and $(II)_k^A$. By enforcing condition (ii) on $[C_k, A]^*$ and (i) on $[C_k, A]$,

$$
\begin{aligned}
& \int_{\mathbb{R}^N} f [C_k, A]^* \cdot (C_k \log h \otimes A \log h) \\
& \qquad \leq c_2 \int_{\mathbb{R}^N} f \, |C_k \log h \otimes A \log h| + c_2 \int_{\mathbb{R}^N} f \, |A(C_k \log h \otimes A \log h)| \\
& \qquad \leq c_2 \int_{\mathbb{R}^N} f \, |C_k \log h \otimes A \log h| + c_2 \int_{\mathbb{R}^N} f \, \left| A^2 \log h \; C_k \log h \right| \\
& \qquad \quad + c_2 \int_{\mathbb{R}^N} f \, |C_k A \log h \; A \log h| + c_2 \int_{\mathbb{R}^N} f \, |[A, C_k] \log h \; A \log h| \\
& \qquad \leq c_2 \int_{\mathbb{R}^N} f \, |C_k \log h \otimes A \log h| + c_2 \int_{\mathbb{R}^N} f \, \left| A^2 \log h \; C_k \log h \right| \\
& \qquad \quad + c_2 \int_{\mathbb{R}^N} f \, |C_k A \log h \; A \log h| + c_1 c_2 \int_{\mathbb{R}^N} f \, |A \log h|^2
\end{aligned}
$$

thanks to which

$$
\begin{aligned}
(I)_k^A \leq & -2\int_{\mathbb{R}^N} f\langle [C_k, A^*]A\log h, C_k\log h\rangle_m \\
& -2\int_{\mathbb{R}^N} f\langle C_kA\log h, [A, C_k]\log h\rangle_{m\times m} \\
& -2\int_{\mathbb{R}^N} f\,|C_kA\log h|^2_{m\times m} \\
& +2c_2\int_{\mathbb{R}^N} f\,|C_k\log h\otimes A\log h| + 2c_2\int_{\mathbb{R}^N} f\left|A^2\log h\ C_k\log h\right| \\
& +2c_2\int_{\mathbb{R}^N} f\,|C_kA\log h\ A\log h| + 2c_1c_2\int_{\mathbb{R}^N} f\,|A\log h|^2 \\
\leq & -2\int_{\mathbb{R}^N} f\,|C_kA\log h|^2 \\
& +2c_3\sqrt{\int_{\mathbb{R}^N} f\,|C_k\log h|^2}\sqrt{\sum_{j=-1}^{k}\int_{\mathbb{R}^N} f\,|C_jA\log h|^2} \\
& +2(c_1+c_2)\sqrt{\int_{\mathbb{R}^N} f\,|C_kA\log h|^2}\sqrt{\int_{\mathbb{R}^N} f\,|A\log h|^2} \\
& +2c_2\sqrt{\int_{\mathbb{R}^N} f\,|C_k\log h|^2}\sqrt{\int_{\mathbb{R}^N} f\,|A\log h|^2} \\
& +2c_2\sqrt{\int_{\mathbb{R}^N} f\,|C_k\log h|^2}\sqrt{\int_{\mathbb{R}^N} f\,|A^2\log h|^2} \\
& +2c_1c_2\int_{\mathbb{R}^N} f\,|A\log h|^2
\end{aligned}
$$

by enforcing conditions (i) and (iii) and compactly denoting $I = C_{-1}$. Similarly, without repeating analogous calculations, one obtains that

$$
\begin{aligned}
(II)_k^A \leq & \,2\sqrt{\int_{\mathbb{R}^N} f\,|C_kA\log h|^2}\sqrt{\int_{\mathbb{R}^N} f\,|C_{k+1}A\log h|^2} \\
& +c_3\sqrt{\int_{\mathbb{R}^N} f\,|C_k\log h|^2}\sqrt{\sum_{j=-1}^{k+1}\int_{\mathbb{R}^N} f\,|C_j\log h|^2} \\
& +c_3\sqrt{\int_{\mathbb{R}^N} f\,|C_{k+1}\log h|^2}\sqrt{\sum_{j=-1}^{k}\int_{\mathbb{R}^N} f\,|C_j\log h|^2}
\end{aligned}
$$

$$
\begin{aligned}
&+ c_1\sqrt{\int_{\mathbb{R}^N} f\,|A\log h|^2}\left(\sqrt{\int_{\mathbb{R}^N} f\,|C_kA\log h|^2}+\sqrt{\int_{\mathbb{R}^N} f\,|C_{k+1}A\log h|^2}\right)\\
&+ c_2\sqrt{\int_{\mathbb{R}^N} f\,|A\log h|^2}\left(\sqrt{\int_{\mathbb{R}^N} f\,|C_k\log h|^2}+\sqrt{\int_{\mathbb{R}^N} f\,|C_{k+1}\log h|^2}\right)\\
&+ c_2\sqrt{\int_{\mathbb{R}^N} f\,|A^2\log h|^2}\left(\sqrt{\int_{\mathbb{R}^N} f\,|C_k\log h|^2}+\sqrt{\int_{\mathbb{R}^N} f\,|C_{k+1}\log h|^2}\right)\\
&+ c_2\sqrt{\int_{\mathbb{R}^N} f\,|A\log h|^2}\left(\sqrt{\int_{\mathbb{R}^N} f\,|C_kA\log h|^2}+\sqrt{\int_{\mathbb{R}^N} f\,|C_{k+1}A\log h|^2}\right)\\
&+ c_1c_2\int_{\mathbb{R}^N}|A\log h|^2\ .
\end{aligned}
$$

This enormous amount of positive terms we have collected should not frighten, since they shall all be dominated in the end; indeed for the sake of brevity we shall no more write the whole right hand side of (2.1), which shall contain as the only negative terms

$$
-2b_k\lambda_{k+1}\int_{\mathbb{R}^N} f\,|C_{k+1}\log h|^2 \quad\text{and}\quad -\,2a_k\int_{\mathbb{R}^N} f\,|C_kA\log h|^2\ .
$$

Here we are always supposing $0\le k\le K-1$, the case of $k=K$ is even easier and should be treated separately. Our strategy will be the following: we will regroup all positive terms in two groups: Group 2 will be made by terms appearing from the estimate of

$$
\begin{aligned}
&2\int_{\mathbb{R}^N} f[C_k,A]^*\cdot(C_k\log h\otimes A\log h)\\
&+\int_{\mathbb{R}^N} f[C_k,A]^*\cdot(C_{k+1}\log h\otimes A\log h)+\int_{\mathbb{R}^N} f[C_{k+1},A]^*\cdot(C_k\log h\otimes A\log h)
\end{aligned}
$$

while Group 1 is the remainder. We shall first focus on Group 1 aiming to prove Group 2 to be negligible with respect to the former: since via Cauchy-Schwarz we have managed to yield everywhere a multiplication times a

$$
\sqrt{\int_{\mathbb{R}^N} f\,|C_j\log h|^2}\quad\text{or a}\quad\sqrt{\int_{\mathbb{R}^N} f\,|C_jA\log h|^2}\ ,
$$

we shall perform Young inequalities on each single term in Group 1, so that we earn a

$$
\varepsilon a_j\int_{\mathbb{R}^N} f\,|C_jA\log h|^2\quad\text{or a}\quad\varepsilon b_{j-1}\int_{\mathbb{R}^N} f\,|C_j\log h|^2
$$

in order to, eventually, compare the sum with

$$
\varepsilon\left(\sum_{j=0}^{K}\left(a_j\int_{\mathbb{R}^N} f\,|C_jA\log h|^2+b_j\int_{\mathbb{R}^N} f\,|C_{j+1}\log h|^2\right)+\int_{\mathbb{R}^N} f\,|A\log h|^2\right)\tag{2.2}
$$

for ε small enough that cancellation with negative terms occurs. Clearly, as a price we shall have the second terms given by Young inequality with conjugated coefficient as an expression of a_j and b_j: for instance, in the fourth line of $(II)^A_k$

we shall estimate

$$
\begin{aligned}
& b_k c_1 \sqrt{\int_{\mathbb{R}^N} f\,|A\log h|^2}\sqrt{\int_{\mathbb{R}^N} f\,|C_k A\log h|^2} \\
& \quad \leq \varepsilon a_k \int_{\mathbb{R}^N} f\,|C_k A\log h|^2 + \frac{b_k^2}{a_k}\frac{c_1^2}{4\varepsilon}\int_{\mathbb{R}^N} f\,|A\log h|^2 \ .
\end{aligned}
$$

Since we may suppose ε to be fixed, we shall try to have the product of a_j and b_j small enough for the second term to be negligible with respect to the first one. Indeed we have already defined the quadratic form with a_j and b_j depending on a parameter $\delta > 0$, and we will ask

$$
\frac{b_k^2}{a_k} \in o_{\delta\to 0}(a_k) \equiv \frac{b_k^2}{a_k} \leq_\delta a_k \ .
$$

Clearly it will be a $\delta(\varepsilon)$, which will make sure that for fixed ε we will have yielded a bound by

$$
\varepsilon a_k \left(\int_{\mathbb{R}^N} f\,|C_k A\log h|^2 + o_\delta(1) \right)
$$

which collected among the whole of Group 1 yields an amount of requirements on $\{a_j\}$ and $\{b_j\}$ to be analysed in the end. Concerning Group 2, we will only be concerned with proving they become negligible as $\delta \to 0$. We shall aim to compare each term in Group 2 with a subset of terms in sum (2.2), without ε. Since comparison will be via Young inequality, we shall compare coefficient in Group 2 with the square root of the product of target coefficients in (2.2). For instance, for the penultimate term in $(II)_k^A$ we shall aim to a comparison with $\int_{\mathbb{R}^N} f\,|A\log h|^2 + a_{k+1}\int_{\mathbb{R}^N} f\,|C_{k+1}A\log h|^2$. To this scope, let us impose

$$
b_k \leq_\delta \sqrt{a_{k+1}}
$$

thanks to which

$$
\begin{aligned}
& b_k \sqrt{\int_{\mathbb{R}^N} f\,|A\log h|^2}\sqrt{\int_{\mathbb{R}^N} f\,|C_{k+1} A\log h|^2} \\
& \quad \leq_\delta \sqrt{a_{k+1}}\sqrt{\int_{\mathbb{R}^N} f\,|A\log h|^2}\sqrt{\int_{\mathbb{R}^N} f\,|C_{k+1} A\log h|^2} \\
& \quad \leq \int_{\mathbb{R}^N} f\,|A\log h|^2 + a_{k+1}\int_{\mathbb{R}^N} f\,|C_{k+1}A\log h|^2 \ ,
\end{aligned}
$$

which means that, as long some conditions on $\{a_j\}$ and $\{b_j\}$ are enforced, we can summarize everything as

$$
\begin{aligned}
&\frac{\mathrm{d}}{\mathrm{d}t}\Big[H_\mu(h\mu) + \sum_{k=0}^{K}\Big(a_k\int_{\mathbb{R}^N}\frac{|C_k h|^2}{h}\,\mathrm{d}\mu + 2b_k\int_{\mathbb{R}^N}\frac{C_k h\cdot C_{k+1}h}{h}\,\mathrm{d}\mu\Big)\Big]\\
&\leq -\int_{\mathbb{R}^N}\frac{|Ah|^2}{h}\,\mathrm{d}\mu - \sum_{k=0}^{K}\Big(2a_k\int_{\mathbb{R}^N} f\,|C_k A\log h|^2 + b_k\lambda_{k+1}\int_{\mathbb{R}^N} f\,|C_{k+1}\log h|^2\Big)\\
&\quad + (\varepsilon + o_\delta(1))\\
&\quad\cdot\Big[\sum_{k=0}^{K}\Big(a_k\int_{\mathbb{R}^N} f\,|C_k A\log h|^2 + b_k\int_{\mathbb{R}^N} f\,|C_{k+1}\log h|^2\Big) + \int_{\mathbb{R}^N}\frac{|Ah|^2}{h}\,\mathrm{d}\mu\Big]\\
&\lesssim -\int_{\mathbb{R}^N}\frac{|Ah|^2}{h}\,\mathrm{d}\mu - \sum_{k=0}^{K}\Big(a_k\int_{\mathbb{R}^N} f\,|C_k A\log h|^2 + b_k\int_{\mathbb{R}^N} f\,|C_{k+1}\log h|^2\Big)\\
&\leq -\int_{\mathbb{R}^N} f\,|C_0\log h|^2 - \sum_{k=1}^{K} b_{k-1}\int_{\mathbb{R}^N} f\,|C_k\log h|^2
\end{aligned}
$$

if ε and δ are small enough, thanks to which the hidden constant is arbitrarily close to 1. We are now only wish to impose enough conditions to close the inequality with $H_\mu + I_{\mu,M}$. For this we shall use (v) and (vi) on half of the last term in the inequality and simply recover $I_{\mu,M}$, giving

$$
\begin{aligned}
&-\int_{\mathbb{R}^N} f\,|C_0\log h|^2 - \sum_{k=1}^{K} b_{k-1}\int_{\mathbb{R}^N} f\,|C_k\log h|^2\\
&\lesssim -\frac{1}{2}\sum_{k=0}^{K}\Big(a_k\int_{\mathbb{R}^N} f\,|C_k\log h|^2 + 2b_k\int_{\mathbb{R}^N} f\ C_k\log h\cdot C_{k+1}\log h\Big)\\
&\quad -\frac{1}{2}\sum_{k=0}^{K}\int_{\mathbb{R}^N} f\,|C_k\log h|^2\\
&\leq -\frac{1}{2}\big(I_{\mu,M}(h) + cI_\mu(h)\big)\\
&\leq -\frac{1}{2}\big(I_{\mu,M}(h) + 2cC_{LSI}H_\mu(h)\big)
\end{aligned}
$$

where the hidden constant only depends on coefficients $\{a_k\}$ and $\{b_k\}$.

The only thing we need to have the theorem proved is enforcing conditions on $\{a_k\}_{k=0}^{K}$ and $\{b_k\}_{k=0}^{K-1}$ themselves and proving their compatibility for existence: such conditions were all derived from asking remainders in Group 1 were δ-negligible with respect to principal terms, and from asking Group 2 to be δ-negligible with respect to Group 1. All in all, these conditions are satisfied if we set

$$a_k := u_{2k+1} \quad \text{and} \quad b_k := u_{2k+2}$$

where $\{u_k\}_{k=0}^{2K+2}$ is a sequence of positive numbers satisfying

$$\begin{cases} u_{k+1} & \leq \delta u_k \\ u_k^2 & \leq \delta u_{k-1}u_{k+1}\ . \end{cases}$$

This is done by asking ratios to tend to 0 and setting $u_k = \eta^{m_k}$, thus transforming the previous conditions into

$$\begin{cases} m_{k+1} - m_k & > 0 \\ 2m_k - (m_{k+1} - m_{k-1}) & > 0 \end{cases}$$

which fulfil our requirements for $\eta \to 0$, since the sequence is finite. In order to prove the existence of m_k, it suffices to notice that one may choose $m_0 = 0$, $m_1 = 1$ and increments such that $0 < m_{k+1} - m_k < m_k - m_{k-1}$. This completes the proof. $\square$

2.2. **Application to kinetic Fokker-Planck equation.** Let us see how the two previous Theorems apply to our Fokker-Planck setting. We shall consider the formal equation

$$\partial_t h + v \cdot \nabla_x h - \nabla_x U(x) \cdot \nabla_v h = \Delta_v h - v \cdot \nabla_v h \ ,$$

on the Hilbert space $L^2(e^{-U(x) - \frac{|v|^2}{2}})$. We wish to apply Theorem 2; here it is easy to prove that $\operatorname{Ker} L$ is made by constant functions (either by integrating Lh against h or by reminding that $\operatorname{Ker} L = \operatorname{Ker} A \cap \operatorname{Ker} B$) so that

$$\mathcal{K}^\perp = \left\{ h : \int_{\mathbb{R}^d_x \times \mathbb{R}^d_v} h \,\mathrm{d}\mu = 0 \right\} \ .$$

Here and in the following, we shall not write the domain of integrals whenever it is meant to be $\mathbb{R}^d_x \times \mathbb{R}^d_v$. Also, $C = [A, B] = \nabla_x$, so that $\mathcal{H}^1 = H^1(\,\mathrm{d}\mu)$ restricted to zero-mean functions. Concerning hypotheses of the theorem to be verified,

- $A = \nabla_v$ and $A^* = -\nabla_v \cdot + v \cdot$ commute with C, since they only interact with the v variable; also, easily $A_i = \partial_{v_i}$ commutes with A_j.
- $[A, A^*] = I : \mathcal{H}^m \to \mathcal{H}^m$, therefore 1-bounded with respect to I and A
- $[B, C] = \nabla^2_x U(x) \cdot \nabla_v$ should be bounded with respect to ∇_v, ∇^2_v, ∇_x, ∇^2_{xv} in $L^2(\mu)$. Indeed, it is enough if $D^2_x U(x) \cdot$ is bounded by I and ∇_x in $L^2_x(e^{-U(x)} dx)$. In other words, we wish $\left|D^2 U\right| \cdot$ to be bounded from $H^1(e^{-U})$ to $L^2(e^{-U})$.

  For this to happen, it is enough to ask for

$$\left|D^2 U\right| \leq c(1 + |\nabla U|) \ , \tag{2.3}$$

  which roughly corresponds to exponential growth at most. To prove that (2.3) implies our desired boundedness, let us prove the same for $|\nabla U|^2$ instead of $\left|D^2 U\right|^2$, which implies our claim by (2.3): pick g smooth and fast-decaying enough and, by reminding that

$$\begin{aligned} (\Delta U)^2 &\leq \sum_{j=1}^d (\partial^2_{jj} U)^2 + \frac{1}{2} \sum_{i \neq j} [(\partial^2_{ii} U)^2 + (\partial^2_{jj} U)^2] \leq d \left|D^2 U\right|^2 \\ &\leq 2dc^2 (1 + |\nabla U|^2) \ , \end{aligned}$$

write $\|g\nabla U\|^2_{L^2(e^{-U})}$ and by integrating by parts twice

$$
\begin{aligned}
\int_{\mathbb{R}^d} |\nabla U|^2 g^2 e^{-U} &= \int_{\mathbb{R}^d} \nabla \cdot (g^2 \nabla U) e^{-U} \\
&= \int_{\mathbb{R}^d} g^2 \Delta U e^{-U} + 2\int_{\mathbb{R}^d} g\nabla g \cdot \nabla U e^{-U} \\
&\leq \sqrt{2d}c \sqrt{\int_{\mathbb{R}^d} g^2 e^{-U}} \sqrt{\int_{\mathbb{R}^d} g^2 e^{-U} + \int_{\mathbb{R}^d} |\nabla U|^2 g^2 e^{-U}} \\
&\quad + 2\sqrt{\int_{\mathbb{R}^d} |\nabla U|^2 g^2 e^{-U}} \sqrt{\int_{\mathbb{R}^d} |\nabla g|^2 e^{-U}} \\
&\leq \frac{1}{2}\int_{\mathbb{R}^d} |\nabla U|^2 g^2 e^{-U} + (\sqrt{2d}c + 2dc^2) \int_{\mathbb{R}^d} g^2 e^{-U} \\
&\quad + 4\int_{\mathbb{R}^d} |\nabla g|^2 e^{-U}
\end{aligned}
$$

by arguing by Cauchy-Schwarz and Young inequality multiple times. This is enough to prove the claim

- Proving that $A^*A + C^*C$ is coercive on $\mathcal{K}^\perp$ amounts to proving that for all $h \in H^1$

$$
\int \left(|\nabla_x h|^2 + |\nabla_v h|^2 \right) \mathrm{d}\mu \geq C \left[\int h^2 \,\mathrm{d}\mu - \left(\int h \,\mathrm{d}\mu \right)^2 \right] ,
$$

  that is, Poincaré inequality for μ. Since Poincaré is satisfied by the Gaussian, this requirement is equivalent to Poincaré inequality for $e^{-U(x)}$.

  Sufficient conditions for Poincaré inequality are a classical exercise: for instance,

$$
\frac{|\nabla U(x)|^2}{2} - \Delta U(x) \xrightarrow[|x|\to\infty]{} \infty
$$

  is taken from Appendix A.19 in [9]. Let us provide a rough outline of its proof: start from that Poincaré holds on bounded sets, then deal by integration by parts with the gradient to yield an estimate of $L^2\big((|\nabla U|^2/2 - \Delta U)\mu\big)$ in terms of $\dot{H}^1(\mu)$, and use it outside the ball.

  According to this criterion, all measures behaving at infinity like $e^{-|x|^\alpha}$, with $\alpha > 1$, satisfy a Poincaré inequality. We are then asking for our potential to grow at least linearly (indeed $e^{-|x|}$ satisfies Poincaré inequality as well) and at most exponentially.

In order to conclude, we only need to produce a $\mathcal{S}$ space, dense and such that $A : \mathcal{S} \to \mathcal{S}^m$ and $B : \mathcal{S} \to \mathcal{S}$ are continuous. This is easy if $U \in C^\infty$; if $U \in C^2$ only, we approximate with V_ε and pass things to the limit. Let us summarize in the following, corresponding to Proposition 35 in [9].

Proposition 4 (H^1 hypocoercivity for Fokker-Planck)**.** *Let $U \in C^2(\mathbb{R}^d)$, satisfying a Poincaré inequality and such that for some $C_U > 0$*

$$
\left|D^2 U(x)\right| \leq C_U (1 + |\nabla U(x)|) , \quad x \in \mathbb{R}^d .
$$

Then there exist C and $\lambda > 0$ such that for every $h_0 \in H^1(\mu)$

$$\left\| h_t - \int h_0 \, d\mu \right\|_{H^1(\mu)} \leq C e^{-\lambda t} \left\| h_0 \right\|_{H^1(\mu)} .$$

The proof of Theorem 3 could be carried out for this particular case, improving explicit estimates on constants.

For less regular initial data, the following hypoellipticity result will be useful. It corresponds to Theorem A.8 in Appendix A.21 in [9], completely devoted to regularization results. We shall here present its proof by expanding some computations and by slightly changing its presentation.

Proposition 5. *Suppose that $U \in C^2$ satisfies*

$$\left| D^2 U \right| \leq C_U (1 + |\nabla U|) .$$

Then there exists $C > 0$, which only depends on d and C_U, such that for all $h_0 \in L^2(\mu)$

$$\|h_t\|_{\dot{H}^{1,3}_{x,v}(\mu)} \leq \frac{C}{t^{3/2}} \|h_0\|_{L^2(\mu)} , \quad 0 < t < 1 .$$

Proof. First, we shall establish a differential inequality on $\|h_t\|^2_{\dot{H}^{1,3}_{x,v}}$. To see this, by writing the generator of the semigroup of h_t, in the sense that $\partial_t h_t + L h_t = 0$,

$$L := v \cdot \nabla_x - \nabla_x U \cdot \nabla_v - \Delta_v + v \cdot \nabla_v ,$$

one can start investigating $\frac{\mathrm{d}}{\mathrm{d}t} \|h_t\|^2_{\dot{H}^1_x(\mu)}$ by adding the term $L\nabla_x$ and studying commutation of L with $\nabla_x h$

$$\begin{aligned} \frac{1}{2} \frac{\mathrm{d}}{\mathrm{d}t} |\nabla_x h|^2 + \nabla_x h \cdot L \nabla_x h &= \nabla_x h \cdot (\partial_t + L) \nabla_x h = \nabla_x h \cdot (-\nabla_x L + L \nabla_x) h \\ &= \nabla_x h \cdot D_x^2 U \cdot \nabla_v h . \end{aligned}$$

By integrating against μ we recover $\frac{\mathrm{d}}{\mathrm{d}t} \|h_t\|^2_{\dot{H}^1_x(\mu)}$ as the first term, while the second one gives

$$\begin{aligned} \int L \nabla_x h \cdot \nabla_x h \, \mathrm{d}\mu &= - \int \Delta_v \nabla_x h \cdot \nabla_x h \ \mathrm{d}\mu + \int v \cdot D_v D_x h \cdot \nabla_x h \, \mathrm{d}\mu \\ &= - \int \Delta_v \nabla_x h \cdot \nabla_x h \, \mathrm{d}\mu + \int \nabla_v \cdot (D_v D_x h \nabla_x h) \, \mathrm{d}\mu \\ &= \int |D_x D_v h|^2 \ \mathrm{d}\mu \end{aligned}$$

since it is easy to prove that transport and spatial confinement terms cancel each other, giving

$$
\begin{aligned}
\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\|h\|^2_{\dot H^1_x(\mu)} + \int \left|D^2_{xv}h\right|^2\,\mathrm{d}\mu &= \int \nabla_v h (D^2_x U)\nabla_x h\,\mathrm{d}\mu \\
&= -\int h D^2_x U\cdot D^2_{xv}h\,\mathrm{d}\mu + \int h v D^2_x U \nabla_x h\,\mathrm{d}\mu \\
&\le \int \left|D^2_x U\right|^2 h^2\,\mathrm{d}\mu + \frac{1}{4}\int \left|D^2_{xv}h\right|^2\,\mathrm{d}\mu \\
&\quad + (\int \left|D^2_x U\right|^2 h^2\,\mathrm{d}\mu\Big)^{\frac{1}{2}}\Big(\int |v|^2\left|\nabla_x h\right|^2\,\mathrm{d}\mu\Big)^{\frac{1}{2}} \\
&\le C\Big(\int h^2\,\mathrm{d}\mu + \int |\nabla_x h|^2\,\mathrm{d}\mu\Big) + \frac{1}{4}\int \left|D^2_{xv}h\right|^2\,\mathrm{d}\mu \\
&\quad + C^2\Big(\int h^2\,\mathrm{d}\mu + \int |\nabla_x h|^2\,\mathrm{d}\mu\Big) + \frac{1}{4}\int |\nabla_x h|^2\,\mathrm{d}\mu + \frac{1}{4}\int \left|D^2_{xv}h\right|^2\,\mathrm{d}\mu
\end{aligned}
$$

where we used that $(\nabla_v)^* = -\nabla_v\cdot + v\cdot : H^1(\mu)\to L^2(\mu)$ and C-boundedness in H^1 of $|v|^2$ and $|\nabla_x U|^2$ operators. On the other hand, for $D^3_v h$ let us deal with each $\partial^3_{v_i v_j v_k} h$ singularly by relabelling it as ∂^3_v

$$
\begin{aligned}
\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\left|\partial^3_v h\right|^2 + \partial^3_v h L \partial^3_v h &= \partial^3_v h(-\partial^3_v L + L\partial^3_v)h \\
&= \partial^3_v h\big(-\partial^3_v(v\cdot\nabla_x h) + v\cdot\nabla_x\partial^3_v h - \partial^3_v(v\cdot\nabla_v h) + v\cdot\partial^3_v\nabla_v h\big) \\
&= -\partial^3_v h(\partial^3_{xvv}h + \partial^3_{vxv}h + \partial^3_{vvx}h) - 3(\partial^3_v h)^2
\end{aligned}
$$

where, for instance, $\partial^3_{xvv}h$ stays for $\partial^3_{x_i v_j v_k}h$. The second term on the left hand side, when integrated in $\mathrm{d}\mu$, gives

$$
\begin{aligned}
\int \partial^3_v h L\partial^3_v h\,\mathrm{d}\mu &= \int \partial^3_v h(-\Delta_v\partial^3_v h + v\cdot\nabla_v\partial^3_v h)\,\mathrm{d}\mu = \int \partial^3_v h(\nabla_v)^*(\nabla_v\partial^3_v h)\,\mathrm{d}\mu \\
&= \int \left|\nabla_v\partial^3_v h\right|^2\,\mathrm{d}\mu\ .
\end{aligned}
$$

When summing over i,j,k, by commuting derivation order in the mixed term,

$$
\begin{aligned}
\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\|h\|^2_{\dot H^3_v(\mu)} + \int \left|D^4_v h\right|^2\,\mathrm{d}\mu + 3\int \left|D^3_v h\right|^2\,\mathrm{d}\mu &= -3\int D^3_v h\cdot D^3_{vvx}h\,\mathrm{d}\mu \\
&= 3\int \Delta_v D^2_v h\cdot D^2_{vx}h\,\mathrm{d}\mu - 3\int v D^3_v h D^2_{vx}h\,\mathrm{d}\mu \\
&\le \frac{1}{4}\int \left|D^4_v h\right|^2\,\mathrm{d}\mu + 9\int \left|D^2_{vx}h\right|^2\,\mathrm{d}\mu \\
&\quad + \frac{1}{4}\int \left|D^3_v h\right|^2\mathrm{d}\mu + \frac{1}{4}\int \left|D^4_v h\right|^2\mathrm{d}\mu + 9C\int \left|D^2_{xv}h\right|^2\mathrm{d}\mu
\end{aligned}
$$

that we can summarize, times a constant ε, together with the previous as

$$\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\int\big(|\nabla_x h|^2+\varepsilon\left|D_v^3 h\right|^2\big)\,\mathrm{d}\mu$$
$$\le (C+C^2)\int h^2\,\mathrm{d}\mu+\left(C+C^2+\frac{1}{4}\right)\int|\nabla_x h|^2\,\mathrm{d}\mu$$
$$+\left(-\frac{1}{2}+9\varepsilon(C+1)\right)\int\left|D_{xv}^2 h\right|^2\,\mathrm{d}\mu-\varepsilon\frac{11}{4}\int\left|D_v^3 h\right|^2\,\mathrm{d}\mu-\frac{\varepsilon}{2}\int\left|D_v^4 h\right|^2\,\mathrm{d}\mu\ .$$

Let us now analyse the evolution of the mixed derivative $\nabla_x h\cdot\nabla_v h$, for which we will only add the 0-average term $L(\nabla_x h\cdot\nabla_v h)$ without multiplying as we did before:

$$(\partial_t+L)\nabla_x h\cdot\nabla_v h=-|\nabla_x h|^2+\nabla_v hD_x^2U\nabla_v h-2D_{xv}^2h\cdot D_v^2h-\nabla_x h\cdot\nabla_v h$$

where the four terms on the right hand side originate respectively from each of the terms in L. After integration we have

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\int\nabla_x h\cdot\nabla_v h\,\mathrm{d}\mu=&-\int|\nabla_x h|^2\,\mathrm{d}\mu-\int hD_x^2U\cdot D_v^2h\,\mathrm{d}\mu+\int h\ vD_x^2U\nabla_v h\,\mathrm{d}\mu\\&-2\int D_{xv}^2h\cdot D_v^2h\,\mathrm{d}\mu-\int\nabla_x h\cdot\nabla_v h\,\mathrm{d}\mu\\\le&-\int|\nabla_x h|^2\,\mathrm{d}\mu+C\delta\Big(\int h^2\,\mathrm{d}\mu+\int|\nabla_x h|^2\,\mathrm{d}\mu\Big)\\&+\frac{C+1}{2\delta}\int|\nabla_v h|^2\,\mathrm{d}\mu+\frac{C}{2\delta}\int\left|D_v^2h\right|^2\,\mathrm{d}\mu\\&+\int\left|D_{xv}^2h\right|^2\,\mathrm{d}\mu+\int\left|D_v^2h\right|^2\,\mathrm{d}\mu\\&+\frac{1}{4}\int|\nabla_x h|^2\,\mathrm{d}\mu+\int|\nabla_v h|^2\,\mathrm{d}\mu\ ,\end{aligned}$$

for all $\delta>0$. By picking it small enough and losing track of all constants, we can summarize it into

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\int\nabla_x h\cdot\nabla_v h\,\mathrm{d}\mu\lesssim&-\int|\nabla_x h|^2\,\mathrm{d}\mu+\int h^2\,\mathrm{d}\mu+\int|\nabla_v h|^2\,\mathrm{d}\mu\\&+\int\left|D_v^2h\right|^2\,\mathrm{d}\mu+\int\left|D_{xv}^2h\right|^2\,\mathrm{d}\mu\ .\end{aligned}$$

Now, by linearity and L^2-nonexpansivess of the Fokker-Planck semigroup, we can suppose that $\|h_t\|_{L^2(\mu)}\le 1$ for all $t\ge 0$. We may then estimate the last term in the last inequality with

$$\begin{aligned}\int h^2\,\mathrm{d}\mu+\int|\nabla_v h|^2\,\mathrm{d}\mu+\int\left|D_v^2h\right|^2\mathrm{d}\mu&\lesssim 1+\Big(\int\left|D_v^3h\right|^2\mathrm{d}\mu\Big)^{\frac{1}{3}}+\Big(\int\left|D_v^3h\right|^2\mathrm{d}\mu\Big)^{\frac{2}{3}}\\&\lesssim\int\left|D_v^3h\right|^2\mathrm{d}\mu\ ,\end{aligned}$$

summarize the previous differential inequality into

$$\frac{\mathrm{d}}{\mathrm{d}t}\int\big(|\nabla_x h|^2+\varepsilon\left|D_v^3h\right|^2\big)\,\mathrm{d}\mu$$
$$\lesssim-\Big(\int\left|D_{xv}^2h\right|^2\mathrm{d}\mu+\int\left|D_v^4h\right|^2\mathrm{d}\mu\Big)+1+\Big(\int|\nabla_x h|^2\,\mathrm{d}\mu+\varepsilon\int\left|D_v^3h\right|^2\mathrm{d}\mu\Big)$$

and thanks to

$$\int \left|D_v^3 h\right|^2 \mathrm{d}\mu \lesssim \left(\int |\nabla_x h|^2 \mathrm{d}\mu + \int \left|D_v^4 h\right|^2 \mathrm{d}\mu + \int \left|D_{xv}^2 h\right|^2 \mathrm{d}\mu \right)^{\frac{3}{4}} ,$$

together with

$$\left| \int \nabla_x h \cdot \nabla_v h \, \mathrm{d}\mu \right| \lesssim \left(\int |\nabla_x h|^2 \mathrm{d}\mu + \varepsilon \int \left|D_v^3 h\right|^2 \mathrm{d}\mu \right)^{\frac{2}{3}} .$$

Now we are done by combining the last four differential inequalities thanks to Lemma A.26 in [9]: this is simply a calculus result, whose proof is not straightforward and will not be presented here. It involves functions $(0,1] \to \mathbb{R}$ and we will use the common symbols $\lesssim$ a and $\asymp$ with the meaning that the ratio of two functions is bounded from above and both from above and from below, respectively. This general result considers $\mathcal{E}$, X, Y, Z, $\mathcal{M}$ functions $(0,1] \to \mathbb{R}$, and we ask for

$$\begin{aligned}
\mathcal{E} &\asymp X + Y \ , \\
|\mathcal{M}| &\lesssim \mathcal{E}^{1-\delta} \ , \\
\frac{\mathrm{d}\mathcal{E}}{\mathrm{d}t} &\lesssim -Z + \mathcal{E} \ , \\
Y &\lesssim (X+Z)^{1-\theta} \ , \\
\frac{\mathrm{d}\mathcal{M}}{\mathrm{d}t} &\lesssim -X + Y + Z
\end{aligned}$$

where δ, θ are fixed constants in $(0,1)$, which turn out to imply that

$$\mathcal{E} \lesssim t^{-\max\left(\frac{1}{\delta}, \frac{1}{\theta} - 1\right)} .$$

In our case one may prove that we need to set

$$\begin{aligned}
\mathcal{E}(t) &= \int |\nabla_x h_t|^2 \mathrm{d}\mu + \varepsilon \int \left|D_v^3 h_t\right|^2 \mathrm{d}\mu \ , \\
X(t) &= \int |\nabla_x h_t|^2 \mathrm{d}\mu \ , \\
Y(t) &= \int \left|D_v^3 h_t\right|^2 \mathrm{d}\mu \ , \\
Z(t) &= \int \left|D_v^4 h_t\right|^2 \mathrm{d}\mu + \int \left|D_{xv}^2 h_t\right|^2 \mathrm{d}\mu \ , \\
\mathcal{M}(t) &= \int \nabla_x h_t \cdot \nabla_v h_t \, \mathrm{d}\mu
\end{aligned}$$

while $\delta = \frac{1}{3}$ and $\theta = \frac{1}{4}$ provide the rate of convergence t^{-3} in the squared inequality, which is the claim. □

Thus convergence still holds even if we start from a $L^2(\mu)$ initial data: indeed, by Proposition 5 it suffices to regularize until any positive time ε, and then apply $H^1(\mu)$ convergence. However this is too strong as an assumption, since $h \in L^2(\mu)$ means Theorem applies to the natural equation for initial data f_0 satisfying $\int f_0^2 e^{U(x) + \frac{|v|^2}{2}} \, \mathrm{d}x \, \mathrm{d}v < \infty$. This is rather strong as a decay assumption, so we look forward to convergence with milder initial assumptions.

Actually we will give up on Sobolev regularization, and take the entropy way, allowing for strong assumptions – essential quadraticness – on the potential U: the following is Theorem 39 in [9], of which it is nice to provide an outline of two possible ways of proof.

Proposition 6. • *Take $U \in C^2(\mathbb{R}^d)$, with $\left|D^2 U\right|$ bounded*

- *Suppose μ satisfies a Logarithmic Sobolev Inequality*
- *Let f_0 be a probability measure on $\mathbb{R}^d_x \times \mathbb{R}^d_v$ with finite second moment:*

$$\int_{\mathbb{R}^d_x \times \mathbb{R}^d_v} f_0(x,v)\big(\,|x|^2 + |v|^2\,\big)\, dx\, dv < \infty \ .$$

Call, as usual, f_t the evolution of f_0 under the Fokker-Planck natural equation

$$\partial_t f + v \cdot \nabla_x f - \nabla_x U(x) \cdot \nabla_v f = \frac{\sigma^2}{2} \Delta_v f + \gamma \nabla_v \cdot (v f) \ ,$$

Then hypocoercivity holds in the weaker entropy sense: upon calling, as usual,

$$h_t(x,v) = \frac{f_t(x,v)}{\mu(x,v)} \doteqdot f_t(x,v) e^{\frac{2\gamma}{\sigma^2}\left(U(x) + \frac{|v|^2}{2}\right)} \ ,$$

we have that there exist positive constants C and λ, which depend from f_0, such that

$$\int_{\mathbb{R}^d_x \times \mathbb{R}^d_v} h_t \log h_t \, d\mu + \int_{\mathbb{R}^d_x \times \mathbb{R}^d_v} \frac{|\nabla h_t|^2}{h_t}\, d\mu \leq C e^{-\lambda t} \ .$$

Proof. The goal is just regularizing f_0 until we get finite H and I; then we will be able to apply Theorem 3 with the usual $K = 1$, $A = \nabla_v$, B concerning transport and confinement term, and we will write

$$[A, B] = \nabla_x + 0$$

and

$$[C_1, B] = 0 - \nabla^2_x U \cdot \nabla_v \cdot$$

that is, with the notation of Theorem 3, $Z_1 = Z_2 = I$, $C_1 = \nabla_x$ $R_1 = 0$ and $R_2 = -\nabla^2_x U \cdot \nabla_v$. Concerning pointwise bounds to prove, all commutators are zero, except for $[A, A^*] = I$, therefore pointwise bounded with respect to I and $\{C_j\}_j$. Just, in order to prove R_2 bounded with respect to A, our assumption makes sure that $|R_2(x,v)| = \left|D^2 U(x)\right| \leq C \lesssim 1 = |A(x,v)|$. Last, pointwise coercivity is equivalent to Poincaré inequality for μ, which is implied by LSI.

There are two routes to prove finiteness in finite time of H and I: in Route 1 we use first Sobolev regularization to yield for $t \to 0$

$$\int \big(\,|\nabla_x f_t|^2 + |\nabla_v f_t|^2\,\big)\, \mathrm{d}x\, \mathrm{d}v = O(t^{-\gamma})$$

for some $\gamma > 0$. By Nash inequality (Lemma A.25 in [9]) we may bound the homogeneous Sobolev norm with a power of L^2 norm (recall that L^1 norm is conserved). Since $\int f \log f \leq \int f^2$, we have that, for small t, $\int f_t \log f_t = O(t^{-\beta})$. On the other hand,

$$\int f_t \log(\mu^{-1}) \lesssim \int f_t(x,v)\big(\,|x|^2 + |v|^2\,\big)\, \mathrm{d}x\, \mathrm{d}v = O(1+t)$$

for small t, by a simple computation. It follows that for all $t_0 > 0$

$$\int f_{t_0}(x,v) \log\left(\frac{f_{t_0}(x,v)}{\mu(x,v)}\right) \mathrm{d}x\, \mathrm{d}v < \infty \ .$$

Hence, again by entropy regularization, we can deduce that for all $t_1 > t_0$ $I_\mu(f_{t_1}) < \infty$ as well, and we can apply Theorem 3.

Next, in Route 2 we shall suppose $U \in C^\infty$ and $\left|\nabla^j U\right| \leq C_j$ for all $j \geq 2$, and f_0 having bounded moments of all orders. This – in particular, bounds on all derivatives of the potential– gives that instantaneously $f_t \in H^{k,l}_{x,v}$ for all k, l positive. Since it may be proved that

$$\int \frac{|\nabla f|^2}{f} \lesssim \|f\|_{H^s_k}$$

for k and s large enough, we have $\int \frac{|\nabla f_t|^2}{f_t} < \infty$ for all $t > 0$.

On the other hand, since

$$I_\mu(f) = \int \frac{|\nabla f|^2}{f} + \nabla f \cdot \nabla E + f\,|\nabla E|^2$$

and we can estimate the middle term in a standard way with the other two, we are left with an estimate for $\int f\,|\nabla E|^2$. Again, by Lipschitz property of the potential $|\nabla E| \leq C\big(\,|x| + |v|\,\big)$, so that an estimate on the second moment will suffice. This will be done as in the previous case, and we have $I_\mu(f_t) < \infty$. Logarithmic Sobolev Inequality gives finiteness of entropy, and we can act as in the previous case.

□

2.3. **Kinetic Fokker-Planck equation with a nonlinearity.** This Subsection is meant to give a quick summary of Part III in [9], in particular the outline of the proof of the main result and its application to self-consistent Vlasov-Fokker-Planck equation: here we shall consider a variant of the usual kinetic Fokker-Planck equation, in the following sense:

- The usual space-confining potential U is replaced by space periodicity of the solution, i.e. $x \in \mathbb{T}^d$;
- An interactive force acts among particles, and it is represented by a small and smooth potential W.

Let us read it then:

$$\begin{cases} \partial_t f + v \cdot \nabla_x f + F[f] \cdot \nabla_v f = \Delta_v f + \nabla_v \cdot (vf) \\ F[f](t,x) = -\displaystyle\int_{\mathbb{R}^d_v} \big(\nabla W(\cdot) *_x f_t(\cdot, v)\big)(x)\,\mathrm{d}v \end{cases} \tag{2.4}$$

Well-posedness is as in the linear case: the tool is regularity of coefficients in the SDE. Here one may prove that $\|F\|_{C^k} \lesssim \|W\|_{C^{k+1}}$, so that it is possible to prove all regularity results as in the linear case.

Also, the Maxwellian $M(v) = (2\pi)^{-d/2} e^{-\frac{|v|^2}{2}}$ is a stationary solution of (2.4) as expected, since x-constant functions lie in $\operatorname{Ker} F$.

Proposition 7 (Convergence to equilibrium in Vlasov-Fokker-Planck)**.** *Let f_0 be a probability measure on $\mathbb{T}^d \times \mathbb{R}^d$, with*

$$\int_{\mathbb{T}^d \times \mathbb{R}^d} f_0(x,v)\,|v|^k \;dx\,dv < \infty$$

for all $k \in \mathbb{N}$. Let $W \in C^\infty(\mathbb{T}^d)$ be even with $\int_{\mathbb{T}^d} W = 0$, and suppose $\|W\|_{L^\infty} < 0.38$. Then

$$\|f_t - M\|_{L^2} = O(t^{-\infty}) \ ,$$

where this is a classical notation meaning that for all $\alpha > 0$ there exists $C_\alpha > 0$ independent from t and f_0 such that $\|f_t - M\|_{L^2} \le C_\alpha t^{-\alpha}$.

This result is a consequence of a way more general theory, of which we introduce characters with an outline of assumptions:

- $\{X_s\}_{s\ge 0}$ is a decreasing family of normed spaces (a high s corresponds to considering very regular functions) such that: X_0 has a Hilbert structure (we will write $\|\cdot\| = \|\cdot\|_0$), $X^{s_2} \hookrightarrow X^{s_1}$ if $s_1 \le s_2$, and they interpolate, in that $\|\cdot\|_{(1-\theta)s_0+\theta s_1} \lesssim \|\cdot\|_{s_0}^{1-\theta} \|\cdot\|_{s_1}^{\theta}$;
- We consider the equation
$$\partial_t f + Bf = \mathcal{C} f$$
where B ("conservative") and $\mathcal{C}$ ("dissipative") are to be thought as differential operators defined on convex and bounded (for all s) $X \subset \bigcap_s X_s$; they allow losses of derivatives, in that they are smooth from large enough s to smaller s; f is quite smooth, namely $f \in C^1(\mathbb{R}_+, X_s) \cap C^0([0,+\infty), X_s)$, and $f_t \in X$ for all t;
- There exists a stationary state, i.e. a $f_\infty \in X$ s.t. $Bf_\infty = \mathcal{C} f_\infty = 0$;
- We have $\{\Pi_j\}_{j=1}^J$ nonlinear projections defined on X, ideally nested, projecting to a set of minimizers of $\mathcal{E}$ with some constraints, and shrinking to f_∞: for all j we ask $\Pi_j(X) \subset \operatorname{Ker} \mathcal{C}$, $\Pi_j f_\infty = f_\infty$, $\Pi_J(X) = f_\infty$ and both Π_j' and Π_j'' allow losses of derivatives.
- $\mathcal{E} : X \to \mathbb{R}$ will serve as a Lyapounov functional, admitting f_∞ as unique minimizer. Also we ask that projections push us to equilibrium, in the following sense: for all $f \in X$
$$\mathcal{E}(f_\infty) \le \mathcal{E}(\Pi_1 f) \le \mathcal{E}(f) \ , \quad f \in X,$$
and for all $\varepsilon \in (0,1)$
$$\mathcal{E}(\Pi_1 f) - \mathcal{E}(f) \lesssim -\|f - \Pi_1 f\|^{2+\varepsilon}$$
and the convergence to equilibrium along projections is essentially quadratic
$$\|\Pi_1 f - f_\infty\|^{2+\varepsilon} \lesssim \mathcal{E}(\Pi_1 f) - \mathcal{E}(f_\infty) \lesssim \|\Pi_1 f - f_\infty\|^{2-\varepsilon}$$
where constants may depend on ε;
- $\mathcal{E}$ is dissipated along $\mathcal{C}$, and coercively so out of Π_1's range:
$$\mathcal{E}'(f) \cdot \mathcal{C}(f) \lesssim -[\mathcal{E}(f) - \mathcal{E}(\Pi_1 f)]^{1+\varepsilon} \ ,$$
while it is conserved along B. Also, we ask that the dissipation of $\mathrm{Id} - \Pi_j$ along B and starting from $\Pi_j f$ dominates $\|(\Pi_j - \Pi_{j+1})f\|^{1+\varepsilon}$:
$$\left\|(\mathrm{Id} - \Pi_j)'_{\Pi_j f} \cdot (B \Pi_j f)\right\| \gtrsim \|\Pi_j f - \Pi_{j+1} f\|^{1+\varepsilon} \ .$$

Theorem 8. *Under the previous assumptions, for every $\beta > 0$,*

$$\mathcal{E}(f_t) - \mathcal{E}(f_\infty) = O(t^{-\beta}) \ .$$

This implies that for every $s \ge 0$ and $\beta > 0$

$$\|f_t - f_\infty\|_s = O(t^{-\beta}) \ .$$

It is easy to see how the second part is implied: we just prove it for $s = 0$, then we argue by interpolation and boundedness of solutions for all s. Then for any $\varepsilon > 0$ (say $\varepsilon = 1$) our conditions on $\mathcal{E}$ give in particular that $\mathcal{E}(f) - \mathcal{E}(f_\infty) \gtrsim \|f_\infty - \Pi_1 f\|^{2+\varepsilon}$ and $\mathcal{E}(f) - \mathcal{E}(f_\infty) \gtrsim \|f - \Pi_1 f\|^{2+\varepsilon}$, so that

$$\|f_t - f_\infty\| \leq \|f - \Pi_1 f\| + \|\Pi_1 f - f_\infty\| \lesssim [\mathcal{E}(f) - \mathcal{E}(f_\infty)]^{\frac{1}{2+\varepsilon}}$$

which easily gives the claim. How to prove the first statement in Theorem 8? Unfortunately $\mathcal{E}$ itself is not enough Lyapounov, so we also have to consider the dissipation of $\|Id - \Pi_j\|^2$ along B:

$$\mathcal{L}(f) := \mathcal{E}(f) - \mathcal{E}(f_\infty) + \sum_{j=1}^{J} a_j \langle (Id - \Pi_j) f, (Id - \Pi_j)'(f) \cdot Bf \rangle \ .$$

Actually this is not Lyapounov either, but we will consider time intervals when $\mathcal{E}(f) - \mathcal{E}(f_\infty)$ is bounded from above and below; on this interval we will fix $\{a_j\}$ so that $\mathcal{L}$ is Lyapounov, and after a sufficient decrease we update $\{a_j\}$ to perform the same trick again. This will give $\frac{\mathrm{d}\mathcal{L}}{\mathrm{d}t} \leq -C\mathcal{L}^{1+\delta}$, which will provide enough estimate for $\mathcal{E}(f_t) - \mathcal{E}(f_\infty)$. In particular a rather technical result states that

Proposition 9. *Suppose that for some $E > 0$*

$$\frac{E}{2} \leq \mathcal{E}(f) - \mathcal{E}(f_\infty) \leq E \ ,$$

and fix $\varepsilon > 0$ small enough. Then there exist choice of the $\{a_j\}$ and a constant K, both depending on ε, such that

$$\frac{E}{4} \leq \mathcal{L}(f) \leq \frac{5}{4} E \ ,$$

and

$$\mathcal{L}'(f) \cdot (\mathcal{C}f - Bf) \leq -K a_{J-1} E^{1+\varepsilon} \ .$$

Proof(of Theorem 8 from Proposition 9). Take $E > 0$ small enough, so that $\mathcal{E}(f_0) - \mathcal{E}(f_\infty) \leq E$ – which is finite because initial data are in X where $\mathcal{E}$ is defined – and let $[t_0, t_0 + T(E)]$ be the time-interval – which is connected from decreasingness of $\mathcal{E}$ – for which $\frac{E}{2} \leq \mathcal{E}(f_t) - \mathcal{E}(f_\infty) \leq E$

Now, since coefficients in the functional are rather large, comparable to E, we expect dissipation to be rather negative, comparable to E. Since then we have upper bounds for the difference we expect time-interval to be small, let us prove that $T(E) \leq CE^{-\lambda\varepsilon}$, for suitable λ: take $\{a_j\}$ from the previous Proposition 9. Then, according to [9], we may take $a_{J-1} \geq K' E^{l\varepsilon}$, for sufficiently universal K' and l. Therefore if $t \in [t_0, t_0 + T(E)]$

$$\frac{\mathrm{d}}{\mathrm{d}t} \mathcal{L}(f_t) \leq -a_{J-1} K E^{1+\varepsilon} \leq -\tilde{K} E^{1+\varepsilon(l+1)} \ .$$

Recalling that we are also provided with upper and lower bounds on $\mathcal{L}$,

$$-E \leq \mathcal{L}(f_{t_0+T(E)}) - \mathcal{L}(f_{t_0}) \leq -\tilde{K} E^{1+\varepsilon(l+1)} \cdot T(E) \ ,$$

which gives the desired estimate

$$T(E) \leq CE^{-\varepsilon\lambda}.$$

Now let us prove that this provides $\mathcal{E}(f_t) - \mathcal{E}(f_\infty) = O(t^{-\frac{2}{\varepsilon\lambda}})$: the argument is classical. Write $E_t := \mathcal{E}(f_t) - \mathcal{E}(f_\infty)$, then, after $T(E_0)$, E_t drops below $E_0/2$, and therefore (by monotonicity) it does after $CE_0^{-\lambda\varepsilon}$ as well; then, after

$$\sum_{n=0}^{m} T(2^{-n}E_0) \leq CE_0^{-\lambda\varepsilon}\left(\frac{1-2^{\lambda\varepsilon(m+1)}}{1-2^{\lambda\varepsilon}}\right) \lesssim E_0^{-\lambda\varepsilon}2^{\lambda\varepsilon m} ,$$

$\mathcal{E}(f_t) - \mathcal{E}(f_\infty)$ has dropped under $E_0 2^{-m-1}$. This gives that, if $t = C2\lambda\varepsilon m$, $E_t \lesssim t^{\frac{-m-1}{\varepsilon\lambda m}}$, which gives $O(t^{-\frac{1+\delta}{\varepsilon\lambda}})$ for all $\delta > 0$, upon taking m large enough. $\square$

Now let us apply this general result to our Vlasov-Fokker-Planck case: we are going to follow the proof of Theorem 56 of [9], aiming to yield better-than-polynomial convergence in Sobolev norms.
fix f_0 as in the hypothesis; we want to work in $X_s = H^s_{x,v}\big((1+|v|^2)^s\big)$, $s \geq 0$: for integer k they may be defined as

$$\|f\|_{X_k} := \sum_{|l|+|m|\leq k} \left\|\nabla_x^l \nabla_v^m f\right\|_{L^2_{x,v}((1+|v|^2)^k)} ,$$

so we wish the solution f_t to be t-bounded in this space. To this scope, we will by start with integer k and argue by interpolation between moments and Sobolev norms.
Concerning the latter, hypoellipticity techniques give that for all $t_0 > 0$ and $k \in \mathbb{N}$,

$$\sup_{t\geq t_0} \|f_t\|_{H^k} < \infty ;$$

concerning the former, let us define the regularized moments

$$M_k(t) := \int f_t(x,v)(1+|v|^2)^{k/2}\,\mathrm{d}x\,\mathrm{d}v .$$

Then one may prove that for all $k \geq 1$

$$M_k' \leq -kM_k + CM_{k-1} + \big(k^2 + k(n-1)\big)M_{k-2} ,$$

where C is a constant arising from the forcing term, and the negative coefficient k comes from the divergence term. We can argue by induction on k, since solutions to $f' \leq -af + b$ are globally bounded; indeed M_0 is mass – which is exactly conserved – and, concerning M_1, it is easy to bound negative moments with mass, and we are done.
We have then that f is bounded in time in all X_k for integer k, and therefore for $s \in \mathbb{R}$ by interpolation. Next, we translate in concrete means the previous formalism: we will take

$$Bf = v \cdot \nabla_x f + F[f]$$

and

$$\mathcal{C}f = \Delta_v f + \nabla_v \cdot (vf) ,$$

and clearly $f_\infty(x,v) = M(v)$. Also, we will consider spatial density $\rho(t,x) := \int_{\mathbb{R}^d} f(t,x,v)\,\mathrm{d}v$, and choose as projection the homogenization in velocity

$$\Pi_1 f(t,x,v) = \rho(t,x)M(v)$$

and $\Pi_2 f(t,x,v) = M(v) = f_\infty$. Our functional will be

$$\mathcal{E}(f) := \int f\log f + \frac{1}{2}\int f(x,v)\,|v|^2\,\mathrm{d}x\,\mathrm{d}v + \frac{1}{2}\int_{\mathbb{T}^d\times\mathbb{T}^d} \rho(x)\rho(y)W(x-y)\,\mathrm{d}x\,\mathrm{d}y .$$

Let us then check it behaves properly with projections:

$$\mathcal{E}(f)-\mathcal{E}(\Pi_1 f)=\int f\log\frac{f}{\rho M}\geq\frac{1}{2}\|f-\rho M\|_{L^1}^2\geq C_\varepsilon\|f-\rho M\|_{H^k}^{-\varepsilon}\|f-\rho M\|_{L^2}^{2+\varepsilon}\ ,$$

where first we just performed a computation, then we used Kullback-Pinsker inequality and L^1-H^k interpolation, for $k(\varepsilon)$ large enough. Similarly,

$$\begin{aligned}\mathcal{E}(\Pi_1 f)-\mathcal{E}(f_\infty)&=\int\rho(x)\log\rho(x)\,\mathrm{d}x+\frac{1}{2}\int\big(\rho(x)-1\big)\big(\rho(y)-1\big)W(x-y)\,\mathrm{d}x\,\mathrm{d}y\\&\geq\frac{1}{2}\big(1-\|W\|_\infty\big)\|\rho-1\|_{L^1_x}^2=C\|\rho M-M\|_{L^1}^2\end{aligned}$$

and then we argue by interpolation by estimating the last term with $\|\rho M-M\|_{L^2}^{2+\varepsilon}$. Also, since $\int W=0$,

$$\mathcal{E}(\Pi_1 f)-\mathcal{E}(f_\infty)\leq\|\rho-1\|_{L^2}+C\|\rho\|_{L^1}^2$$

and again we argue by interpolation as before. Notice that we did not introduce our workspace X. Indeed, we already know that our solution is X_s-bounded in time, for all $s\geq 0$. Also, let us remind that we have a hypoelliptical lower bound on the solution f. Then ρ as well is bounded from below, uniformly in x, which allows to choose as workspace, in order to avoid issues with $H(f|\rho M)$,

$$X=\{f:\|f\|_s\leq C_s\ \forall s,\ \rho\geq c>0\}\ ,$$

where C_s and c are depending from our equation. Now, this definition allows to easily prove that many of our assumptions are satisfied, leaving us with dissipation issues. First, let us compute that, with the L^2 differential structure, and thanks to f and ρ being bounded from below,

$$\mathcal{E}'(f)\cdot h=\int\left(\log f(x,v)+1+\frac{|v|^2}{2}\right)h(x,v)\,\mathrm{d}x\,\mathrm{d}v+\int\rho(y)h(x,v)W(x-y)\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}v\ ,$$

so that, taking f satisfying $\partial_t f=\mathcal{C}f$

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{E}(f_t)=\mathcal{E}'(f)\cdot\mathcal{C}f=-\int f\left|\nabla_v\log\left(\frac{f}{M}\right)\right|^2\leq-2\int f\log\frac{f}{\rho M}=\mathcal{E}(\Pi_1 f)-\mathcal{E}(f)$$

where we used log-Sobolev inequality x-wise and a previous computation. Also, it is easy to show that $\mathcal{E}$ is invariant on the flow of B, so we are left to show that $\left\|(Id-\Pi_1)'_{\Pi_1 f}\cdot(B\Pi_1 f)\right\|\geq K_\varepsilon\|\Pi_1 f-f_\infty\|^{1+\varepsilon}$.
In fact, notice that Π vanishes on the range of B, and that by linearity $\Pi'=\Pi$, so that after some manipulations, and using the lower bound on ρ, we yield

$$\big\|(Id-\Pi)'_{\Pi f}\cdot(B\Pi f)\big\|^2\geq C\int_{\mathbb{T}^n}\rho\left|\frac{\nabla\rho}{\rho}+\nabla(\rho*W)\right|^2\ .$$

Now, write $\nabla(\rho*W)=\nabla\big(-\log(ce^{-\rho*W})\big)$ where c is normalizing. It is very easy to see that $\rho*W$ is bounded from above and below, so that log-Sobolev applies, and we have

$$\int\rho\left|\frac{\nabla\rho}{\rho}+\nabla(\rho*W)\right|^2\geq K\bigg(\int\rho\log\rho+\int\rho(\rho*W)+\log\int e^{-W*\rho}\bigg)\ .$$

We estimate the first term with Cziszar-Kullback-Pinsker inequality, and the other two with standard means, and we yield, for $\delta = \|W\|_\infty$,

$$\left\|(Id - \Pi_1)'_{\Pi_1 f} \cdot (B\Pi_1 f)\right\|^2 \geq K\left(\frac{1}{2} - \delta - \frac{\delta^2 e^\delta}{2}\right) \|\rho - 1\|^2_{L^1} \ ,$$

then integrate times $M(v)\,\mathrm{d}v$ and interpolate with L^2 and H^k as before.

Remark. Let us remark that, indeed, we proved that better-than-polynomial convergence holds for L^2 and for all $H^k\big((1+|v|)^k\big)$. However, by recalling the precise statement of the Theorem, we yield entropic and therefore L^1 convergence as well.

3. Improvement of quantitative estimates on entropical relaxation

Here we will be studying kinetic Fokker-Planck equation in relation with its long-time behaviour by mainly following [5]. In addition to the review of the work and proofs developed in it in a general framework, we shall focus on the proof of entropical relaxation by particularly emphasizing the case of kinetic Fokker-Planck equation, providing a sharper proof and a remarkable improvement in the quantitative computation of the relaxation constant with respect to the computation in [6], both of which things constitute a novelty.

The outline of the Section is as follows: in Subsection 3.1 we shall expose classical features of Γ-calculus; in Subsection 3.2 we shall show the proof of Theorem 11, which is the main content of [5] and which provides entropical relaxation in a very general context; in Subsection 3.3 we shall perform a precise estimate on the relaxation constant for kinetic Fokker-Planck equation provided in the previous Subsection.

3.1. Markov semigroups and generalized Γ-calculus.

We shall briefly sketch some classical concepts on Markov semigroups, a more precise explanation and further details may be found for instance in [2]. Let $(X_t)_{t\geq 0}$ be process on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$ taking values in $\mathbb{R}^n$, such that $X_0 = x \in \mathbb{R}^n$ a.s. and adapted to the filtration $\{\mathcal{F}_t\}_{t\geq 0}$. By denoting the law of the random variable X_t by $\mathcal{L}X_t$, we shall call $(X_t)_t$ a (time-homogeneous) *Markov process* if, for $0 \leq s < t$, $\mathcal{L}X_t$ given $\mathcal{F}_s$ may be identified with $\mathcal{L}X_t$ given X_s and with $\mathcal{L}X_{t-s}$ in the following sense: for all smooth and bounded f,

$$\mathbb{E}(f(X_t)|\mathcal{F}_s) = \mu_{t-s}(X_s, f) \ .$$

Next, given a Markov process $(X_t)_t$, we call P_t its associated *Markov semigroup* on $\mathbb{R}^n$, by defining for all suitable measurable functions $f : \mathbb{R}^n \to \mathbb{R}$

$$P_t f(x) := \mathbb{E}[f(X_t)|X_0 = x] \quad t \geq 0 \ , \ x \in \mathbb{R}^n$$

and define its generator L as

$$Lf(x) := \lim_{t\to 0} \frac{P_t f(x) - f(x)}{t} \quad x \in \mathbb{R}^n$$

as long as f belongs to some $\mathcal{D}(L)$, the domain of L.

Suppose that both L and P_t fix a some set $\mathcal{A}$ – in that they map $\mathcal{A}$ into itself – and take Φ as a functional defined on $\mathcal{A}_+$, that is, positive functions of $\mathcal{A}$. Here we will also need to take the structure of the function space into consideration, namely we ask $\mathcal{A}$ to be included in a Banach space $\mathcal{B}$ which in this memoire will be a natural L^2 space whenever not specified. Suppose $\Phi : \mathcal{A}_+ \to \mathcal{A}_+$ – which we will call gauge function throughout this essay – is smooth enough and define

$$\Gamma_{L,\Phi}(f) := \frac{L\Phi(f) - \mathrm{d}\Phi(f).Lf}{2} \ , \quad f \in \mathcal{A}_+$$

where $\mathrm{d}\Phi(f).Lf$ denotes the dual coupling between the differential $\mathrm{d}\Phi(f) \in \mathcal{A}^*$ and $Lf \in \mathcal{A}$ according to the structure of $\mathcal{B}$.

From the definition it follows immediately that Γ is (L, Φ)-bilinear, in that a linear combination of either a finite set of generators or a finite set of gauge functions transfers to the operator Γ. It is also easy to show that $\Gamma_{L,\Phi}(f) \geq 0$ for all f if Φ

is a positive and convex function $\mathbb{R} \to \mathbb{R}$. Indeed for all x

$$\begin{aligned}\Gamma_{L,\Phi}(f)(x) &= \frac{[\partial_{t|_{t=0}} P_t \Phi(f) - \partial_{t|_{t=0}} \Phi(P_t f)](x)}{2} \\ &= \lim_{t\to 0} \frac{[P_t\Phi(f) - \Phi(f) - (\Phi(P_t f) - \Phi(f))](x)}{2} \\ &= \lim_{t\to 0} \frac{\mathbb{E}[\Phi(f(X_t))|X_0 = x] - \Phi(\mathbb{E}[f(X_t)|X_0 = x])}{2} \geq 0\end{aligned}$$

from Jensen inequality.

While dependence from L may be omitted since we will always be studying Γ operators with respect to some fixed evolution P_t, we shall always write Γ_Φ in order to keep track of the gauge function.

A remarkable case of gauge function is $\Phi(f) = f^2$, which has been first introduced by Bakry and Émery in [1] and which makes the functional Γ polarizable into a two-variable one. In this case the traditional notation reads

$$\Gamma := \Gamma_{\cdot^2}$$

so that what is traditionally called *carré du champ*

$$\Gamma(f,g) = \frac{L(fg) - fLg - gLf}{2}$$

measures the defect of Leibniz property of the generator L. We will also write

$$\Gamma_2 := \Gamma_{\Gamma(\cdot,\cdot)} \ ,$$

that is,

$$\Gamma_2(f,g) = \frac{L\Gamma(f,g) - \Gamma(f,Lg) - \Gamma(g,Lf)}{2} \ ,$$

where the number 2 referring to an iteration is not to be confused with the aforementioned square operation $\cdot^2$. First, let us notice that Γ is insensitive to affine perturbations of the functional: for all $a, b \in \mathbb{R}$, if $\Phi(f) = f^2 + af + b$, $\mathrm{d}\Phi(f).g = 2fg + ag$ for all f and g, giving that

$$\Gamma_{\cdot^2+a\cdot+b}(f) = \frac{L(f^2 + af + b) - (2f.Lf + aLf)}{2} = \frac{Lf^2 + Lb - 2f.Lf}{2}$$

and, since by Markov property $Lb = 0$, we have $\Gamma_{\cdot^2+a\cdot+b}(f) = \Gamma(f,f)$.

Let us now consider the classical case of $L = \Delta - \nabla U \cdot \nabla$, induced by the Markov process

$$\mathrm{d}X_t = -\nabla U(X_t)\,\mathrm{d}t + \mathrm{d}B_t$$

where B_t is a n-dimensional Brownian motion, and compute for all f and g

$$\Gamma(f,g) = \frac{1}{2}[\Delta(fg) - \nabla U \cdot \nabla(fg) - g\Delta f + g\nabla U \cdot \nabla f - f\Delta g - f\nabla U \cdot \nabla g] = \nabla f \cdot \nabla g$$

where obtaining the standard Dirichlet form shows the reason of the factor $\frac{1}{2}$ multiplying Γ. Then

$$\begin{aligned}\Gamma_2(f,g) = \frac{1}{2}[&(\Delta - \nabla U \cdot \nabla)(\nabla f \cdot \nabla g) \\ &- \nabla f \cdot \nabla(\Delta g - \nabla U \cdot \nabla g) - \nabla g \cdot \nabla(\Delta f - \nabla U \cdot \nabla f)] \ .\end{aligned}$$

Since

$$\nabla(\nabla f \cdot \nabla g) = D^2 f \nabla g + D^2 g \nabla f \ ,$$

we have

$$
\begin{aligned}
\Gamma_2(f,g) =& \frac{1}{2}[\nabla\cdot(D^2f\nabla g + D^2g\nabla f) - \nabla U\cdot(D^2f\nabla g + D^2g\nabla f)\\
&- \nabla f\cdot(\nabla\Delta g - D^2U\cdot\nabla g - D^2g\cdot\nabla U)\\
&- \nabla g\cdot(\nabla\Delta f - D^2U\cdot\nabla f - D^2f\cdot\nabla U)]\\
=& \frac{1}{2}[\nabla\cdot(D^2f\nabla g + D^2g\nabla f) + 2\nabla f D^2U\nabla g - \nabla f\cdot\nabla\Delta g - \nabla g\cdot\nabla\Delta f]\ .
\end{aligned}
$$

Also,

$$
\begin{aligned}
\nabla\cdot(D^2f\nabla g + D^2g\nabla f) &= \frac{1}{2}\Big[\sum_j \partial_j\Big(\sum_i \partial^2_{ij}f\partial_i g + \partial^2_{ij}g\partial_i f\Big)\Big]\\
&= \sum_{i,j}\partial^3_{ijj}f\partial_i g + \partial^2_{ij}f\partial^2_{ij}g + \partial^3_{ijj}g\partial_i f + \partial^2_{ij}g\partial^2_{ij}f\\
&= \nabla\Delta f\cdot\nabla g + \nabla f\cdot\nabla\Delta g + 2D^2f : D^2g\ ,
\end{aligned}
$$

where : denotes the termwise product of matrices, that is $A : B := \sum_{i,j=1}^n a_{ij}b_{ij}$. This means that

$$
\Gamma_2(f,g) = D^2f : D^2g + \nabla f D^2U\nabla g\ .
$$

Let us now expose some general applications of Γ-calculus: indeed for $\Phi(f) = |\nabla f|^2$ the next result is classically due to Bakry and Émery in [1].

Proposition 10. *The pointwise exponential subcommutation*

$$
\Phi(P_tf) \le e^{-2\rho t}P_t[\Phi(f)]\ ,t\ge 0, f\in\mathcal{A}_+ \tag{3.1}
$$

for some $\rho\in\mathbb{R}$ is equivalent to the curvature condition

$$
\Gamma_\Phi \ge \rho\Phi\ . \tag{3.2}
$$

Proof. Indeed consider for fixed $t>0$, $f\ge 0$ and $x\in\mathbb{R}^n$ the function

$$
\psi_t(s) = P_s\Phi(P_{t-s}f)(x) \quad ,0\le s\le t\ .
$$

Then, by writing $\Phi(s,g) = P_s\Phi(g)$,

$$
\begin{aligned}
\psi_t'(s) &= \frac{d}{ds}[\Phi(s,P_{t-s}f)(x)] = L[P_s\Phi(P_{t-s}f)](x) + \mathrm{d}\Phi(s,P_{t-s}f).(-LP_{t-s}f)(x)\\
&= P_s[L(\Phi(P_{t-s}f)) - \mathrm{d}\Phi(P_{t-s}f).LP_{t-s}f](x) = 2P_s\Gamma_\Phi(P_{t-s}f)(x)
\end{aligned}
$$

from the commutation between L and P_s. Then the curvature condition (3.2) implies

$$
\psi_t'(s) \ge 2\rho P_s\Phi(P_{t-s}f) = 2\rho\psi_t(s)\ ,
$$

which integrated in s gives

$$
P_t\Phi(f)(x) = P_t\Phi(P_0f)(x) = \psi_t(t) \ge e^{2\rho t}\psi_t(0) = e^{2\rho t}P_0\Phi(P_tf)(x) = e^{2\rho t}\Phi(P_tf)(x)
$$

which is the subcommutation (3.1). Conversely, if (3.1) holds,

$$
\begin{aligned}
\mathrm{d}\Phi(f).Lf &= \frac{d}{dt}_{|t=0}\Phi(P_tf) = \lim_{t\to 0}\frac{\Phi(P_tf)-\Phi(f)}{t} \le \lim_{t\to 0}\frac{e^{-2\rho t}P_t\Phi(f)-\Phi(f)}{t}\\
&= \frac{d}{dt}_{|t=0}[e^{-2\rho t}P_t\Phi(f)] = -2\rho\Phi(f) + L\Phi(f)\ ,
\end{aligned}
$$

that is, $\Gamma_\Phi \ge 2\rho\Phi$. ∎

In order to get the link with our long-time behaviour problem, suppose P_t admits an *invariant* law μ, that is, such that for all $f \geq 0$, $t \geq 0$

$$\int_{\mathbb{R}^n} P_t f \, \mathrm{d}\mu = \int_{\mathbb{R}^n} f \, \mathrm{d}\mu$$

or, equivalently, $\int_{\mathbb{R}^n} Lf \, \mathrm{d}\mu = 0$ for $f \in \mathcal{A}$. Then integrating in space (3.1) gives

$$\int_{\mathbb{R}^n} \Phi(P_t f) \, \mathrm{d}\mu \leq e^{-2\rho t} \int_{\mathbb{R}^n} P_t[\Phi(f)] \, \mathrm{d}\mu = e^{-2\rho t} \int_{\mathbb{R}^n} \Phi(f) \, \mathrm{d}\mu$$

which is, upon proper choice of Φ, our goal in this essay.

Also, it is classical that

$$|\nabla P_t f| \leq e^{-\rho t} P_t \, |\nabla f|$$

for P_t linked with $L = \Delta - \nabla V \cdot \nabla$. By Jensen inequality applied to P_t it then follows that $|\nabla P_t f|^2 \leq e^{-2\rho t} P_t \, |\nabla f|$ which implies by Proposition 10 that it indeed holds $\Gamma_{|\nabla \cdot|^2} \geq \rho \, |\nabla \cdot|^2$.

These last facts should highlight the interest of establishing inequalities as the curvature condition.

Next, let us see a criterion to establish inequalities as (3.2): consider Φ_1 and Φ_2 gauge functions and $\gamma : \mathbb{R}_+ \to \mathbb{R}_+$ integrable. Suppose that

$$\Gamma_{\Phi_1}(P_t f) \leq \gamma(t) P_t \Phi_2(f) \tag{3.3}$$

for all $f \geq 0$ and $t \geq 0$. Suppose also that P_t is *ergodic*, that is, it admits a unique invariant law μ and

$$P_t f(x) \xrightarrow[t \to \infty]{} \int_{\mathbb{R}^n} f \, \mathrm{d}\mu$$

for all f and for all $x \in \mathbb{R}^n$. Then

$$\int_{\mathbb{R}^n} \Phi_1(f) \, \mathrm{d}\mu - \Phi_1\Big(\int_{\mathbb{R}^n} f \, \mathrm{d}\mu \Big) \leq 2 \Big(\int_0^\infty \gamma(s) \, \mathrm{d}s \Big) \int_{\mathbb{R}^n} \Phi_2(f) \, \mathrm{d}\mu$$

for all $f \geq 0$.

To prove this fix *una tantum* $x \in \mathbb{R}^n$, define as before $\psi_t(s) = P_s[\Phi_1(P_{t-s} f)](x)$ and compute

$$\begin{aligned}
P_t \Phi_1(f)(x) - \Phi_1(P_t f)(x) &= P_t \Phi_1(P_0 f)(x) - P_0[\Phi_1(P_t f)](x) = \psi_t(t) - \psi_t(0) \\
&= \int_0^t \psi_t'(s) \, \mathrm{d}s = 2 \int_0^t P_s \Gamma_\Phi(P_{t-s} f)(x) \, \mathrm{d}s \\
&\leq 2 \int_0^t P_s[\gamma(t-s) P_{t-s} \Phi_2(f)](x) \, \mathrm{d}s \\
&= 2 \Big(\int_0^t \gamma(s) \, \mathrm{d}s \Big) P_t \Phi_2(f)(x)
\end{aligned}$$

since P_s is sign-preserving. By letting $t \to \infty$, the left hand side gives

$$P_t \Phi_1(f)(x) - \Phi_1(P_t f)(x) \xrightarrow[t \to \infty]{} \int_{\mathbb{R}^n} \Phi_1(f) \, \mathrm{d}\mu - \Phi_1\Big(\int_{\mathbb{R}^n} f \, \mathrm{d}\mu \Big) \, ,$$

where the last term is to be meant as Φ_1 applied to the constant function $\int_{\mathbb{R}^n} f \, \mathrm{d}\mu$. In the same fashion,

$$P_t \Phi_1(f)(x) \xrightarrow[t \to \infty]{} \int_{\mathbb{R}^n} \Phi_1(f) \, \mathrm{d}\mu$$

giving thus our claim.

Notice that (3.3) is interesting in itself before time integration gives space globalization: for instance the Laplacian $L = \Delta$ satisfies of course $\Gamma_2 \geq 0$, giving that 0-curvature condition holds, for $\Phi = \Gamma$. Since

$$\Gamma_\Delta(f) = |\nabla f|^2 \ ,$$

$\Gamma_2 \geq 0$ is equivalent by Proposition 10 to

$$|\nabla P_t f|^2 \leq P_t \, |\nabla f|^2 \ .$$

This condition is (3.3) with $\gamma(t) = 1$, $\Phi_1(f) = f^2$ – that is, $\Gamma_{\Phi_1}(f) = \Gamma(f) = |\nabla f|^2$ – and $\Phi_2(f) = \Gamma(f)$. Then for all $t > 0$

$$P_t(f^2) - (P_t f)^2 \leq 2t P_t(|\nabla f|^2) \ .$$

Since in this case $\gamma(t) = t \notin L^1(\mathbb{R}_+)$, this information is not useful for large time. However for small t it provides a version of Poincaré inequality, where the integration is given by the conditional expectation of the Markov semigroup P_t.

This application with Γ_2 suggests to apply the result in the following Γ_2 setting: suppose that, in addition to ergodicity of P_t,

$$\Gamma_{\Gamma_{\Phi_1}} \geq \rho \Gamma_{\Phi_1} \tag{3.4}$$

with $\rho > 0$. Then we know this is equivalent to

$$\Gamma_{\Phi_1}(P_t f) \leq e^{-2\rho t} P_t[\Gamma_{\Phi_1}(f)] \ .$$

which is (3.3), with $\Phi_2 = \Gamma_{\Phi_1}$ and $\gamma(t) = e^{-2\rho t}$. Then

$$\int_{\mathbb{R}^n} \Phi_1(f) \, \mathrm{d}\mu - \Phi_1\biggl(\int_{\mathbb{R}^n} f \, \mathrm{d}\mu \biggr) \leq 2 \int_0^\infty e^{-2\rho s} \, \mathrm{d}s \int_{\mathbb{R}^n} \Gamma_{\Phi_1}(f) \, \mathrm{d}\mu = \frac{1}{\rho} \int_{\mathbb{R}^n} \Gamma_{\Phi_1}(f) \, \mathrm{d}\mu \ .$$

Let us highlight its character of necessarity for the generalized Γ_2 condition: take the semigroup

$$Lf(x) = Bx \cdot \nabla f(x) + \nabla \cdot (D \nabla f)(x)$$

where B is the linear drift field and D is the constant, positive semidefinite diffusion matrix. Then

$$\begin{aligned} \Gamma(f) &= \frac{L(f^2) - 2fLf}{2} = \frac{\nabla \cdot (D\nabla(f^2)) - 2f\nabla \cdot (D\nabla f)}{2} \\ &= \nabla \cdot (fD\nabla f) - f\nabla \cdot (D\nabla f) = \nabla f D \nabla f \ , \end{aligned}$$

where the drift does not appear since it is a simple derivation. Then, if D admits a nontrivial kernel – as it is the case for kinetic Fokker-Planck equation – there may not be a positive Γ_2 curvature, even without computing the Γ_2 functional. Indeed, if (3.4) held for some $\rho > 0$ and $\Phi_1(f) = f^2$, we would have also the modified Poincaré inequality

$$\int_{\mathbb{R}^n} f^2 \, \mathrm{d}\mu - \biggl(\int_{\mathbb{R}^n} f \, \mathrm{d}\mu \biggr)^2 \leq \frac{1}{\rho} \int_{\mathbb{R}^n} \Gamma(f) \, \mathrm{d}\mu = \frac{1}{\rho} \int_{\mathbb{R}^n} \nabla f D \nabla f \, \mathrm{d}\mu$$

for all $f \geq 0$. However this is not possible by just considering f not μ-a.e. constant such that $\nabla f(x) \in \ker D$ for all x – for instance pick $\psi : \mathbb{R} \to \mathbb{R}_+$ smooth, $w \in \ker D$ and consider $x \mapsto \psi(x \cdot w)$. Thus the left hand side is positive while the right hand side is null.

Let us now introduce *diffusion* semigroups by defining them as P_t generated by an operator L such that the following change of variable formulas hold:

(i) For all f and all $\psi \in C_c^\infty$

$$L(\psi(f)) = \psi'(f)Lf + \psi''(f)\Gamma(f,f)$$

(ii) For all f and g and all $\psi \in C_c^\infty$

$$\Gamma(\psi(f), g) = \psi'(f)\Gamma(f,g)$$

While this definition may look too abstract, the most common kind of generators – namely second order derivative operators – satisfy the diffusion property: fix

$$L = \sum_{i,j=1}^n a^{ij}(x)\partial_{ij}^2 + \sum_{j=1}^n b^j(x)\partial_j$$

with $a^{ij} = a^{ji}$ and b^j smooth coefficients. The constant coefficient is imposed to be null in order to keep the Markov property

$$L1 = 0 \ .$$

In this case

$$\begin{aligned}
\Gamma(f,g) &= \frac{L(fg) - fLg - gLf}{2} \\
&= \frac{1}{2}\Big[\sum_{i,j=1}^n a^{ij}\partial_{ij}^2(fg) - fa^{ij}\partial_{ij}^2 g - ga^{ij}\partial_{ij}^2 f + \sum_{j=1}^n b^j\partial_j(fg) - b^j f\partial_j g - b^j g\partial_j f\Big] \\
&= \sum_{i,j=1}^n a^{ij}\partial_i f\partial_j g \ .
\end{aligned}$$

Then, if $\psi \in C_c^\infty$, for all smooth f and g

$$\Gamma(\psi(f), g) = \sum_{i,j=1}^n a^{ij}\partial_i(\psi(f))\partial_j g = \sum_{i,j=1}^n a^{ij}\psi'(f)\partial_i\partial_j g = \psi'(f)\Gamma(f,g)$$

proving thus that (ii) holds, while concerning (i)

$$\begin{aligned}
L[\psi(f)] &= \sum_{i,j=1}^n a^{ij}\partial_{ij}^2[\psi(f)] + \sum_{j=1}^n b^j\partial_j[\psi(f)] \\
&= \sum_{i,j=1}^n a^{ij}\partial_i[\psi'(f)\partial_j f] + \sum_{j=1}^n b^j\psi'(f)\partial_j f \\
&= \sum_{i,j=1}^n a^{ij}[\psi''(f)\partial_i f\partial_j f + \psi'(f)\partial_{ij}^2 f] + \sum_{j=1}^n b^j\psi'(f)\partial_j f \\
&= \psi''(f)\Gamma(f) + \psi'(f)Lf
\end{aligned}$$

from the previous computation of Γ.

Back to general diffusion generators L, it is possible to establish a link between $\Gamma_{a(\cdot)}$, for convex $a : \mathbb{R} \to \mathbb{R}$, and Γ. Since

$$\mathrm{d}[a](f).g = \lim_{h\to 0}\frac{a(f+hg) - a(f)}{h} = a'(f)g \ ,$$

we have that

$$\Gamma_a(f) = \frac{L[a(f)] - \mathrm{d}[a](f).Lf}{2} = \frac{L[a(f)] - a'(f)Lf}{2} = \frac{a''(f)}{2}\Gamma(f) \ . \tag{3.5}$$

More generally, if Φ and a are given, we can apply (i) to the function $\Phi(f)$ and yield

$$\begin{aligned}
\Gamma_{a(\Phi(\cdot))}(f) &= \frac{L[a(\Phi(f))] - \mathrm{d}[a(\Phi)]f.Lf}{2} = \frac{L[a(\Phi(f))] - a'(\Phi(f))\mathrm{d}\Phi f.Lf}{2} \\
&= \frac{a'(\Phi(f))L\Phi(f) + a''(\Phi(f))\Gamma(\Phi(f)) - a'(\Phi(f))\mathrm{d}\Phi f.Lf}{2} \\
&= \frac{a''(\Phi(f))}{2}\Gamma(\Phi(f)) + a'(\Phi(f))\Gamma_\Phi(f) \ . \qquad (3.6)
\end{aligned}$$

since $\mathrm{d}[a(\Phi)] = a'(\Phi)\mathrm{d}\Phi$ through the previous simple reasoning. Setting $\Phi(f) = f$ gives (3.5), since $\Gamma_{\mathrm{id}} = 0$.

For instance, suppose that Φ satisfies the curvature condition $\Gamma_\Phi \geq \rho\Phi$ and that $\Phi \geq 0$. Then, for $p > 1$, (3.6) gives

$$\Gamma_{\Phi^p} f = \frac{p(p-1)}{2}\Phi^{p-2}(f)\Gamma(\Phi(f)) + p\Phi^{p-1}(f)\Gamma_\Phi(f) \geq p\Phi^{p-1}(f)\Gamma_\Phi(f) \geq p\rho\Phi^p(f)$$

since $\Gamma_\Phi \geq 0$. It follows that, as long as P_t is a diffusion semigroup,

$$\Phi^p(P_t f) \leq e^{-2p\rho t} P_t[\Phi^p(f)] \ ,$$

that is, Φ^p satisfies the curvature condition with constant $p\rho$. This is not new but is consistent with

$$\Phi(P_t f) \leq e^{-2\rho t} P_t[\Phi(f)] \leq e^{-2\rho t}[P_t(\Phi^p(f))]^{1/p}$$

which just comes from Proposition 10 and from Jensen inequality applied to the Markov semigroup P_t.

If $L = \sum_{j=1}^d A_j^2 + B$ where $A_j = a_j \cdot \nabla$ and $B = b \cdot \nabla$

$$\Gamma_L = \Gamma_B + \sum_{j=1}^d \Gamma_{A_j^2} = \sum_{j=1}^d \Gamma_{A_j^2}$$

since B is a derivation, and

$$\begin{aligned}
\Gamma_{A_j^2} f &= \frac{A_j^2(f^2)}{2} - f\ A_j^2 f = \frac{a_j \cdot \nabla(a_j \cdot \nabla f^2)}{2} - f\ a_j \cdot \nabla(a_j \cdot \nabla f) \\
&= a_j \cdot \nabla(f a_j \cdot \nabla f) - f\ a_j \cdot \nabla(a_j \cdot \nabla f = (a_j \cdot \nabla f)^2 = (A_j f)^2
\end{aligned}$$

so that

$$\Gamma_L f = |Af|^2 = A^T f\ Af \ .$$

3.2. **A general result of convergence for second-order equations.** Let $a : \mathbb{R}_+ \to \mathbb{R}$ be a non-strictly convex function and let $\nu \in \mathcal{P}_{ac}(\mathbb{R}^n)$. Then for all nonnegative $f : \mathbb{R}^n \to \mathbb{R}_+$ we may define the a-entropy with respect to ν as

$$\mathcal{E}_\nu^a(f) := \int_{\mathbb{R}^n} a(f)\,\mathrm{d}\nu - a\left(\int_{\mathbb{R}^n} f\,\mathrm{d}\nu\right)$$

which is of course nonnegative by Jensen inequality. In the following we will suppose that

- $a \in C^4(\mathbb{R}_+)$
- $a'' > 0$
- $(1/a'')'' \leq 0$.

Notice that this does not rule out neither the logarithmic entropy, for which $a(x) = x\log x$ and $1/a''(x) = x$, nor the quadratic case of $a(x) = x^2$, where $1/a'' = 1/2$. Indeed it is easy to prove that $a(x) = x^\alpha$ is admissible for $\alpha \geq 2$.

With these two main examples of gauge function in mind, we are in the position to state the main result of the essay, which is also general in L and corresponds to Theorem 10 of [5]. It will be easy to restrict to the kinetic Fokker-Planck case, and we shall indeed prove a rather technical part in our particular case only. This result will be subsequently used to provide a quantitatively better estimate in the convergence rate.

Theorem 11. *Let $L = \sum_{j=1}^r A_j^2 + B$ and suppose there exists $C_0 : \mathcal{A} \to \mathcal{A}^p$ such that, upon writing*

$$[B, C_i] = Z_{i+1}C_{i+1} + R_{i+1}$$

for some operator $0 < \lambda \leq Z_i \leq \Lambda$ and some remainder R_i, one can set, for some $I \geq 1$,

$$C_{I+1} = 0$$

such that for all i and j

(i)
$$[A_j, C_i] = 0$$

and in such a way that

(ii)
$$C_0^T C_0 \leq m_1 A^T A \ ,$$

for all $i \geq 1$

(iii)
$$R_i^T R_i \leq m_2 \sum_{j=0}^{i-1} C_j^T C_j$$

and

(iv)
$$\sum_{i=0}^{I} C_i^T C_i \geq \rho$$

which is as usual to be meant as $\sum_{i=1}^d \sum_{k=1}^d (c_{i,k}(x)\partial_k f(x))^2 \geq \rho\, |\nabla f(x)|^2$ for all f and all x. Suppose also that there exists a probability measure μ invariant under P_t such that the following inequality is satisfied, which should remind of logarithmic Sobolev inequality:

(v)
$$\mathcal{E}_\mu^a(f) \leq \frac{1}{K} \int_{\mathbb{R}^n} a''(f)\, |\nabla f|^2 \ d\mu \ .$$

Then there exists $C > 0$ such that for all f

$$\mathcal{E}_\mu^a(P_t f) \leq e^{-C\int_0^t (1-e^{-s})^{2I}\, ds}\, \mathcal{E}_\mu^a(f)$$

Remark 2. Let us compare this result with its very close analogue in Section 1, that is, Theorem 3: first, this one has as a strong point that it actually involves relative entropy, instead of a sum of it and Fisher information. Also, no geometric structure of the function space is required, so that there is no need of computing the adjoint of the operator A.

On the other hand, the convergence constant provided by this last Theorem – even if sharpened as in Subsection 3.3 – is incredibly small, of the order of 10^{-7}. In addition to this, the convergence is slightly worse than e^{-Ct}, even though just by a multiplicative constant. However the most remarkable issue is that condition (i) is quite stringent on coefficients of A: in most cases the only admissible configuration is where A has constant coefficients, which is luckily our case.

Proof. Take $\varepsilon_I > 0$ to be later chosen, and set $\varepsilon_i \in (0, \varepsilon_I)$ and $\lambda_i > 0$ to be later determined. Define

$$\Phi_{(t)}(f) := \sum_{i=0}^{I} \lambda_i \Phi_{i,(t)}(f)$$

where

$$\Phi_{0,(t)}(f) := a(f) + \varepsilon_0^2(1-e^{-t})a''(f)\,|C_0 f|^2 \ ,$$

while for $i \geq 1$

$$\Phi_{i,(t)}(f) := \varepsilon_i(1-e^{-t})^{2i-1}a''(f)\left|(C_{i-1}+\varepsilon_i(1-e^{-t})C_i)f\right|^2 \ .$$

We want to prove monotonicity, for all $t > 0$, of the function

$$s \mapsto \psi_t(s) := P_s\Phi_{(t-s)}(P_{t-s}f) \ .$$

Indeed

$$\begin{aligned}\psi_t(0) =& \Phi_{(t)}P_t f = a(P_t f) + a''(P_t f)\Big[\varepsilon_0^2(1-e^{-t})\,|C_0 P_t f|^2 \\ &+ \sum_{i\geq 1}\lambda_i\varepsilon_i(1-e^{-t})^{2i-1}\left|(C_{i-1}+\varepsilon_i(1-e^{-t})C_i)P_t f\right|^2\Big]\end{aligned}$$

and

$$\psi_t(t) = P_t\Phi_{(0)}(f) = P_t a(f) \ .$$

Since

$$\begin{aligned}\psi_t'(s) &= P_s[(L\Phi_{(t-s)} - \mathrm{d}\Phi_{(t-s)}.L - \partial_\tau\Phi_{(t-s)})(P_{t-s}f)] \\ &= P_s[(2\Gamma_{\Phi_{(t-s)}} - \partial_\tau\Phi_{(t-s)})(P_{t-s}f)] \ ,\end{aligned}$$

in order to prove that $\psi_t'(s) \geq 0$ it will be enough to show that, for all $r \geq 0$ and all function g, $[2\Gamma_{\Phi_{(r)}} - \partial_\tau\Phi_{(r)}](g) \geq 0$ and argue by sign-preservation of the semigroup. Indeed

$$\begin{aligned}[2\Gamma_{\Phi_{(r)}} - \partial_\tau\Phi_{(r)}](g) =& \sum_{i=0}^{I}\lambda_i(2\Gamma_{\Phi_{i,(r)}} - \partial_\tau\Phi_{i,(r)})(g) \\ \geq& a''(g)\sum_{i=0}^{I} - b_1\lambda_i\Big(\sum_{j=0}^{i-1}(1-e^{-r})^{2j}\,|C_j g|^2\Big) \\ &+ b_2\lambda_i\varepsilon_i^2(1-e^{-r})^{2i}\,|C_i g|^2 - b_3\lambda_i\varepsilon_i^4(1-e^{-r})^{2i+2}\,|C_{i+1}g|^2 \\ =& a''(g)\sum_{i=0}^{I}|C_i g|^2\,(1-e^{-r})^{2i}\Big[-b_1\Big(\sum_{j=i+1}^{I}\lambda_j\Big) \\ &+ b_2\lambda_i\varepsilon_i^2 - b_3\lambda_{i-1}\varepsilon_{i-1}^4\Big]\end{aligned}$$

thanks to Lemma 12 below, where we have set $\lambda_{-1}\varepsilon_{-1}^4 = 0$; as already stated, even if we are studying a general time r and function g, one may think them as t and $P_t f$ respectively. We want to bound the last term from below. Let us set

$\lambda_0 = 1$, $\lambda_i = \varepsilon_{i-1}^\alpha \lambda_{i-1}$ and $\varepsilon_{i-1} = \varepsilon_i^\beta$ with $\alpha > 0, \beta > 1$ to be determined. Then, if $\varepsilon_I^\alpha \leq 1/2$, with the crude estimate $\varepsilon_k^\alpha \leq \varepsilon_I^\alpha \leq 1/2$, for all $j \geq i+1$

$$\lambda_j = \lambda_{i+1} \prod_{k=i+1}^{j-1} \varepsilon_k^\alpha \leq \lambda_{i+1} 2^{-j+i+1} = \lambda_i \varepsilon_i^\alpha 2^{-j+i+1}$$

so that the term we are interested in gives

$$-b_1 \Big(\sum_{j=i+1}^{I} \lambda_j \Big) + b_2 \lambda_i \varepsilon_i^2 - b_3 \lambda_{i-1} \varepsilon_{i-1}^4 \geq \lambda_i \varepsilon_i^2 (-2b_1 \varepsilon_i^{\alpha-2} + b_2 - b_3 \varepsilon_i^{\beta(4-\alpha)-2}) \tag{3.7}$$

and we want this expression to be positive for all ε_i sufficiently small. Then the choices $2 < \alpha < 4$ and $\beta > 2/(4-\alpha)$, together with $b_2 > 0$, yield the existence of some $\varepsilon_* \in (0,1)$ and $c = c(\alpha, \beta, \varepsilon_*) > 0$ such that if all $\varepsilon_i \in (0, \varepsilon_*)$

$$[2\Gamma_{\Phi_{(r)}} - \partial_\tau \Phi_{(r)}](g) \geq c a''(g) \sum_{i=0}^{I} |C_i g|^2 (1 - e^{-r})^{2i} \lambda_i \varepsilon_i^2 \ , \tag{3.8}$$

for a suitable value of ε_I.

Our goal is turning (3.8) into a Gronwall inequality. This is motivated by the identity

$$\begin{aligned} \int_{\mathbb{R}^n} [2\Gamma_{\Phi_{(t)}} - \partial_t \Phi_{(t)}](P_t f) \, \mathrm{d}\mu &= - \int_{\mathbb{R}^n} [\mathrm{d}\Phi_{(t)}.L + \partial_t \Phi_{(t)}](P_t f) \, \mathrm{d}\mu \\ &= - \partial_t \int_{\mathbb{R}^n} [\Phi_{(t)}(P_t f)] \, \mathrm{d}\mu \end{aligned}$$

since μ being an invariant measure gives $\int_{\mathbb{R}^n} Lg \, \mathrm{d}\mu = 0$ for all g. In order to close the inequality let us integrate (3.8) with a crude bound on time

$$\begin{aligned} \int_{\mathbb{R}^n} [2\Gamma_{\Phi_{(t)}} - \partial_t \Phi_{(t)}](P_t f) \, \mathrm{d}\mu &\geq c \int_{\mathbb{R}^n} a''(P_t f) \sum_{i=0}^{I} |C_i P_t f|^2 (1 - e^{-t})^{2i} \lambda_i \varepsilon_i^2 \, \mathrm{d}\mu \\ &\geq c \lambda_I \varepsilon_I^2 (1 - e^{-t})^{2I} \int_{\mathbb{R}^n} a''(P_t f) \sum_{i=0}^{I} |C_i P_t f|^2 \, \mathrm{d}\mu \end{aligned}$$

where, in order to keep track of constants, we have used $\lambda_{i+1} \varepsilon_{i+1}^2 = \lambda_i \varepsilon_i^2 \varepsilon_i^{\alpha-2} \varepsilon_{i+1}^2 < \lambda_i \varepsilon_i^2$ since $\alpha > 2$ and $\varepsilon_k < 1$. Now for all $\delta \in (0,1)$ let us use δ times hypothesis (iv) to have

$$\begin{aligned} &\int_{\mathbb{R}^n} [2\Gamma_{\Phi_{(t)}} - \partial_t \Phi_{(t)}](P_t f) \, \mathrm{d}\mu \\ &\qquad \geq c \lambda_I \varepsilon_I^2 (1 - e^{-t})^{2I} \int_{\mathbb{R}^n} a''(P_t f) \Big[(1 - \delta) \sum_{i=0}^{I} |C_i P_t f|^2 + \delta \rho |\nabla P_t f|^2 \Big] \, \mathrm{d}\mu \ . \end{aligned}$$

On the Fisher-like term use (v), while for the first term use the bound

$$\eta \sum_{i=0}^{I} |c_i|^2 \geq (1 - e^{-t}) \varepsilon_0^2 |c_0|^2 + \sum_{i=1}^{I} \varepsilon_i (1 - e^{-t})^{2i-1} \left| c_{i-1} + (1 - e^{-t}) \varepsilon_i c_i \right|^2$$

for some $\eta = \eta(\beta) > 0$ as shown in the following, predictable, Lemma 14, so that

$$
\begin{aligned}
(1-\delta)\int_{\mathbb{R}^n} & a''(P_tf)\sum_{i=0}^{I}|C_iP_tf|^2 \,\mathrm{d}\mu \\
& \geq \frac{1-\delta}{\eta}\int_{\mathbb{R}^n} a''(P_tf)\Big[\varepsilon_0^2(1-e^{-t})\,|C_0P_tf|^2 \\
& \qquad + \sum_{i\geq 1}\varepsilon_i(1-e^{-t})^{2i-1}\left|(C_{i-1}+\varepsilon_i(1-e^{-t})C_i)P_tf\right|^2\Big]\,\mathrm{d}\mu\ .
\end{aligned}
$$

Then

$$
\begin{aligned}
\int_{\mathbb{R}^n}[2\Gamma_{\Phi_{(t)}}-\partial_t\Phi_{(t)}](P_tf)\,\mathrm{d}\mu \geq c\lambda_I\varepsilon_I^2(1-e^{-t})^{2I}\Big[&\delta\rho K\mathcal{E}_\mu^a(P_tf) \\
& + \frac{1-\delta}{\eta}\int_{\mathbb{R}^n} a''(P_tf)\Big(\varepsilon_0^2\,|C_0P_tf|^2 \\
& + \sum_{i\geq 1}\varepsilon_i(1-e^{-t})^{2i-1}\left|(C_{i-1}+(1-e^{-t})\varepsilon_iC_i)P_tf\right|^2\Big)\,\mathrm{d}\mu\Big]\ .
\end{aligned}
$$

Now, since $\lambda_i \leq \lambda_0 = 1$ for all i,

$$
\begin{aligned}
&\int_{\mathbb{R}^n}[2\Gamma_{\Phi_{(t)}}-\partial_t\Phi_{(t)}](P_tf)\,\mathrm{d}\mu \\
&\geq c\lambda_I\varepsilon_I^2(1-e^{-t})^{2I}\min\left(\delta\rho K,\frac{1-\delta}{\eta}\right)\left[\int_{\mathbb{R}^n}\Phi_{(t)}(P_tf)\,\mathrm{d}\mu - a\Big(\int_{\mathbb{R}^n}P_tf\,\mathrm{d}\mu\Big)\right]
\end{aligned}
$$

where the minimum, thanks to an optimization in δ, will be taken as the optimal value $\frac{\rho K}{\rho K\eta+1}$.

Now we only need a term as $a(\int_{\mathbb{R}^n}P_tf\,\mathrm{d}\mu)$ in $\int_{\mathbb{R}^n}\Phi_{(t)}(P_tf)\,\mathrm{d}\mu$, but notice that we can indeed change the functional into

$$
\int_{\mathbb{R}^n}\Phi_{(t)}(P_tf)\,\mathrm{d}\mu - a\Big(\int_{\mathbb{R}^n}P_tf\,\mathrm{d}\mu\Big) =: H(t)
$$

since $\partial_t a(\int_{\mathbb{R}^n}P_tf\,\mathrm{d}\mu) = 0$ by invariance of μ. Therefore, by letting $C = \frac{c\lambda_I\varepsilon_I^2\rho K}{1+\rho K\eta}$, we have obtained

$$
-H'(t) \geq C(1-e^{-t})^{2I}H(t)\ , \quad t\geq 0
$$

which means

$$
H(t) \leq H(0)e^{-C\int_0^t(1-e^{-s})^{2I}\,\mathrm{d}s}\ , \quad t \geq 0\ .
$$

It now suffices to notice that

$$
\begin{aligned}
H(t) = \int_{\mathbb{R}^n}&\Big(a(P_tf)+\varepsilon_0^2(1-e^{-t})a''(P_tf)\,|C_0P_tf|^2 \\
&+\sum_{i=1}^{I}\lambda_i\varepsilon_i(1-e^{-t})^{2i-1}a''(P_tf)\left|(C_{i-1}+\varepsilon_i(1-e^{-t})C_i)(P_tf)\right|^2\Big)\,\mathrm{d}\mu \\
&- a\Big(\int_{\mathbb{R}^n}P_tf\,\mathrm{d}\mu\Big) \\
\geq & \mathcal{E}_\mu^a(P_tf)
\end{aligned}
$$

and

$$H(0) = \mathcal{E}^a_\mu(f)$$

to conclude. $\square$

Remark 3 (A short-time effect). Let us focus on

$$[2\Gamma_{\Phi_{(r)}} - \partial_\tau \Phi_{(r)}](g) \geq 0$$

which gave, in the beginning of the proof of Theorem 11, that $\psi_t'(s) \geq 0$ for all $s \geq 0$. In particular from $\psi_t(t) \geq \psi_t(0)$ we get that, for all x,

$$\begin{aligned} P_t(a(f))(x) - a(P_t f)(x) \geq a''(P_t f)\Big[&\varepsilon_0^2(1-e^{-t})\left|C_0 P_t f\right|^2 \\ &+ \sum_{i\geq 1} \lambda_i \varepsilon_i (1-e^{-t})^{2i-1}\left|(C_{i-1} + \varepsilon_i(1-e^{-t})C_i)P_t f\right|^2\Big], \end{aligned}$$

where we have omitted dependence on x in the right hand side. Let now $t > 0$; in the quadratic form in C_i on the right hand side all coefficients of $|C_i|^2$ are multiplied by $(1-e^{-t})^{2i+1}$ and mixed terms by $(1-e^{-t})^{2i}$, so that the right hand side is indeed $(1-e^{-t})$ times a positive definite quadratic form of $(1-e^{-t})^i C_i$:

$$P_t(a(f))(x) - a(P_t f)(x) \geq \eta_3\ a''(P_t f)(1-e^{-t}) \sum_{i=0}^{I} (1-e^{-t})^{2i}\left|C_i P_t f\right|^2 . \tag{3.9}$$

where η_3 is linked to the quadratic form in a similar fashion of Lemma 14 and does not depend on time.

To get the regularizing meaning of this last part, take the Poincaré setting, with $a(f) = f^2$: for $t \to \infty$ we expect, from $P_t f(x) \to \int_{\mathbb{R}^n} f \,\mathrm{d}\mu$ for all x, that $C_i P_t f \to 0$, so that the inequality is not stronger than $\int_{\mathbb{R}^n} f^2 \,\mathrm{d}\mu - \left(\int_{\mathbb{R}^n} f \,\mathrm{d}\mu\right)^2 \geq 0$, that is the well-known Cauchy-Schwarz inequality. On the other hand the statement for small t becomes

$$\sum_{i=0}^{I} t^{2i}\left|C_i P_t f\right|^2 \leq C \frac{P_t(f^2) - (P_t f)^2}{t}$$

and since of course the semigroup maps L^2 into L^2, the left-hand side is finite, giving H^1 regularization.

The following result is crucial for Theorem 11 itself, but it shall be presented separately since it is rather technical. Also, part of the proof will be focused on the particular case of kinetic Fokker-Planck we have in exam.

Lemma 12. *Consider $\Phi_{i,(t)}$, $0 \leq i \leq I$ as defined in the proof of Theorem 11, supposing all hypotheses on L are satisfied. Then there exist $b_1, b_2, b_3 > 0$ and $\varepsilon_* \in (0,1)$ such that, for all $\varepsilon_i \in (0, \varepsilon_*)$ for $0 \leq i \leq I$*

$$\begin{aligned} (2\Gamma_{\Phi_i} - \partial_t \Phi_{i,(t)})(f) \geq a''(f)\Big[&-b_1\Big(\sum_{j=0}^{i-1}(1-e^{-t})^{2j}\left|C_j f\right|^2\Big) \\ &+ b_2\varepsilon_i^2(1-e^{-t})^{2i}\left|C_i f\right|^2 - b_3\varepsilon_i^4(1-e^{-t})^{2i+2}\left|C_{i+1} f\right|^2\Big]. \end{aligned}$$

Proof. For $i=0$ we want to prove

$$(2\Gamma_{\Phi_0} - \partial_t \Phi_{0,(t)})(f) \geq a''(f)\Big[b_2\varepsilon_0^2 \left|C_0 f\right|^2 - b_3\varepsilon_0^4(1-e^{-t})^2 \left|C_1 f\right|^2\Big] \ .$$

By linearity

$$\Gamma_{\Phi_0} = \Gamma_{a(\cdot)} + \varepsilon_0^2(1-e^{-t})\Gamma_{a''(\cdot)|C_0\cdot|^2} \ ;$$

For the first term we use diffusion property

$$2\Gamma_{a(\cdot)} f = a''(f)\Gamma f = a''(f)\left|Af\right|^2 \geq a''(f)\frac{\left|C_0 f\right|^2}{m_1} \ ,$$

while for the second Γ operator we shall use the following Lemma 13 to estimate

$$\Gamma_{a''(\cdot)|C_0\cdot|^2} f \geq a''(f)[L, C_0]f \cdot C_0 f \ .$$

Also, from $[A_j, C_0] = 0$ it also easily follows $[A_j^2, C_0] = 0$, so that

$$[L, C_0] = [B, C_0] = Z_1 C_1 + R_1 \ .$$

Last, for time-derivative, clearly

$$-\partial_t \Phi_{0,(t)} = -\varepsilon_0^2 e^{-t} a'' \left|C_0\right|^2 \ .$$

Then

$$\frac{2\Gamma_{\Phi_0} - \partial_t \Phi_{0,(t)}}{a''} \geq \frac{\left|C_0\right|^2}{m_1} + 2\varepsilon_0^2(1-e^{-t})C_0 \cdot (Z_1 C_1 + R_1) - \varepsilon_0^2 e^{-t}\left|C_0\right|^2$$

Now Young inequality, with a parameter δ to be soon determined, gives

$$\begin{aligned}
2C_0 \cdot \varepsilon_0^2(1-e^{-t})(Z_1 C_1 + R_1) &\geq -2\left|C_0\right| \cdot \varepsilon_0^2(1-e^{-t})(\left|Z_1 C_1\right| + \left|R_1\right|) \\
&\geq -\frac{\left|C_0\right|^2}{\delta} - \delta\varepsilon_0^4(1-e^{-t})^2(\left|Z_1 C_1\right|^2 + \left|R_1\right|^2) \\
&\geq -\frac{\left|C_0\right|^2}{c} - \delta\varepsilon_0^4(1-e^{-t})^2(\Lambda^2\left|C_1\right|^2 + m_2\left|C_0\right|^2)
\end{aligned}$$

where we also used initial hypotheses on Z_1 and R_1. Now let $\delta = 2m_1$ so that

$$\begin{aligned}
\frac{2\Gamma_{\Phi_{0,(t)}} - \partial_t \Phi_{0,(t)}}{a''} \geq & \left|C_0\right|^2\left(\frac{1}{2m_1} - 2m_1 m_2 \varepsilon_0^4(1-e^{-t})^2 - \varepsilon_0^2 e^{-t}\right) \\
& - 2m_1\Lambda^2\varepsilon_0^4(1-e^{-t})^2\left|C_1\right|^2 \\
\geq & \left|C_0\right|^2\left(\frac{1}{2m_1} - 2m_1 m_2\varepsilon_0^4 - \varepsilon_0^2\right) - 2\Lambda^2 m_1\varepsilon_0^4(1-e^{-t})^2\left|C_1\right|^2
\end{aligned}$$

and we are done if we force ε_0, via ε_*, to be small enough. In particular it is easy to show that we need

$$\varepsilon_0^2 < \frac{\sqrt{1+4m_2} - 1}{4m_1 m_2} \ .$$

Now let $i \geq 1$. By reminding that

$$\Phi_{i,(t)}(f) = \varepsilon_i(1-e^{-t})^{2i-1}a''(f)\left|(C_{i-1} + \varepsilon_i(1-e^{-t})C_i)f\right|^2 \ ,$$

we can use Lemma 13 with $C = C_{i-1} + \varepsilon_i(1-e^{-t})C_i$ and $b(f) = \varepsilon_i(1-e^{-t})^{2i-1}a''(f)$. Again $[L, C_i] = [B, C_i]$ from hypothesis (i) gives

$$\begin{aligned}
\Gamma_{\Phi_{i,(t)}} \geq & \varepsilon_i(1-e^{-t})^{2i-1}a''(C_{i-1} + \varepsilon_i(1-e^{-t})C_i) \cdot [L, C_{i-1} + \varepsilon_i(1-e^{-t})C_i] \\
= & \varepsilon_i(1-e^{-t})^{2i-1}a''(C_{i-1} + \varepsilon_i(1-e^{-t})C_i) \\
& \cdot (Z_i C_i + R_i + \varepsilon_i(1-e^{-t})(Z_{i+1}C_{i+1} + R_{i+1}))
\end{aligned}$$

while

$$
\begin{aligned}
-\partial_t \Phi_{i,(t)} =& -\varepsilon_i a'' \Big[(2i-1) e^{-t} (1-e^{-t})^{2i-2} \left| C_{i-1} + \varepsilon_i (1-e^{-t}) C_i \right|^2 \\
&+ (1-e^{-t})^{2i-1} 2\varepsilon_i e^{-t} C_i (C_{i-1} + \varepsilon_i (1-e^{-t}) C_i) \Big]
\end{aligned}
$$

so that

$$
\begin{aligned}
2\Gamma_{\Phi_{i,(t)}} - \partial_t \Phi_{i,(t)} \geq & \varepsilon_i a'' (1-e^{-t})^{2i-2} (C_{i-1} + \varepsilon_i (1-e^{-t}) C_i) \\
& \cdot \Big(2(1-e^{-t}) \big(Z_i C_i + R_i + \varepsilon_i (1-e^{-t})(Z_{i+1} C_{i+1} + R_{i+1}) \big) \\
& - (2i-1) e^{-t} (C_{i-1} + \varepsilon_i (1-e^{-t}) C_i) - 2(1-e^{-t}) \varepsilon_i e^{-t} C_i \Big) \ .
\end{aligned}
$$

From now on let us restrict ourselves to the kinetic Fokker-Planck case, as described in Subsection 3.3, where $i = I = 1$. We want to prove that

$$
(2\Gamma_{\Phi_1} - \partial_t \Phi_{1,(t)})(f) \geq a''(f) \Big[-b_1 \left| C_0 f \right|^2 + b_2 \varepsilon_1^2 (1-e^{-t})^2 \left| C_1 f \right|^2 \Big] \ .
$$

Kinetic Fokker-Planck amounts, as computed in Subsection 3.3, to $R_1 = \gamma \nabla_v = \gamma C_0$ and $|R_2| = \left| D_x^2 U \cdot \nabla_v \right| \leq c \left| C_0 \right|$ for some $c = C_U > 0$, so, just by applying these identities and expanding the product,

$$
\begin{aligned}
\frac{2\Gamma_{\Phi_{1,(t)}} - \partial_t \Phi_{1,(t)}}{a''} \geq & \varepsilon_1 \big(C_0 + \varepsilon_1 (1-e^{-t}) C_1 \big) \\
& \cdot \Big(2 \big(1-e^{-t}\big) \big(C_1 + \gamma C_0 + \varepsilon_1 (1-e^{-t}) R_2 \big) \\
& - e^{-t} \big(C_0 + \varepsilon_1 (1-e^{-t}) C_1 \big) - 2(1-e^{-t}) \varepsilon_1 e^{-t} C_1 \Big) \\
= & \varepsilon_1 \big(C_0 + \varepsilon_1 (1-e^{-t}) C_1 \big) \\
& \cdot \Big(\big(2\gamma (1-e^{-t}) - e^{-t} \big) C_0 + 2\varepsilon_1 (1-e^{-t})^2 R_2 \\
& + (1-e^{-t})(2 - 3\varepsilon_1 e^{-t}) C_1 \Big) \\
= & \varepsilon_1 \big(2\gamma (1-e^{-t}) - e^{-t} \big) \left| C_0 \right|^2 + \varepsilon_1^2 (1-e^{-t})^2 (2 - 3\varepsilon_1 e^{-t}) \left| C_1 \right|^2 \\
& + \varepsilon_1 (1-e^{-t}) \big((2 - 3\varepsilon_1 e^{-t}) + \varepsilon_1 \big(2\gamma (1-e^{-t}) - e^{-t} \big) \big) C_0 \cdot C_1 \\
& + 2\varepsilon_1^2 (1-e^{-t})^2 R_2 \big(C_0 + \varepsilon_1 (1-e^{-t}) C_1 \big) \\
\geq & \varepsilon_1 \big(2\gamma (1-e^{-t}) - e^{-t} - 2c\varepsilon_1 (1-e^{-t})^2 \big) \left| C_0 \right|^2 \\
& + \varepsilon_1^2 (1-e^{-t})^2 (2 - 3\varepsilon_1 e^{-t}) \left| C_1 \right|^2 \\
& - \varepsilon_1 (1-e^{-t}) \Big((2 - 3\varepsilon_1 e^{-t}) \\
& + \varepsilon_1 \left| 2\gamma (1-e^{-t}) - e^{-t} \right| + 2c\varepsilon_1^2 (1-e^{-t})^2 \Big) \left| C_0 \right| \left| C_1 \right| \ .
\end{aligned}
$$

We can already see that the coefficient of $\left| C_1 \right|^2$ has already the desired order, so we have to perform Young estimates on the mixed term in such a way that the current arrangement with orders is not perturbated. Of the three terms composing $\left| C_0 \right| \left| C_1 \right|$, the first one reads

$$
\begin{aligned}
-\varepsilon_1 (1-e^{-t})(2 - 3\varepsilon_1 e^{-t}) \left| C_0 \right| \left| C_1 \right| \geq & -\frac{1}{2} \varepsilon_1 (1-e^{-t}) \left| C_0 \right| \left| C_1 \right| \\
\geq & -\frac{1}{2} \left| C_0 \right|^2 - \frac{1}{8} \varepsilon_1^2 (1-e^{-t})^2 \left| C_1 \right|^2
\end{aligned}
$$

since $\varepsilon_i \le \frac{1}{2}$, say. Concerning the second one,

$$\begin{aligned} -\varepsilon_1^2(1-e^{-t})\left|2\gamma(1-e^{-t})-e^{-t}\right||C_0||C_1| &\ge -c_\gamma\varepsilon_1^2(1-e^{-t})|C_0||C_1| \\ &\ge -c_\gamma^2|C_0|^2-\frac{1}{4}\varepsilon_1^4(1-e^{-t})^2|C_1|^2 \end{aligned}$$

where

$$c_\gamma=\sup_{t>0}\left|2\gamma(1-e^{-t})-e^{-t}\right|=\max\{1,2\gamma\}=\begin{cases}1 & \text{if } \gamma<\frac{1}{2}\\ 2\gamma & \text{if } \gamma\ge\frac{1}{2}\end{cases}$$

while the third term gives

$$-2c\varepsilon_1^3(1-e^{-t})^3|C_0||C_1|\ge-\varepsilon_1^4(1-e^{-t})^2|C_1|^2-c^2\varepsilon_1^2(1-e^{-t})^4|C_0|^2$$

so that summing everything up we have

$$\begin{aligned} &\frac{2\Gamma_{\Phi_{1,(t)}}-\partial_t\Phi_{1,(t)}}{a''} \\ &\ge\left(2\varepsilon_1\gamma(1-e^{-t})-\varepsilon_1e^{-t}-2c\varepsilon_1^2(1-e^{-t})^2-\frac{1}{2}-c_\gamma^2-c^2\varepsilon_1^2(1-e^{-t})^4\right)|C_0|^2 \\ &\qquad+\varepsilon_1^2(1-e^{-t})^2\left(2-3\varepsilon_1e^{-t}-\frac{1}{8}-\frac{\varepsilon_1^2}{4}\right)|C_1|^2 \end{aligned}$$

and, by using $\varepsilon_i \le 1/2$,

$$2-3\varepsilon_1e^{-t}-\frac{1}{8}-\frac{\varepsilon_1^2}{4}\ge2-3\varepsilon_1-\frac{1}{8}-\frac{\varepsilon_1^2}{4}\ge\frac{5}{16}\ .$$

On the other hand the coefficient of $|C_0|^2$ is clearly bounded from below, namely by

$$-\varepsilon_1-2c\varepsilon_1^2-\frac{1}{2}-c_\gamma^2-c^2\varepsilon_1^2\ge-1-\frac{c}{2}-c_\gamma^2-\frac{c^2}{4}\ .$$

This concludes the proof for the particular case of kinetic Fokker-Planck. □

Remark 4. For $i=0$ we reported the proof in [5] for the sake of showing analogies with other indexes, but it was enough to notice that $\Gamma_{a''(\cdot)|C_0\cdot|^2}\ge0$ to conclude, instead of using Lemma 13.

Let us again focus on the case of $I=1$, and keep supposing that $R_1=\gamma C_0$ and $|R_2|\le C_U|C_0|$. Then we may choose

$$b_1=1+\frac{C_U}{2}+\frac{C_U^2}{4}+c_\gamma^2$$

since it only appears in the computations of Γ_{Φ_1}, and $b_3=0$ thanks to Remark 4. This last fact gives that we can widen our conditions on α and β to $\alpha>2$ and $\beta\ge1$. Concerning b_2, we should take into consideration $\frac{1}{2m_1}-2m_1m_2\varepsilon_0^4-\varepsilon_0^2$ and $\frac{5}{16}$. In order to bound the first expression away from 0 pick

$$\varepsilon_*\le\min\left\{\frac{c_{1,m}}{2},\frac{1}{2}\right\}$$

where $c_{1,m} := \sqrt{\frac{\sqrt{1+4m_2}-1}{4m_1 m_2}}$ is the positive root of $\frac{1}{2m_1} - 2m_1 m_2 \varepsilon_0^4 - \varepsilon_0^2 = 0$. Thus, from $\varepsilon_0 \leq \varepsilon_*$,

$$\begin{aligned}
\frac{1}{2m_1} - 2m_1 m_2 \varepsilon_0^4 - \varepsilon_0^2 &\geq \frac{1}{2m_1} - \frac{m_1 m_2}{8}\left(\frac{\sqrt{1+4m_2}-1}{4m_1 m_2}\right)^2 - \frac{\sqrt{1+4m_2}-1}{16 m_1 m_2} \\
&= \frac{1}{2m_1} + \frac{3 - 2m_2 - 3\sqrt{1+4m_2}}{64 m_1 m_2} \\
&= \frac{3 + 30 m_2 - 3\sqrt{1+4m_2}}{64 m_1 m_2} =: c_{2,m} \ .
\end{aligned}$$

We can then set $b_2 = \min\{c_{2,m}, \frac{5}{16}\}$. Notice as well that, from the proof of Theorem 11, we need ε_* to be small enough that for some c

$$-2b_1 \varepsilon_i^{\alpha-2} + b_2 - b_3 \varepsilon_i^{\beta(4-\alpha)-2} \geq c$$

as long that $\varepsilon_i \leq \varepsilon_*$. In particular we shall choose

$$\varepsilon_* \leq \left(\frac{b_2}{4b_1}\right)^{\frac{1}{\alpha-2}}$$

for some λ to be chosen later, so that

$$-2b_1 \varepsilon_1^{\alpha-2} + b_2 - b_3 \varepsilon_1^{\beta(4-\alpha)-2} = -2b_1 \varepsilon_1^{\alpha-2} + b_2 \geq -2b_1 \varepsilon_*^{\alpha-2} + b_2 = \frac{b_2}{2} =: c$$

and we may choose

$$\varepsilon_* = \min\left\{\frac{c_{1,m}}{2}, \frac{1}{2}, \left(\frac{b_2}{4b_1}\right)^{\frac{1}{\alpha-2}}\right\} .$$

Further choices will be made in Subsection 3.3 for our specific case.

In view of the proof of Lemma 12 we need the following result, which will be stated with a general function b instead of a''.

Lemma 13. *Let L be the generator of any diffusion semigroup, and let $b : \mathbb{R}_+ \to \mathbb{R}_+$ be a positive, C^2 function such that $\frac{1}{b}$ is concave. Let also C be a linear operator on a space of smooth functions $\mathcal{A}$ to some $\mathcal{A}^p$. Then for all f*

$$\Gamma_{b(\cdot)|C\cdot|^2} f \geq b(f) Cf \cdot [L, C] f \ .$$

Proof. Fix f and write

$$L(b(f)\,|Cf|^2) = |Cf|^2\, L(b(f)) + L(|Cf|^2) b(f) + 2\Gamma(b(f), |Cf|^2)$$

and by diffusion property (i)

$$L(b(f)) = b'(f)\ Lf + b''(f)\ \Gamma f$$

so that

$$L(b(f)\,|Cf|^2) = |Cf|^2\, b'(f)\ Lf + |Cf|^2\, b''(f)\ \Gamma f + L(|Cf|^2) b(f) + 2\Gamma(b(f), |Cf|^2) \ .$$

Also

$$\mathrm{d}(|C|^2\, b)(f) Lf = Lf\Big[\mathrm{d}(|C|^2)(f)\ b(f) + |Cf|^2\, b'(f)\Big] \ ,$$

so that

$$\begin{aligned}\Gamma_{b(\cdot)|C\cdot|^2} f =& \frac{1}{2}\Big[\, |Cf|^2\, b''(f)\ \Gamma f + L(|Cf|^2) b(f) + 2\Gamma(b(f), |Cf|^2) \\ & - Lf \mathrm{d}(|C|^2)(f)\ b(f)\Big] \\ =& \frac{1}{2}\, |Cf|^2\, b''(f)\ \Gamma f + b(f)\Gamma_{|C\cdot|^2} f + \Gamma(b(f), |Cf|^2)\ .\end{aligned}$$

Concerning the last term, let us apply diffusion property (ii) and

$$\Gamma(b(f), |Cf|^2) = b'(f) 2Cf \cdot \Gamma(Cf, f)\ .$$

For the second one let us compute

$$\begin{aligned}\Gamma_{|C|^2} f &= \frac{L(Cf\cdot Cf) - \mathrm{d}\,|C|^2\,(f)(Lf)}{2} = \frac{L(Cf\cdot Cf) - 2Cf\cdot CLf}{2} \\ &= \Gamma(Cf) - Cf\cdot CLf + Cf\cdot LCf = \Gamma(Cf) + Cf\cdot [L, C] f\ .\end{aligned}$$

Now let us write, since $(1/b)'' \leq 0$,

$$b'' = \left(\frac{1}{1/b}\right)'' = -\frac{(1/b)''}{(1/b)^2} + 2\frac{((1/b)')^2}{(1/b)^3} \geq 2\frac{((1/b)')^2}{(1/b)^3}$$

so that, reminding that $\Gamma f \geq 0$,

$$\begin{aligned}\Gamma_{b(\cdot)|C\cdot|^2} f \geq& \frac{((1/b)')^2}{(1/b)^3}\, |Cf|^2\, \Gamma f + b(f)[\Gamma(Cf) + Cf\cdot [L, C] f] + 2b'(f) Cf\cdot \Gamma(Cf, f) \\ =& \frac{((1/b)')^2}{(1/b)^3}\, |Cf|^2\, \Gamma f + b(f)[\Gamma(Cf) + Cf\cdot [L, C] f] \\ & - 2\frac{(1/b)'}{(1/b)^2}(f) Cf\cdot \Gamma(Cf, f)\ .\end{aligned}$$

Now we want to apply Cauchy-Schwarz inequality to $\Gamma(Cf, f)$ in order to get rid of the last term thanks to the first two ones. Indeed

$$\begin{aligned}-2\frac{(1/b)'}{(1/b)^2}(f) Cf\cdot \Gamma(Cf, f) &\geq -2\frac{|(1/b)'|}{(1/b)^2}(f)\, |Cf|\, |\Gamma(Cf, f)| \\ &\geq -2\frac{|(1/b)'|}{(1/b)^2}(f)\, |Cf|\, \sqrt{\Gamma(Cf)\Gamma f} \\ &\geq -\frac{((1/b)')^2}{(1/b)^3}(f)\, |Cf|^2\, \Gamma f - b(f)\Gamma(Cf)\end{aligned}$$

so that thanks to Young inequality in the last passage we have the initial claim. □

Remark 5. Notice that the inequality is rather sharp in the usual cases of logarithmic entropy and L^2 since, except for classical arithmetical estimates, we only used $(1/a'')'' \leq 0$. Indeed in these two cases $1/(a''(x))$ is respectively equal to x and $1/2$.

Lemma 14. *There exists $\eta = \eta_I(\beta) > 0$, independent from t, such that for all $c_i \in \mathbb{R}^d$, $0 \leq i \leq I$,*

$$(1 - e^{-t})\varepsilon_0^2\, |c_0|^2 + \sum_{i=1}^{I} \varepsilon_i (1 - e^{-t})^{2i-1}\, \left|c_{i-1} + (1 - e^{-t})\varepsilon_i c_i\right|^2 \leq \eta (1 - e^{-t}) \sum_{i=0}^{I} |c_i|^2$$

Proof.

$$
\begin{aligned}
&(1-e^{-t})\varepsilon_0^2 \left|c_0\right|^2 + \sum_{i=1}^{I} \varepsilon_i (1-e^{-t})^{2i-1} \left|c_{i-1} + (1-e^{-t})\varepsilon_i c_i\right|^2 \\
&\qquad \leq (1-e^{-t})\varepsilon_0^2 \left|c_0\right|^2 + 2\sum_{i=1}^{I} (1-e^{-t})^{2i-1} \big(\varepsilon_i \left|c_{i-1}\right|^2 + (1-e^{-t})^2 \varepsilon_i^3 \left|c_i\right|^2 \big) \\
&\qquad = (1-e^{-t})(\varepsilon_0^2 + 2\varepsilon_1) \left|c_0\right|^2 \\
&\qquad\quad + 2\sum_{i=1}^{I-1} (1-e^{-t})^{2i+1} (\varepsilon_i^3 + \varepsilon_{i+1}) \left|c_i\right|^2 + 2(1-e^{-t})^{2I+1} \varepsilon_I^3 \left|c_I\right|^2 \\
&\qquad \leq (1-e^{-t}) \Big[(\varepsilon_0^2 + 2\varepsilon_1) \left|c_0\right|^2 + 2\sum_{i=1}^{I-1} (\varepsilon_i^3 + \varepsilon_{i+1}) \left|c_i\right|^2 + 2\varepsilon_I^3 \left|c_I\right|^2 \Big]
\end{aligned}
$$

where the formula also holds, with no sum, for $I = 1$. We want to find the largest coefficient in it, so start by supposing that $I \geq 2$. Concerning terms in the sum notice that, since $\varepsilon_{i-1} = \varepsilon_i^\beta < \varepsilon_i$ where $\beta \geq 1$, $\varepsilon_{i-1}^3 + \varepsilon_i < \varepsilon_i^3 + \varepsilon_{i+1}$. Also, clearly $\varepsilon_I^3 < \varepsilon_I < \varepsilon_{I-1}^3 + \varepsilon_I$. Last let us compare coefficients of $\left|c_0\right|^2$ and $\left|c_1\right|^2$. For this suppose that $\beta \geq 2$, which does not rule out the standard case $\alpha = \beta = 3$. Then, reminding that $\varepsilon_i \leq 1/2$ for all i,

$$
\begin{aligned}
\varepsilon_0^2 + 2\varepsilon_1 &= \varepsilon_2^{2\beta^2} + 2\varepsilon_2^\beta \leq \varepsilon_2 (2^{-2\beta^2+1} + 2^{-\beta}) \leq \varepsilon_2 (2^{-7} + 1) < 2\varepsilon_2 < 2(\varepsilon_2 + \varepsilon_2^{3\beta}) \\
&= 2(\varepsilon_2 + \varepsilon_1^3)
\end{aligned}
$$

which implies that we can choose $\eta = 2(\varepsilon_I + \varepsilon_I^{3\beta})$.

However when $I = 1$ there is no such term in the sum, we end up with

$$\varepsilon_0^2 + 2\varepsilon_1 > 2\varepsilon_1^3$$

so that $\eta_1 = \varepsilon_1^{2\beta} + 2\varepsilon_1$. □

Remark 6. Let us explain the choice of the exponential term in Theorem 11: indeed we could have proved that

$$\mathcal{E}_\mu^a(P_t f) \leq e^{-C_\alpha \int_0^t (\alpha(s))^{2I}\, \mathrm{d}s} \mathcal{E}_\mu^a(f)$$

for all $\alpha(t)$ such that $\alpha(0) = 0$, $0 \leq \alpha(t) \leq 1$ and $|\alpha'(t)| \leq 1$.

3.3. **Application to kinetic Fokker-Planck equation.** Now let us focus our concrete case:

Here we take $A_j = \frac{\sigma}{\sqrt{2}} \partial_{v_j}$ and $B = -v \cdot \nabla_x + \nabla_x U \cdot \nabla_v - \gamma v \cdot \nabla_v$ in order to reach the $\partial_t = \sum_{j=1}^d A_j^2 + B$ form. Let us choose

$$C_0 = \nabla_v$$

and compute, reminding that $U = U(x)$, the commutator

$$
\begin{aligned}
[B, C_0]_i &= [-v \cdot \nabla_x + (\nabla_x U - \gamma v) \cdot \nabla_v, \partial_{v_i}] = [-v \cdot \nabla_x - \gamma v \cdot \nabla_v, \partial_{v_i}] \\
&= -v \cdot \nabla_x \partial_{v_i} + \partial_{v_i}(v \cdot \nabla_x) - \gamma v \cdot \nabla_v \partial_{v_i} + \partial_{v_i}(\gamma v \cdot \nabla_v) = \partial_{x_i} + \gamma \partial_{v_i} \ .
\end{aligned}
$$

Let us set then $C_1 = \nabla_x$, $Z_1 = \mathrm{Id}$ and $R_1 = \gamma \nabla_v$. Also

$$
\begin{aligned}
[B, C_1]_i &= [-v \cdot \nabla_x + (\nabla_x U - \gamma v) \cdot \nabla_v, \partial_{x_i}] = [\nabla_x U \cdot \nabla_v, \partial_{x_i}] \\
&= \nabla_x U \cdot \nabla_v \partial_{x_i} - \sum_{j=1}^{d} \partial_{x_i} (\partial_{x_j} U \partial_{v_j}) = - \sum_{j=1}^{d} \partial^2_{x_i x_j} U \partial_{v_j}
\end{aligned}
$$

so that we can choose $C_2 = 0$ and $R_2 = -D_x^2 U \cdot \nabla_v$ with, of course, $Z_2 = \mathrm{Id}$.

One can now check hypotheses from Theorem 11 with $I = 1$, since of course

$$
[A_j, C_0] = [\partial_{v_j}, \nabla_v] = 0 = [\partial_{v_j}, \nabla_x] = [A_j, C_1] \ ,
$$

$$
C_0^T C_0 = |\nabla_v|^2 = \frac{2}{\sigma^2} A^T A \ ,
$$

and

$$
C_0^T C_0 + C_1^T C_1 = |\nabla_x|^2 + |\nabla_v|^2 = |\nabla|^2
$$

so that (i), (ii) and (iv) are satisfied with $m_1 = \frac{2}{\sigma^2}$ and $\rho = 1$. Concerning (iii),

$$
R_1^T R_1 = \gamma^2 \nabla_v^T \nabla_v = \gamma^2 C_0^T C_0
$$

and

$$
R_2^T R_2 = \nabla_v \cdot [D_x^2 U]^2 \cdot \nabla_v \ .
$$

This last equality tells us that that, unfortunately, in order to have hypothesis (iii) fulfilled it is strictly necessary to suppose that there exists some $C_U > 0$ such that for all $x \in \mathbb{R}^d$

$$
\left| w \cdot D_x^2 U(x) \cdot w \right| \leq C_U \, |w|^2 \quad , \quad w \in \mathbb{R}^d \ ,
$$

as in [9]. We can then set $m_2 := \max\{\gamma^2, C_U^2\}$.

Next let us focus on conditions on a and μ, included in hypothesis (v): we shall choose, up to a normalizing constant,

$$
\mu = f_\infty(x, v) = e^{-\frac{2\gamma}{\sigma^2}\left(U(x) + \frac{|v|^2}{2}\right)} \ ,
$$

which is clearly invariant under P_t, and $a(x) = x \log x$. Thus

$$
\mathcal{E}_\mu^a(f) = \int_{\mathbb{R}_x^d \times \mathbb{R}_v^d} f \log f \, \mathrm{d}\mu - \int_{\mathbb{R}_x^d \times \mathbb{R}_v^d} f \, \mathrm{d}\mu \ \log \left(\int_{\mathbb{R}_x^d \times \mathbb{R}_v^d} f \, \mathrm{d}\mu \right) \quad , \quad f \in \mathcal{A}_+
$$

but since $f\mu$ satisfies a Fokker-Planck equation in divergence form $\int_{\mathbb{R}^n} f \, \mathrm{d}\mu$ is conserved, so by choosing $f \in \mathcal{P}(\mathrm{d}\mu)$ we have $\mathcal{E}_\mu^{f \log f} = H_\mu$, the Kullback information with respect to μ. We then want to prove that the product measure $e^{-\frac{2\gamma}{\sigma^2} U(x)} \otimes e^{-\frac{2\gamma}{\sigma^2} \frac{|v|^2}{2}}$ satisfies a logarithmic Sobolev inequality. Indeed, it is well-known that Logarithmic Sobolev Inequality tensorizes and that we just need to find the smallest between the two constants.

Proposition 15. *Let ν_λ be the gaussian measure on $\mathbb{R}^n$ with mean 0 and variance λ^2. Then ν_λ satisfies a logarithmic Sobolev inequality*

$$
\int_{\mathbb{R}^n} f \log f \, d\nu_\lambda \leq \frac{\lambda^2}{2} \int_{\mathbb{R}^n} \frac{|\nabla f|^2}{f} \, d\nu_\lambda
$$

for all $f \in \mathcal{P}(d\nu_\lambda)$.

Proof. Let us deal with $n = 1$ and then argue by tensorization. By reminding that the standard gaussian measure satisfies a logarithmic Sobolev inequality with constant 1, that is, for all $f \geq 0$,

$$\int_{\mathbb{R}^n} f(x)\log f(x) e^{-\frac{|x|^2}{2}}\,\mathrm{d}x - \int_{\mathbb{R}^n} f(x) e^{-\frac{|x|^2}{2}}\,\mathrm{d}x \log\left(\int_{\mathbb{R}^n} f(x) e^{-\frac{|x|^2}{2}}\,\mathrm{d}x\right)$$
$$\leq \frac{1}{2}\int_{\mathbb{R}^n} \frac{|\nabla f(x)|^2}{f(x)} e^{-\frac{|x|^2}{2}}\,\mathrm{d}x$$

fix $\lambda > 0$ and take f such that

$$\int_{\mathbb{R}^n} \frac{f(x)e^{-\frac{|x|^2}{2\lambda^2}}}{\sqrt{2\pi}\lambda}\,\mathrm{d}x = 1\ .$$

Then

$$\int_{\mathbb{R}^n} f(x)\log f(x)\frac{e^{-\frac{|x|^2}{2\lambda^2}}}{\sqrt{2\pi}\lambda}\,\mathrm{d}x = \int_{\mathbb{R}^n} f(\lambda y)\log f(\lambda y)\frac{e^{-\frac{|y|^2}{2}}}{\sqrt{2\pi}}\,\mathrm{d}y$$
$$\leq \frac{1}{2}\int_{\mathbb{R}^n} \frac{|\nabla_y f(\lambda y)|^2}{f(\lambda y)}\frac{e^{-\frac{|y|^2}{2}}}{\sqrt{2\pi}}\,\mathrm{d}y + \int_{\mathbb{R}^n} f(\lambda y)\frac{e^{-\frac{|y|^2}{2}}}{\sqrt{2\pi}}\,\mathrm{d}y\ \log\left(\int_{\mathbb{R}^n} f(\lambda y)\frac{e^{-\frac{|y|^2}{2}}}{\sqrt{2\pi}}\,\mathrm{d}y\right)\ .$$

Since

$$\int_{\mathbb{R}^n} f(\lambda y)\frac{e^{-\frac{|y|^2}{2}}}{\sqrt{2\pi}}\,\mathrm{d}y = \int_{\mathbb{R}^n} f(y)\frac{e^{-\frac{|y|^2}{2\lambda^2}}}{\sqrt{2\pi}\lambda}\,\mathrm{d}y = 1\ ,$$

we have

$$\begin{aligned}\int_{\mathbb{R}^n} f(x)\log f(x)\frac{e^{-\frac{|x|^2}{2\lambda^2}}}{\sqrt{2\pi}\lambda}\,\mathrm{d}x &\leq \frac{1}{2}\int_{\mathbb{R}^n} \frac{|\nabla_y f(\lambda y)|^2}{f(\lambda y)}\frac{e^{-\frac{|y|^2}{2}}}{\sqrt{2\pi}}\,\mathrm{d}y \\ &= \frac{\lambda^2}{2}\int_{\mathbb{R}^n} \frac{|\nabla_{\lambda y} f(\lambda y)|^2}{f(\lambda y)}\frac{e^{-\frac{|y|^2}{2}}}{\sqrt{2\pi}}\,\mathrm{d}y \\ &= \frac{\lambda^2}{2}\int_{\mathbb{R}^n} \frac{|\nabla_x f(x)|^2}{f(x)}\frac{e^{-\frac{|x|^2}{2\lambda^2}}}{\sqrt{2\pi}\lambda}\,\mathrm{d}x\ .\end{aligned}$$

□

The gaussian $e^{-\frac{2\gamma}{\sigma^2}\frac{|v|^2}{2}} = \nu_{\sigma^2/2\gamma}$ satisfies therefore a logarithmic Sobolev inequality with constant $\frac{\lambda^2}{2} = \frac{\sigma^2}{4\gamma}$, that is,

$$\int_{\mathbb{R}^d_v} f\log f\,\mathrm{d}\nu_{\sigma^2/2\gamma} \leq \frac{\sigma^2}{4\gamma}\int_{\mathbb{R}^d_v}\frac{|\nabla f|^2}{f}\,\mathrm{d}\nu_{\sigma^2/2\gamma}$$

for all $f \in \mathcal{P}(d\nu_{\sigma^2/2\gamma})$. Concerning $e^{-\frac{2\gamma}{\sigma^2}U(x)}$, we wish to apply Bakry-Émery criterion: we shall therefore suppose that $U \in C^2(\mathbb{R}^d)$ with

$$w\cdot D^2U(x)w \geq c_U\,|w|^2\ ,\ w\in\mathbb{R}^d\ ,$$

uniformly in $x \in \mathbb{R}^d$. This gives that $e^{-\frac{2\gamma}{\sigma^2}U(x)}$ satisfies a logarithmic Sobolev inequality with constant $\frac{\sigma^2}{4\gamma c_U}$, so that for all $f \in \mathcal{P}(\mu)$

$$\begin{aligned}\int_{\mathbb{R}^d_x\times\mathbb{R}^d_v} f\log f \,\mathrm{d}\mu \le& \max\left\{\frac{\sigma^2}{4\gamma c_U}, \frac{\sigma^2}{4\gamma}\right\}\int_{\mathbb{R}^d_x\times\mathbb{R}^d_v}\frac{|\nabla f|^2}{f}\,\mathrm{d}\mu\\ =&\frac{\sigma^2}{4\gamma}\frac{1}{\min\{1,c_U\}}\int_{\mathbb{R}^d_x\times\mathbb{R}^d_v}\frac{|\nabla f|^2}{f}\,\mathrm{d}\mu\ ,\end{aligned}$$

that is, $K = \frac{4\gamma}{\sigma^2}\min\{1, c_U\}$. The appearance of the absolute constant 1 should not surprise, since it is linked with the transport term.

We are now able to compute explicitly in this case the constant C of Theorem 11, that is, such that for all f

$$\mathcal{E}^a_\mu(P_t f) \le e^{-C\int_0^t (1-e^{-s})^2\,\mathrm{d}s}\mathcal{E}^a_\mu(f)\ .$$

Let us recall that, for general I,

$$C = \frac{c\lambda_I\varepsilon_I^2\rho K}{1+\rho K\eta}\ ,$$

and that in this setting

$$\lambda_I = \lambda_1 = \varepsilon_0^\alpha\lambda_0 = \varepsilon_1^{\alpha\beta}$$

while $\eta = \varepsilon_1^{2\beta} + 2\varepsilon_1$ as shown in the proof of Lemma 14. Thus, by relabelling $\min\{1, c_U\}$ with c_U,

$$C = \frac{4\gamma c\varepsilon_1^{2+\alpha\beta}c_U}{\sigma^2 + 4\gamma(\varepsilon_1^{2\beta}+2\varepsilon_1)c_U}\ .$$

It is then clear that β should be as small as possible, namely $\beta = 1$.

We now only need to find $c = c(\alpha, \beta, \varepsilon_*, \{b_i\}_i)$ as in the proof of Theorem 11 and ε_* in order to be done. We have already proved that in this case we may set

$$c = \frac{b_2}{2}$$

and

$$\begin{aligned}b_2 =& \min\left\{\frac{5}{16}, c_{2,m}\right\} = \min\left\{\frac{5}{16}, \frac{3+30m_2-3\sqrt{1+4m_2}}{64m_1m_2}\right\}\\ =& \min\left\{\frac{5}{16}, \frac{\sigma^2}{2}\frac{3+30m_2-3\sqrt{1+4m_2}}{64m_2}\right\} =: \min\left\{\frac{5}{16}, \frac{3\sigma^2}{128}c_{3,m}\right\}\ .\end{aligned}$$

It is elementary to prove that $8 < c_{3,m} < 10$ for all $\max\{\gamma^2, C_U^2\} = m_2 > 0$, therefore upon picking $\sigma^2 \ge \frac{5}{3}$ we can suppose

$$b_2 = \frac{5}{16}$$

and so

$$c = \frac{5}{32}\ .$$

Now let us take into consideration the value taken by

$$\varepsilon_* = \min\left\{\frac{1}{2}, \frac{c_{1,m}}{2}, \left(\frac{b_2}{4b_1}\right)^{\frac{1}{\alpha-2}}\right\} =: \min\left\{\frac{1}{2}, \varepsilon_{(*,1)}, \varepsilon_{(*,2)}\right\}$$

by analysing the different items. Since $m_1 = \frac{2}{\sigma^2}$,

$$\varepsilon_{(*,1)} = \frac{1}{4}\sqrt{\frac{\sqrt{1+4m_2}-1}{m_1 m_2}} = \frac{\sigma}{4\sqrt{2}}\sqrt{\frac{\sqrt{1+4m_2}-1}{m_2}}$$

and it is easy to notice that $0 < \sqrt{\frac{\sqrt{1+4m_2}-1}{m_2}} < \sqrt{2}$ and that it is a decreasing function of $m_2 = \max\{\gamma^2, C_U^2\}$. Last,

$$\varepsilon_{(*,2)} = \left(\frac{b_2}{4b_1}\right)^{\frac{1}{\alpha-2}} = \left(\frac{5}{64\big(1+\frac{C_U}{2}+\frac{C_U^2}{4}+c_\gamma^2\big)}\right)^{\frac{1}{\alpha-2}} .$$

Since $c_\gamma = \max\{1, 2\gamma\} \geq 2\gamma$, we have

$$C_U + \frac{C_U^2}{4} + c_\gamma^2 \geq \frac{m_2}{4}$$

whence, by reminding that $\alpha > 2$,

$$\varepsilon_{(*,2)} \leq \left(\frac{5}{64\big(1+\frac{m_2}{4}\big)}\right)^{\frac{1}{\alpha-2}} =: \tilde{\varepsilon}_{(*,2)} .$$

Let us now choose $\alpha = 4$ – which is an admissible choice because $b_3 = 0$ – and compare $\varepsilon_{(*,1)}$ and $\tilde{\varepsilon}_{(*,2)}$. Then it is elementary to prove that

$$\tilde{\varepsilon}_{(*,2)} \leq \varepsilon_{(*,1)}$$

for all choice of $m_2 > 0$ as long as $\sigma^2 \geq \frac{\sqrt{19}-2}{3}$. Since this last value is clearly smaller than $\frac{5}{3}$ and $\varepsilon_{(*,2)} \leq \sqrt{\frac{5}{64}} < \frac{1}{2}$, we may pick

(3.10)
$$\varepsilon_* = \varepsilon_{(*,2)} = \sqrt{\frac{5}{64\big(1+\frac{C_U}{2}+\frac{C_U^2}{4}+c_\gamma^2\big)}} = \sqrt{\frac{5}{64\big(1+\frac{C_U}{2}+\frac{C_U^2}{4}+\max\{1,4\gamma^2\}\big)}} .$$

We have therefore proved the following

Theorem 16. *Consider the equation*

$$\partial_t f + v\cdot\nabla_x f - \nabla_x U\cdot\nabla_v f = \frac{\sigma^2}{2}\Delta_v f + \gamma\nabla_v\cdot(vf) , \quad t \geq 0,\ (x,v)\in\mathbb{R}^d\times\mathbb{R}^d ,$$

and suppose that $\sigma^2 \geq \frac{5}{3}$, $\gamma > 0$ and $U \in C^2(\mathbb{R}^d_x)$ such that

$$c_U \leq D_x^2 U(x) \leq C_U , \ x \in \mathbb{R}^d ,$$

where $0 < c_U \leq 1$ and $C_U < \infty$ are constants which do not depend on x. Then, by writing f_t as the evolution at time t of f_0, for all $t > 0$ and for all $f_0 \in \mathcal{P}(\,d\mu)$

$$\int_{\mathbb{R}^d_x\times\mathbb{R}^d_v} f_t \log\left(\frac{df_t}{d\mu}\right) d\mu \leq e^{-C\int_0^t (1-e^{-s})^2\,ds}\int_{\mathbb{R}^d_x\times\mathbb{R}^d_v} f_0\log\left(\frac{df_0}{d\mu}\right) d\mu .$$

where

(3.11)
$$C = \frac{5\gamma\varepsilon_*^6 c_U}{8\big(\sigma^2 + 4\gamma(\varepsilon_*^2 + 2\varepsilon_*)c_U\big)}$$

and $\varepsilon_* = \varepsilon_*(C_U,\gamma)$ is under the form (3.10).

In particular notice that the standard case $\sigma^2 = 2, \gamma = 1$ with the quadratic potential $U(x) = \frac{|x|^2}{2}$, for which $c_U = C_U = 1$, is covered by our assumptions. Then $\varepsilon_* = \frac{\sqrt{5}}{4\sqrt{23}} = 0.11656...$ and $C = 5.2491... \cdot 10^{-7}$, which is considerably better than the order 10^{-46} that is obtained in [6].

4. Entropic convergence for self-consistent Vlasov-Fokker-Planck equation by particle method

In this Section we shall provide a result on Wasserstein-2 hypocoercivity result for a self-interacting system, and we will closely follow [6].
We shall consider the Vlasov-Fokker-Planck equation, namely

$$\partial_t f_t + v\cdot\nabla_x f_t = \frac{\sigma^2}{2}\Delta_v f_t + \nabla_v\cdot\left(\gamma f_t v + f_t\int_{\mathbb{R}^d_x\times\mathbb{R}^d_v}\nabla_x\mathcal{U}(x,y)f_t(y,v)\,\mathrm{d}y\,\mathrm{d}v\right), \tag{4.1}$$

where $\gamma,\sigma>0$, $x,v\in\mathbb{R}^d$, $f_t=f_t(x,v)$ denotes the unknown density of particles at time $t>0$ and $f_0\in\mathcal{P}_{ac}(\mathbb{R}^d_x\times\mathbb{R}^d_v)$ is given. Here $\mathcal{U}:\mathbb{R}^d_x\times\mathbb{R}^d_y\to\mathbb{R}$ is a two-variable potential modelling external force in the first variable x and self-interaction in the second variable y. If $\mathcal{U}=U(x)$ this reduces to kinetic Fokker-Planck equation with external potential $\mathcal{U}$, by mass conservation from the divergence form. The existence of an equilibrium measure f_∞ may be proved as in Proposition 2 of [3], which relies on the compactness argument in Proposition 3.1 in [8].
Analogously to it, one can see f_t as the law at time t of the stochastic process $(\bar X_t,\bar V_t)$ in $\mathbb{R}^d_x\times\mathbb{R}^d_v$ where

$$\begin{cases}\mathrm{d}\bar X_t &= \bar V_t\,\mathrm{d}t\\ \mathrm{d}\bar V_t &= -\gamma\bar V_t\,\mathrm{d}t - \left(\int\nabla_x\mathcal{U}(\bar X_t,y)f_t(y,v)\,\mathrm{d}y\,\mathrm{d}v\right)\mathrm{d}t+\sigma\,\mathrm{d}B_t\end{cases} \tag{4.2}$$

and $(\bar X_0,\bar V_0)$ has law f_0. As usual B_t denotes a d-dimensional Brownian motion. In this Section we will study long-time behaviour of densities f_t by approximating tensorized solutions $(\bar X,\bar V)^{\otimes N}$ of (4.2) with $(X,V)\in\mathbb{R}^{2dN}$, where for all $1\le i\le N$ $(X^i_t,V^i_t)\in\mathbb{R}^{2d}$ solves

$$\begin{cases}\mathrm{d}X^i_t &= V^i_t\,\mathrm{d}t\\ \mathrm{d}V^i_t &= -\gamma V^i_t\,\mathrm{d}t - \frac{1}{N}\sum_{j=1}^N\nabla_x\mathcal{U}(X^i_t,X^j_t)\,\mathrm{d}t+\sigma\,\mathrm{d}B^i_t\end{cases}, \tag{4.3}$$

where $(X^i_0,V^i_0)\in\mathbb{R}^{2d}$ are independent and identically distributed as f_0, and $\{B^i_t\}_i$ are N independent d-dimensional Brownian motions. We shall call $f^{(N)}_t:\mathbb{R}^{2dN}\to\mathbb{R}_+$ the law of (X_t,V_t) such that, as seen in Section 1, $f^{(N)}_t$ satisfies a diffusion equation, namely

$$\partial_t f^{(N)}_t + v\cdot\nabla_x f^{(N)}_t - \frac{1}{N}\sum_{i,j=1}^N\nabla_x\mathcal{U}(x_i,x_j)\cdot\nabla_{v_i}f^{(N)}_t = \frac{\sigma^2}{2}\Delta_v f^{(N)}_t + \gamma\nabla_v\cdot(vf^{(N)}_t) \tag{4.4}$$

with $f^{(N)}_0=f_0^{\otimes N}$. This means that, upon calling $\mathcal{U}_N(x):=\frac{1}{N}\sum_{i,j=1}^N\nabla_x\mathcal{U}(x_i,x_j):\mathbb{R}^{dN}\to\mathbb{R}$, the equilibrium of (4.4) is given, up to a constant, by

$$f^{(N)}_\infty(x,v)=e^{-\frac{2\gamma}{\sigma^2}\left(\mathcal{U}_N(x)+\frac{|v|^2}{2}\right)}.$$

Throughout this Section we will consider potentials

$$\mathcal{U}(x,y) = U(x) + W(x-y) + U(y)$$

and we will always suppose the following:

(i) U is smooth, and there exists a constant $c_U > 0$ independent from x such that

$$D^2U(x) \geq c_U$$

in the sense that $z \cdot D^2U(x) \cdot z \geq c_U |z|^2$ for all $z \in \mathbb{R}^d$ and for all $x \in \mathbb{R}^d$, and there exists a finite constant C_U such that

$$\left\| D^2U(x) \right\| \leq C_U \ , \quad x \in \mathbb{R}^d$$

for some norm $\|\cdot\|$ on $\mathbb{R}^{d\times d}$;

(ii) W is even, smooth and there exists a constant $c_W \in \mathbb{R}$ with

$$c_W < \frac{c_U}{2}$$

such that

$$D^2W \geq -c_W$$

and a constant C_W such that

$$\left\| D^2W(x) \right\| \leq C_W \ , \quad x \in \mathbb{R}^d \ .$$

In this Section one of the main features we will be dealing with is Wasserstein distance, which has been increasingly used in the last two decades. We shall recall here main features, for – much – more information one may view, for instance Chapter 7 in [10] or Chapter 5 in [7]: given μ and ν probability measures on $\mathbb{R}^n$ and $1 \leq p < \infty$ define

$$\begin{aligned} W_p(\mu,\nu) &:= \inf\left\{ \left(\int_{\mathbb{R}^n\times\mathbb{R}^n} |x-y|^p \, \mathrm{d}\gamma(x,y) \right)^{\frac{1}{p}} : \gamma \in \mathcal{P}(\mu,\nu) \right\} \\ &= \inf\left\{ \left[\mathbb{E}\big(|X-Y|^p \big) \right]^{\frac{1}{p}} : \mathrm{law}(X) = \mu \ , \quad \mathrm{law}(Y) = \nu \right\} \end{aligned}$$

where with $\mathcal{P}(\mu,\nu)$ we denote the set of probability measures on $\mathbb{R}^n_x \times \mathbb{R}^n_y$ whose marginals on $\mathbb{R}^n_x$ and $\mathbb{R}^n_y$ are equal to μ and ν respectively. Such probability measures are generalizations of maps which push forward μ on ν, so that W_p represents the best transport cost of μ to ν; the second equality is simply a probabilistic rephrasing. It is trivial to prove that $W_p(\mu,\nu) < \infty$ as long as both μ and ν have finite p-th moment, and with a bit more work one may prove W_p to be a distance. Convergence under this distance turns out to be equivalent to weak-* convergence coupled with convergence of the p-th moment. Simple variational arguments also show the existence of a minimizer $\bar{\gamma}$, or $(\bar{X},\bar{V})$. We need no other basic tool to enter the main discussion.

Lemma 17. *Let $p \in [1,+\infty)$. For all μ and ν probability measures on $\mathbb{R}^n$ with finite p-th moment and all $N \geq 1$,*

$$W_p^p(\mu^{\otimes N}, \nu^{\otimes N}) = N W_p^p(\mu,\nu) \ .$$

Proof. Clearly by considering a minimizer $\bar{\gamma}$ in $W_p(\mu,\nu)$, we have that $\bar{\gamma}^{\otimes N}$ has marginals $\mu^{\otimes N}$ and $\nu^{\otimes N}$ so that, by writing $x=(x_1,\dots,x_N)\in\mathbb{R}^{nN}$

$$\begin{aligned} W_p^p(\mu^{\otimes N},\nu^{\otimes N}) &\le \int_{\mathbb{R}_x^{nN}\times\mathbb{R}_y^{nN}} \sum_{i=i}^{N} |x_i-y_i|^p \,\mathrm{d}\bar{\gamma}^{\otimes N}(x,y) \\ &= N\int_{\mathbb{R}^n_{x_1}\times\mathbb{R}^n_{y_1}} |x_1-y_1|^p \,\mathrm{d}\bar{\gamma}(x_1,y_1) = NW_p^p(\mu,\nu)\ . \end{aligned}$$

Conversely, for all $\gamma^{(N)}\in\mathcal{P}(\mu^{\otimes N},\nu^{\otimes N})$, by calling $\gamma_i^{(N)}$ the measure on $\mathbb{R}^n\times\mathbb{R}^n$ obtained by projecting $\gamma^{(N)}$ along all but the i-th coordinate, it is clear that $\gamma_i^{(N)}\in\mathcal{P}(\mu,\nu)$ so that

$$\begin{aligned} \int_{\mathbb{R}_x^{nN}\times\mathbb{R}_y^{nN}} |x-y|^p \,\mathrm{d}\gamma^{(N)}(x,y) &= \sum_{i=1}^{N}\int_{\mathbb{R}_x^{nN}\times\mathbb{R}_y^{nN}} |x_i-y_i|^p \,\mathrm{d}\gamma^{(N)}(x,y) \\ &= \sum_{i=1}^{N}\int_{\mathbb{R}^N_{x_i}\times\mathbb{R}^N_{y_i}} |x_i-y_i|^p \,\mathrm{d}\gamma_i^{(N)}(x_i,y_i) \ge NW_p^p(\mu,\nu) \end{aligned}$$

and the claim follows by just taking the infimum in $\gamma^{(N)}$. □

Theorem 18. *Suppose that f_t solves (4.1) and that the initial data f_0 satisfies*

$$M_2(f_0) := \int_{\mathbb{R}_x^d\times\mathbb{R}_v^d} f_0(x,v)\big(|x|^2+|v|^2\big)\,dx\,dv < \infty\ .$$

Then there exist constants $C_1>0$, depending on $M_2(f_0)$ and parameters of the equation, and C_2, depending only on the parameters of the equation, such that

$$W_2^2(f_t,f_\infty)\le C_1 e^{-C_2 t}\ .$$

Proof. First, by Lemma 17 squared Wasserstein distance tensorizes as a sum, so we can establish a link between $W_2^2\big(f_t,f_\infty\big)$ and $W_2^2\big(f_t^{\otimes N},f_\infty^{\otimes N}\big)$ for arbitrary N to be eventually chosen. Let us then triangulate the latter with the other measures on $\mathbb{R}^{2dN}$ $f_t^{(N)}$ and $f_\infty^{(N)}$ and

$$\begin{aligned} W_2^2(f_t,f_\infty) &= \frac{1}{N}W_2^2\big(f_t^{\otimes N},f_\infty^{\otimes N}\big) \\ &\le \frac{1}{N}\Big(W_2\big(f_t^{\otimes N},f_t^{(N)}\big)+W_2\big(f_t^{(N)},f_\infty^{(N)}\big)+W_2\big(f_\infty^{(N)},f_\infty^{\otimes N}\big)\Big)^2\ . \end{aligned} \tag{4.5}$$

By focusing, at first, on the second term, the equilibrium

$$f_\infty^{(N)}(x,v) = e^{-\frac{2\gamma}{\sigma^2}\left(\mathcal{U}_N(x)+\frac{|v|^2}{2}\right)}$$

satisfies a Talagrand inequality: the potential is $\frac{2\gamma c}{\sigma^2}$-convex, where we relabel $c:=\min\{c_U+2c_W,1\}>0$. Then, by Bakry-Émery criterion, it satisfies a Logarithmic Sobolev Inequality with constant $\frac{2\gamma c}{\sigma^2}$, and therefore it also satisfies a Talagrand inequality with the same constant – for such and other interesting inequalities see Chapter 9 in [2]. Therefore

$$W_2^2\big(f_t^{(N)},f_\infty^{(N)}\big)\le\frac{\sigma^2}{\gamma c}H_{f_\infty^{(N)}}\big(f_t^{(N)}\big)\ .$$

Since we are now dealing with $f_t^{(N)}$ solution of a Fokker-Planck equation with potential $\mathcal{U}_N$, by Theorem 16 the decay of the right hand side is close to exponential, with decay constant C_N which, by the explicit formula (3.11), is an increasing function of the convexity $c_{\mathcal{U}_N}$ of the potential $\mathcal{U}_N$. Such increasingness gives that a growth of the dimension will just improve the convergence, giving a more than satisfactory uniformity as the dimension grows. It then follows from our assumptions on the potential that all $f_t^{(N)}$ decay with some rate $C > 0$, giving that

$$W_2^2\big(f_t^{(N)}, f_\infty^{(N)}\big) \leq \frac{\sigma^2}{\gamma c} H_{f_\infty^{(N)}}(f_0^{(N)}) e^{-Ct(1-e^{-t})^2} \ .$$

We now wish to give a somehow explicit bound to the relative entropy of $f_0^{(N)}$, that we may factorize into $f_0^{\otimes N}$ by the assumption of independence at time 0: by reminding that $f_\infty^{(N)}(x,v) = \frac{e^{-\frac{2\gamma}{\sigma^2} U_N(x,v)}}{C_N}$ where $U_N(x,v) = \mathcal{U}_N(x) + \frac{|v|^2}{2}$ and C_N is just the renormalizing constant,

$$\begin{aligned} H_{f_\infty^{(N)}}(f_0^{\otimes N}) &= \int f_0^{\otimes N} \log\Big(\frac{f_0^{\otimes N}}{f_\infty^{(N)}}\Big) \\ &= \int f_0^{\otimes N} \log f_0^{\otimes N} + \frac{2\gamma}{\sigma^2} \int f_0^{\otimes N}(x,v)[\mathcal{U}_N(x) + \frac{|v|^2}{2}] \,\mathrm{d}x\,\mathrm{d}v \\ &+ \log(C_N) \int f_0^{\otimes N} \ . \end{aligned}$$

For brevity we will always write $\mathbb{R}^{2dN}$ for $\mathbb{R}_x^{dN} \times \mathbb{R}_v^{dN}$. The tensorization of the H functional gives

$$\begin{aligned} \int_{\mathbb{R}^{2dN}} f_0^{\otimes N} \log f_0^{\otimes N} \,\mathrm{d}x\,\mathrm{d}v &= \sum_{i=1}^N \int_{\mathbb{R}^{2dN}} \prod_{j=1}^N f_0(x_j, v_j) \log f_0(x_i, v_i) \,\mathrm{d}x\,\mathrm{d}v \\ &= \sum_{i=1}^N \int_{\mathbb{R}^{2d}} f_0(x_i, v_i) \log f_0(x_i, v_i) \,\mathrm{d}x_i\,\mathrm{d}v_i \\ &= N \int_{\mathbb{R}^{2d}} f_0(x,v) \log f_0(x,v) \,\mathrm{d}x\,\mathrm{d}v \end{aligned}$$

since f_0 has mass 1 on $\mathbb{R}^{2d}$. Next,

$$\begin{aligned} &\int_{\mathbb{R}^{2d}} f_0^{\otimes N} \mathcal{U}_N \\ &\quad = \int_{\mathbb{R}^{2dN}} f_0(x_1, v_1) \dots f_0(x_N, v_N) \Big[\sum_{i=1}^N U(x_i) + \frac{1}{2N} \sum_{i,j=1}^N W(x_i - x_j)\Big] \,\mathrm{d}x\,\mathrm{d}v \ . \end{aligned}$$

By supposing that $U(x) \leq a_U + b_U \left|x\right|^2$ and $W(x) \leq a_W + b_W \left|x\right|^2$,

$$\begin{aligned} &\int_{\mathbb{R}^{2dN}} f_0(x_1, v_1) \dots f_0(x_N, v_N) \Big[\sum_{i=1}^N U(x_i) \,\mathrm{d}x\,\mathrm{d}v\Big] \\ &\qquad = N \int_{\mathbb{R}^{2d}} f_0(x,v) U(x) \,\mathrm{d}x\,\mathrm{d}v \leq N\Big[a_U + b_U \int_{\mathbb{R}^{2d}} f_0(x,v) \left|x\right|^2 \,\mathrm{d}x\,\mathrm{d}v\Big] \end{aligned}$$

and

$$\int_{\mathbb{R}^{2dN}} f_0(x_1,v_1)\dots f_0(x_N,v_N)\Big[\frac{1}{2N}\sum_{i,j=1}^{N} W(x_i-x_j)\Big]\,\mathrm{d}x\,\mathrm{d}v$$

$$\le \frac{1}{2N}\int_{\mathbb{R}^{2dN}} f_0(x_1,v_1)\dots f_0(x_N,v_N)\Big[\sum_{i,j=1}^{N} a_W + b_W\,|x_i-x_j|^2\Big]\,\mathrm{d}x\,\mathrm{d}v$$

$$\le \frac{N}{2}a_W + \frac{b_W}{N}\sum_{i,j=1}^{N}\int_{\mathbb{R}^{2dN}} f_0(x_1,v_1)\dots f_0(x_N,v_N)\big(\,|x_i|^2+|x_j|^2\,\big)\,\mathrm{d}x\,\mathrm{d}v$$

$$= \frac{N}{2}a_W + 2Nb_W\int_{\mathbb{R}^{2d}} f_0(x,v)\,|x|^2\,\mathrm{d}x\,\mathrm{d}v\ ,$$

while clearly

$$\int_{\mathbb{R}^{2dN}} f_0^{\otimes N}(x,v)\frac{|v|^2}{2}\,\mathrm{d}x\,\mathrm{d}v = \frac{N}{2}\int_{\mathbb{R}^{2d}} f_0(x,v)\,|v|^2\,\mathrm{d}x\,\mathrm{d}v\ ,$$

so that

$$\int_{\mathbb{R}^{2dN}} f_0^{\otimes N}U_N$$
$$\le N\Big[a_U+\frac{a_W}{2}+(b_U+2b_W)\int_{\mathbb{R}^{2d}} f_0(x,v)\,|x|^2\,\mathrm{d}x\,\mathrm{d}v+\frac{1}{2}\int_{\mathbb{R}^{2d}} f_0(x,v)\,|v|^2\,\mathrm{d}x\,\mathrm{d}v\Big]\ .$$

Last, let us focus on $\log(C_N)$: by our assumptions we may suppose that $\mathcal{U}_N(x)\ge a+b\,|x|^2$

$$C_N = \int_{\mathbb{R}^{2dN}} e^{-\frac{2\gamma}{\sigma^2}U_N(x,v)}\,\mathrm{d}x\,\mathrm{d}v = \int_{\mathbb{R}^{dN}} e^{-\frac{2\gamma}{\sigma^2}\mathcal{U}_N(x)}\,\mathrm{d}x\int_{\mathbb{R}^{dN}} e^{-\frac{2\gamma}{\sigma^2}\frac{|v|^2}{2}}\,\mathrm{d}v$$
$$\le \int_{\mathbb{R}^{dN}} e^{-\frac{2\gamma}{\sigma^2}(a+b|x|^2)}\,\mathrm{d}x\int_{\mathbb{R}^{dN}} e^{-\frac{2\gamma}{\sigma^2}\frac{|v|^2}{2}}\,\mathrm{d}v = e^{-\frac{2\gamma a}{\sigma^2}}\left[\frac{\pi\sigma^2}{2\sqrt{b}\gamma}\right]^{dN}$$

giving that

$$\log(C_N)\int_{\mathbb{R}^{2dN}} f_0^{\otimes N} \le -\frac{2\gamma a}{\sigma^2} + Nd\log\left[\frac{\pi\sigma^2}{2\sqrt{d}\gamma}\right]\ .$$

All in all,

$$H_{f_\infty^{(N)}}(f_0^{\otimes N}) \le N\Big[\frac{2\gamma}{\sigma^2}\Big(a_U+\frac{a_W}{2}+(b_U+2b_W)\int_{\mathbb{R}^{2d}} f_0(x,v)\,|x|^2\,\mathrm{d}x\,\mathrm{d}v$$
$$+\frac{1}{2}\int_{\mathbb{R}^{2d}} f_0(x,v)\,|v|^2\,\mathrm{d}x\,\mathrm{d}v\Big)+\int_{\mathbb{R}^{2d}} f_0\log f_0 + d\Big|\log\left(\frac{\pi\sigma^2}{2\sqrt{d}\gamma}\right)\Big|\Big]$$

or, more compactly,

$$H_{f_\infty^{(N)}}(f_0^{\otimes N}) \le K(f_0)N\ ,$$

where K depends on the second moment and the free entropy of f_0. Back to our system, we have yielded that

$$W_2^2(f_t^{(N)}, f_\infty^{(N)}) \le \frac{\sigma^2}{\gamma c}K(f_0)Ne^{-Ct(1-e^{-t})} \tag{4.6}$$

which is enough for the second term in (4.5).

Concerning the third term, we refer to [4]. By considering t approaching infinity, we can deduce that

$$W_2(f_\infty^{(N)}, f_\infty^{\otimes N}) \leq K \ , \tag{4.7}$$

where K depends on f_0 and parameters of the equation, but not on N.
We are now left only with the first term in (4.5), with which we will deal in the following way: consider $Z_N(t) = (X_t, V_t)$ and $\bar{Z}_N(t) = (\bar{X}_t, \bar{V}_t)$ solutions of equations (4.3), (4.2) and driven by the same Brownian motion B_t. We want to establish an estimate, uniform in N, on $\mathbb{E}\big[\left|Z_N(t) - \bar{Z}_N(t)\right|^2\big]$: for all $1 \leq i \leq N$, by recalling that $\mathcal{U}(x,y) = U(x) + U(y) + W(x-y)$,

$$\begin{aligned}
\partial_t(\left|X_t^i - \bar{X}_t^i\right|^2 + \left|V_t^i - \bar{V}_t^i\right|^2) &= 2(X - \bar{X}_t^i)\cdot(V - \bar{V}_t^i) + 2(V_t^i - \bar{V}_t^i) \\
\cdot \Bigg[-\gamma(V_t^i - \bar{V}_t^i) - \Bigg(\frac{1}{N}\sum_{j=1}^N \nabla_x \mathcal{U}(X_t^i, X_t^j) &- \int_{\mathbb{R}^{2d}} \nabla_x \mathcal{U}(\bar{X}_t^i, y) f_t(y,v)\,\mathrm{d}y\,\mathrm{d}v\Bigg)\Bigg] \\
&\leq 2\left|X_t^i - \bar{X}_t^i\right|\left|V_t^i - \bar{V}_t^i\right| - 2\gamma\left|V_t^i - \bar{V}_t^i\right|^2 \\
&\quad + 2\left|V_t^i - \bar{V}_t^i\right| \Big[\left|\nabla_x U(X_t^i) - \nabla_x U(\bar{X}_t^i)\right| \\
&\quad + \left|\frac{1}{N}\sum_{j=1}^N \nabla W(X_t^i - X_t^j) - \int_{\mathbb{R}^{2d}} \nabla W(\bar{X}_t^i - y) f_t(y,v)\,\mathrm{d}y\,\mathrm{d}v\right|\Big] \ .
\end{aligned}$$

Concerning the exterior potential, we have that

$$2\left|V_t^i - \bar{V}_t^i\right|[\nabla_x U(X_t^i) - \nabla_x U(\bar{X}_t^i)] \leq 2\left\|D^2 U\right\|_{L^\infty}\left|X_t^i - \bar{X}_t^i\right|\left|V_t^i - \bar{V}_t^i\right| \ ,$$

while, for the interactive part, triangulate with $\nabla W(\bar{X}_t^i - \bar{X}_t^j)$ and

$$\begin{aligned}
&\left|\frac{1}{N}\sum_{j=1}^N \nabla_x W(X_t^i - X_t^j) - \int_{\mathbb{R}^{2d}} \nabla_x W(\bar{X}_t^i - y) f_t(y,v)\,\mathrm{d}y\,\mathrm{d}v\right| \\
=& \frac{1}{N}\left|\sum_{j=1}^N \left[\nabla_x W(X_t^i - X_t^j) - \int_{\mathbb{R}^{2d}} \nabla_x W(\bar{X}_t^i - y) f_t(y,v)\,\mathrm{d}y\,\mathrm{d}v\right]\right| \\
\leq& \frac{1}{N}\left|\sum_{j=1}^N \left[\nabla_x W(X_t^i - X_t^j) - \nabla W(\bar{X}_t^i - \bar{X}_t^j)\right]\right| \\
&+ \frac{1}{N}\left|\sum_{j=1}^N \left[\nabla_x W(\bar{X}_t^i - \bar{X}_t^j) - \int_{\mathbb{R}^{2d}} \nabla_x W(\bar{X}_t^i - y) f_t(y,v)\,\mathrm{d}y\,\mathrm{d}v\right]\right| \\
\leq& \frac{1}{N}\left\|D^2 W\right\|_{L^\infty}\sum_{j=1}^N \left|X_t^i - \bar{X}_t^i + X_t^j - \bar{X}_t^j\right| \\
&+ \frac{1}{N}\left|\sum_{j=1}^N \left[\nabla_x W(\bar{X}_t^i - \bar{X}_t^j) - \int_{\mathbb{R}^{2d}} \nabla_x W(\bar{X}_t^i - y) f_t(y,v)\,\mathrm{d}y\,\mathrm{d}v\right]\right|
\end{aligned}$$

so that, by taking expectation,

$$\partial_t \mathbb{E}\big[\,\big|X_t^i - \bar{X}_t^i\big|^2 + \big|V_t^i - \bar{V}_t^i\big|^2\,\big] \le \big(1 + \big\|D^2 U\big\|_\infty\big)\mathbb{E}\big[\,\big|X_t^i - \bar{X}_t^i\big|^2 + \big|V_t^i - \bar{V}_t^i\big|^2\,\big] \tag{4.8}$$
$$+ \frac{1}{N}\big\|D^2 W\big\|_{L^\infty} \sum_{j=1}^N \mathbb{E}\,\big|V_t^i - \bar{V}_t^i\big|\,\Big|X_t^i - \bar{X}_t^i + X_t^j - \bar{X}_t^j\Big|$$
$$+ \frac{1}{N}\mathbb{E}\left[\big|V_t^i - \bar{V}_t^i\big|\left|\sum_{j=1}^N \left[\nabla_x W(\bar{X}_t^i - \bar{X}_t^j) - \int_{\mathbb{R}^{2d}} \nabla_x W(\bar{X}_t^i - y) f_t(y,v)\,\mathrm{d}y\,\mathrm{d}v\right]\right|\right].$$

The last term gives, by Cauchy-Schwarz,

$$\frac{1}{N}\mathbb{E}\left[\big|V_t^i - \bar{V}_t^i\big|\left|\sum_{j=1}^N \left[\nabla_x W(\bar{X}_t^i - \bar{X}_t^j) - \int_{\mathbb{R}^{2d}} \nabla_x W(\bar{X}_t^i - y) f_t(y,v)\,\mathrm{d}y\,\mathrm{d}v\right]\right|\right]$$
$$\le \frac{1}{2}\mathbb{E}\big[\,\big|V_t^i - \bar{V}_t^i\big|^2\,\big]$$
$$+ \frac{1}{2N^2}\mathbb{E}\left|\sum_{j=1}^N \left[\nabla_x W(\bar{X}_t^i - \bar{X}_t^j) - \int_{\mathbb{R}^{2d}} \nabla_x W(\bar{X}_t^i - y) f_t(y,v)\,\mathrm{d}y\,\mathrm{d}v\right]\right|^2.$$

The N random variables $\big\{\nabla_x W(\bar{X}_t^i - \bar{X}_t^j) - \int_{\mathbb{R}^{2d}} \nabla_x W(\bar{X}_t^i - y) f_t(y,v)\,\mathrm{d}y\,\mathrm{d}v\big\}_{j=1}^N$ are independent conditionally to $\{\bar{X}_t^i\}_i$, since $\bar{X}_t^j$ are, and of zero average:

$$\mathbb{E}\big[\nabla_x W(\bar{X}_t^i - \bar{X}_t^j)\big] = \int_{\mathbb{R}^{2d}} \nabla_x W(x-y) f_t(y,v) f_t(x,w)\,\mathrm{d}y\,\mathrm{d}v\,\mathrm{d}x\,\mathrm{d}w$$
$$= \mathbb{E}\left[\int_{\mathbb{R}^{2d}} \nabla_x W(\bar{X}_t^i - y) f_t(y,v)\,\mathrm{d}y\,\mathrm{d}v\right],$$

so that orthogonality gives

$$\mathbb{E}\left[\left|\sum_{j=1}^N \left[\nabla_x W(\bar{X}_t^i - \bar{X}_t^j) - \int_{\mathbb{R}^{2d}} \nabla_x W(\bar{X}_t^i - y) f_t(y,v)\,\mathrm{d}y\,\mathrm{d}v\right]\right|^2\right]$$
$$= \sum_{j=1}^N \mathbb{E}\left|\nabla_x W(\bar{X}_t^i - \bar{X}_t^j) - \int_{\mathbb{R}^{2d}} \nabla_x W(\bar{X}_t^i - y) f_t(y,v)\,\mathrm{d}y\,\mathrm{d}v\right|^2.$$

Since

$$\left|\nabla_x W(\bar{X}_t^i - \bar{X}_t^j) - \int_{\mathbb{R}^{2d}} \nabla_x W(\bar{X}_t^i - y) f_t(y,v)\,\mathrm{d}y\,\mathrm{d}v\right|$$
$$= \left|\int_{\mathbb{R}^{2d}} [\nabla_x W(\bar{X}_t^i - \bar{X}_t^j) - \nabla_x W(\bar{X}_t^i - y)] f_t(y,v)\,\mathrm{d}y\,\mathrm{d}v\right|$$
$$\le \big\|D^2 W\big\|_\infty \int_{\mathbb{R}^{2d}} \Big|\bar{X}_t^j - y\Big|\, f_t(y,v)\,\mathrm{d}y\,\mathrm{d}v$$

by the usual mean value argument,

$$
\begin{aligned}
&\mathbb{E}\left[\Big|\sum_{j=1}^{N}\Big[\nabla_x W(\bar{X}_t^i-\bar{X}_t^j)-\int_{\mathbb{R}^{2d}}\nabla_x W(\bar{X}_t^i-y)f_t(y,v)\,\mathrm{d}y\,\mathrm{d}v\Big]\Big|^2\right]\\
&\qquad\le \left\|D^2W\right\|_\infty^2\sum_{j=1}^{N}\mathbb{E}\Big(\int_{\mathbb{R}^{2d}}\left|\bar{X}_t^j-y\right| f_t(y,v)\,\mathrm{d}y\,\mathrm{d}v\Big)^2\\
&\qquad\le \left\|D^2W\right\|_\infty^2\sum_{j=1}^{N}\mathbb{E}\int_{\mathbb{R}^{2d}}\left|\bar{X}_t^j-y\right|^2 f_t(y,v)\,\mathrm{d}y\,\mathrm{d}v \qquad (4.9)\\
&\qquad= N\left\|D^2W\right\|_\infty^2\int_{\mathbb{R}^{2d}}|x-y|^2 f_t(y,v)f_t(x,w)\,\mathrm{d}y\,\mathrm{d}v\,\mathrm{d}x\,\mathrm{d}w\\
&\qquad\le N\left\|D^2W\right\|_\infty^2\int_{\mathbb{R}^{2d}}2\big(|x|^2+|y|^2\big)f_t(y,v)f_t(x,w)\,\mathrm{d}y\,\mathrm{d}v\,\mathrm{d}x\,\mathrm{d}w\\
&\qquad= 4N\left\|D^2W\right\|_\infty^2\int_{\mathbb{R}^{2d}}|y|^2 f_t(y,v)\,\mathrm{d}y\,\mathrm{d}v = 4N\left\|D^2W\right\|_\infty^2\mathbb{E}\big[\,|\bar{X}_t^i|^2\,\big]\ .
\end{aligned}
$$

Concerning the second line in (4.8),

$$
\begin{aligned}
&\frac{1}{N}\left\|D^2W\right\|_{L^\infty}\sum_{j=1}^{N}\mathbb{E}\left|V_t^i-\bar{V}_t^i\right|\left|X_t^i-\bar{X}_t^i+X_t^j-\bar{X}_t^j\right|\\
&\quad=\left\|D^2W\right\|_{L^\infty}\mathbb{E}\Big[\left|V_t^i-\bar{V}_t^i\right|\frac{1}{N}\sum_{j=1}^{N}\left|X_t^i-\bar{X}_t^i+X_t^j-\bar{X}_t^j\right|\Big]\\
&\quad\le\frac{1}{2}\left\|D^2W\right\|_{L^\infty}^2\Big[\mathbb{E}\left|V_t^i-\bar{V}_t^i\right|^2+\frac{1}{N^2}\sum_{j=1}^{N}\left|X_t^i-\bar{X}_t^i\right|^2+\left|X_t^j-\bar{X}_t^j\right|^2\Big]\ .
\end{aligned}
$$

This gives that

$$
\begin{aligned}
&\partial_t\mathbb{E}\big[\left|X_t^i-\bar{X}_t^i\right|^2+\left|V_t^i-\bar{V}_t^i\right|^2\big]\\
&\quad\le\big(2+\left\|D^2U\right\|_\infty\big)\mathbb{E}\big[\left|X_t^i-\bar{X}_t^i\right|^2+\left|V_t^i-\bar{V}_t^i\right|^2\big]\\
&\qquad+\frac{1}{2}\left\|D^2W\right\|_\infty^2\Big[\mathbb{E}\left|V_t^i-\bar{V}_t^i\right|^2+\frac{1}{N^2}\sum_{j=1}^{N}\Big(\left|X_t^i-\bar{X}_t^i\right|^2+\left|X_t^j-\bar{X}_t^j\right|^2\Big)\Big]\\
&\qquad+\frac{2}{N}\left\|D^2W\right\|_\infty^2\mathbb{E}\big[\,|\bar{X}_t^i|^2\,\big]
\end{aligned}
$$

and, by taking the sum on i,

$$
\begin{aligned}
\partial_t \mathbb{E}\big[\,\big|X_t - \bar{X}_t\big|^2 + \big|V_t - \bar{V}_t\big|^2\,\big]
&\le \big(2 + \big\|D^2U\big\|_\infty\big)\mathbb{E}\big[\,\big|X_t - \bar{X}_t\big|^2 + \big|V_t - \bar{V}_t\big|^2\,\big] \\
&\quad + \big\|D^2W\big\|_\infty^2 \Bigg[\frac{1}{2}\mathbb{E}\left|V_t - \bar{V}_t\right|^2 \\
&\qquad + \frac{1}{2N^2}\sum_{i,j=1}^N \Big(\big|X_t^i - \bar{X}_t^i\big|^2 + \Big|X_t^j - \bar{X}_t^j\Big|^2\Big) + \frac{2}{N}\sum_{i=1}^N \mathbb{E}\left|\bar{X}_t^i\right|^2\Bigg] \\
&= \big(2 + \big\|D^2U\big\|_\infty\big)\mathbb{E}\big[\,\big|X_t - \bar{X}_t\big|^2 + \big|V_t - \bar{V}_t\big|^2\,\big] \\
&\quad + \big\|D^2W\big\|_\infty^2 \left[\frac{1}{2}\mathbb{E}\left|V_t - \bar{V}_t\right|^2 + \frac{1}{N}\left|X_t - \bar{X}_t\right|^2 + \frac{2}{N}\mathbb{E}\left|\bar{X}_t\right|^2\right] \\
&\le \Big(2 + \big\|D^2U\big\|_\infty + \frac{1}{2}\big\|D^2W\big\|_\infty^2\Big) \\
&\quad \cdot \mathbb{E}\Big[\,\big|X_t - \bar{X}_t\big|^2 + \big|V_t - \bar{V}_t\big|^2 + \frac{2}{N}\big\|D^2W\big\|_\infty^2\,\mathbb{E}\left|\bar{X}_t\right|^2\,\Big]
\end{aligned}
$$

since the $\bar{X}_t^i$ are independent. We then only need to give a bound on the second moment of $\bar{X}$, which comes from Lemma 19, to reach an inequality of the form

$$
\partial_t \mathbb{E}\big[\,\big|Z_N - \bar{Z}_N\big|^2\,\big] \le a\mathbb{E}[\big|Z_N - \bar{Z}_N\big|^2] + b\ .
$$

Since for all real, positive function g satisfying

$$
g' \le ag + b\ , \qquad a, b \in \mathbb{R}\ ,
$$

it holds

$$
g(t) \le g(0)e^{at} + \frac{b}{a}e^{at} - \frac{b}{a}\ ,
$$

and since

$$
\mathbb{E}\big[\,\big|Z_N(0) - \bar{Z}_N(0)\big|^2\,\big] = 0\ ,
$$

it finally follows that

$$
\mathbb{E}\big[\,\big|Z_N(t) - \bar{Z}_N(t)\big|^2\,\big] \le \frac{b}{a}(e^{at} - 1)\ ,
$$

giving us the desired bound

$$
W_2^2(f_t^{\otimes N}, f_t^{(N)}) \le \mathbb{E}\big[\,\big|X_t - \bar{X}_t\big|^2 + \big|V_t - \bar{V}_t\big|^2\,\big] \le \frac{b}{a}(e^{at} - 1) \tag{4.10}
$$

which is evident from the definition of Wasserstein distance. We can now collect (4.6), (4.7) and (4.10) so that (4.5) gives

$$
\begin{aligned}
W_2^2(f_t, f_\infty) &\le \frac{1}{N}[W_2(f_t^{\otimes N}, f_t^{(N)}) + W_2(f_t^{(N)}, f_\infty^{(N)}) + W_2(f_\infty^{(N)}, f_\infty^{\otimes N})]^2 \\
&\le \frac{1}{N}[C_1e^{at} + C_2N^{\frac{1}{2}}e^{-\frac{C}{2}t(1-e^{-t})} + C_3]^2
\end{aligned}
$$

which, upon taking $N = N(t)$ large enough, gives the claim.

□

Lemma 19. *Let (X, V) and $(\bar{X}, \bar{V})$ solve (4.3) and (4.2) respectively. Then there exists a constant C, which depends on the initial data $(\bar{X}_0, \bar{V}_0)$ and on parameters of the equation, such that for all t and all N*

$$\mathbb{E}[\left|\bar{X}_t\right|^2] \leq CN$$

Proof. In order to give a bound to the moment of $\bar{X}$ let us study the modified moment

$$H_\varepsilon(\bar{X}, \bar{V}) = U_N(\bar{X}, \bar{V}) + \varepsilon \bar{X} \cdot \bar{V} = \mathcal{U}_N(\bar{X}) + \frac{1}{2}\left|\bar{V}\right|^2 + \varepsilon \bar{X} \cdot \bar{V} \ .$$

We want to establish a differential inequality on $\mathbb{E}[H_\varepsilon(\bar{X}, \bar{V})]$ in order to prove that it is bounded in time for sufficiently small ε, that will also imply boundedness of $\mathbb{E}[\left|\bar{X}\right|^2]$ by uniform convexity of $\mathcal{U}_N$. By calling $\bar{L}_N$ the generator of $\frac{f_t^{\otimes N}}{f_\infty^{\otimes N}}$ and $\bar{L}_N^*$ its adjoint in $L^2(f_\infty^{\otimes N})$

$$\begin{aligned}
\partial_t \mathbb{E}[H_\varepsilon(\bar{X}, \bar{V})] &= \int_{\mathbb{R}^{2dN}} H_\varepsilon(x, v) \partial_t \frac{f_t^{\otimes N}}{f_\infty^{\otimes N}} f_\infty^{\otimes N} \,\mathrm{d}x \,\mathrm{d}v \\
&= \int_{\mathbb{R}^{2dN}} H_\varepsilon(x, v) \bar{L}_N \left(\frac{f_t^{\otimes N}}{f_\infty^{\otimes N}} \right) f_\infty^{\otimes N} \,\mathrm{d}x \,\mathrm{d}v \\
&= \int_{\mathbb{R}^{2dN}} \bar{L}_N^*[H_\varepsilon(x, v)] f_t^{\otimes N} \,\mathrm{d}x \,\mathrm{d}v \ .
\end{aligned}$$

Since

$$\bar{L}_N = -v \cdot \nabla_x + \nabla_x \bar{\mathcal{U}}_N \cdot \nabla_v - \gamma v \cdot \nabla_v + \frac{\sigma^2}{2} \Delta_v \quad ,$$

where

$$\bar{\mathcal{U}}_N(x) = \sum_{i=1}^N \int_{\mathbb{R}^{2d}} \mathcal{U}(x_i, y) f_t(y, w) \,\mathrm{d}y \,\mathrm{d}w \ ,$$

it follows from easy computations that

$$\bar{L}_N^* = v \cdot \nabla_x - \nabla_x \bar{\mathcal{U}}_N \cdot \nabla_v - \gamma v \cdot \nabla_v + \frac{\sigma^2}{2} \Delta_v = L_N^* + (\nabla_x \mathcal{U}_N - \nabla_x \bar{\mathcal{U}}_N) \cdot \nabla_v$$

where L_N denotes the generator of $f_t^{(N)}$, and the last adjoint is to be meant in $L^2(f_\infty^{(N)})$. First,

$$\begin{aligned}
L_N^*[H_\varepsilon(x, v)] &= \varepsilon \left|v\right|^2 - \varepsilon x \cdot \nabla_x \mathcal{U}_N - \gamma\big(\left|v\right|^2 + \varepsilon x v\big) + \frac{\sigma^2}{2} dN \\
&\leq \big(-\gamma + \varepsilon + 2\gamma^2 \sqrt{\varepsilon} \big) \left|v\right|^2 + \big(-C_1 \varepsilon + \varepsilon \sqrt{\varepsilon} \big) \left|x\right|^2 + \frac{\sigma^2}{2} dN
\end{aligned}$$

where we used that $\mathcal{U}_N$ attains its minimum in 0 and that the value of the minimum is 0. This may easily be supposed by considering translations of f if needed and by noticing that the potential, as such, is defined up to a constant. C_1 here is the lower bound to the convexity of $\mathcal{U}_N$.

$$L_N^*[H_\varepsilon(x, v)] \leq -\frac{\gamma}{4} \left|v\right|^2 - \frac{C_1 \varepsilon}{2} \left|x\right|^2 + \frac{\sigma^2}{2} dN$$

Also, concerning the other term in $\bar{L}_N^*$,

$$
\begin{aligned}
(\nabla_x \mathcal{U}_N - \nabla_x \bar{\mathcal{U}}_N) \cdot \nabla_v H_\varepsilon &= (\nabla_x \mathcal{U}_N - \nabla_x \bar{\mathcal{U}}_N) \cdot (v + \varepsilon x) \\
&\leq C_\varepsilon \left| \nabla_x \mathcal{U}_N - \nabla_x \bar{\mathcal{U}}_N \right|^2 + \frac{\gamma}{8} |v|^2 + \frac{C_1 \varepsilon}{4} |x|^2
\end{aligned}
$$

thanks to Young inequality with a sufficiently large constant C_ε, giving that

$$
\begin{aligned}
\bar{L}_N^*[H_\varepsilon(x,v)] &= L_N^* + (\nabla_x \mathcal{U}_N - \nabla_x \bar{\mathcal{U}}_N) \cdot \nabla_v H_\varepsilon \\
&\leq -\frac{\gamma}{8} |v|^2 - \frac{C_1 \varepsilon}{4} |x|^2 + \frac{\sigma^2}{2} dN + C_\varepsilon \left| \nabla_x \mathcal{U}_N - \nabla_x \bar{\mathcal{U}}_N \right|^2 \\
&\leq -\frac{\gamma}{8} |v|^2 - \frac{C_1}{2C_2} \varepsilon \mathcal{U}_N(x) + \frac{\sigma^2}{2} dN + C_\varepsilon \left| \nabla_x \mathcal{U}_N - \nabla_x \bar{\mathcal{U}}_N \right|^2 \\
&\leq -c_\varepsilon H_\varepsilon(x,v) + \frac{\sigma^2}{2} dN + C_\varepsilon \left| \nabla_x \mathcal{U}_N - \nabla_x \bar{\mathcal{U}}_N \right|^2 \ .
\end{aligned}
$$

where we called C_2 the upper bound of $D^2 \mathcal{U}_N$ and c_ε a sufficiently small constant. It follows that, for sufficiently small ε,

(4.11)

$$
\partial_t \mathbb{E}[H_\varepsilon(\bar{X}_t, \bar{V}_t)] \leq -c_\varepsilon \mathbb{E}[H_\varepsilon(\bar{X}_t, \bar{V}_t)] + C_\varepsilon \mathbb{E} \left| \nabla_x \mathcal{U}_N(\bar{X}_t) - \nabla_x \bar{\mathcal{U}}_N(\bar{X}_t) \right|^2 + \frac{\sigma^2}{2} dN \ .
$$

We just need to compute the term in the middle: for all $1 \leq k \leq N$, by writing $\nabla_k = \nabla_{x_k} \in \mathbb{R}^d$,

$$
\begin{aligned}
&\nabla_k \mathcal{U}_N(x) - \nabla_k \bar{\mathcal{U}}_N(x) \\
=&\nabla_k \Big[\sum_{j=1}^N U(x_j) + \frac{1}{2N} \sum_{i,j=1}^N W(x_i - x_j) \Big] \\
&- \sum_{j=1}^N \nabla_k \left[U(x_j) + \int_{\mathbb{R}^{2d}} W(x_j - y) f_t(y,v) \, \mathrm{d}y \, \mathrm{d}v \right] \\
=&\frac{1}{2N} \sum_{i,j=1}^N \nabla_k W(x_i - x_j) - \sum_{j=1}^N \int_{\mathbb{R}^{2d}} \nabla_k W(x_j - y) f_t(y,v) \, \mathrm{d}y \, \mathrm{d}v \\
=&\frac{1}{N} \sum_{j=1}^N \left[\nabla W(x_j - x_k) - \int_{\mathbb{R}^{2d}} \nabla W(x_k - y) f_t(y,v) \, \mathrm{d}y \, \mathrm{d}v \right] ,
\end{aligned}
$$

since it may be supposed that $\nabla W(0) = 0$, and then we have that

$$
\begin{aligned}
&\mathbb{E} \left| \nabla_x \mathcal{U}_N(\bar{X}_t) - \nabla_x \bar{\mathcal{U}}_N(\bar{X}_t) \right|^2 \\
&= \frac{1}{N^2} \sum_{k=1}^N \mathbb{E} \bigg| \sum_{j=1}^N \left[\nabla_x W(\bar{X}_t^j - \bar{X}_t^k) - \int_{\mathbb{R}^{2d}} \nabla_x W(\bar{X}_t^k - y) f_t(y,v) \, \mathrm{d}y \, \mathrm{d}v \right] \bigg|^2 \\
&= \frac{1}{N} \mathbb{E} \bigg| \sum_{j=1}^N \left[\nabla_x W(\bar{X}_t^j - \bar{X}_t^1) - \int_{\mathbb{R}^{2d}} \nabla_x W(\bar{X}_t^1 - y) f_t(y,v) \, \mathrm{d}y \, \mathrm{d}v \right] \bigg|^2 \\
&\leq \frac{1}{N} \left\| D^2 W \right\|_\infty^2 \sum_{j=1}^N \mathbb{E} \int_{\mathbb{R}^{2d}} \left| \bar{X}_t^j - y \right|^2 f_t(y,v) \, \mathrm{d}y \, \mathrm{d}v \leq \frac{4}{N} \left\| D^2 W \right\|_\infty^2 \mathbb{E} \big[\left| \bar{X}_t \right|^2 \big]
\end{aligned}
$$

where we acted as in (4.9) since the $\bar{X}^k$ are identically distributed. We can then go back to (4.11) and, supposing ε to have been fixed small enough,

$$\begin{aligned}\partial_t \mathbb{E}[H(\bar{X}_t,\bar{V}_t)] &\leq -c\mathbb{E}[H(\bar{X}_t,\bar{V}_t)] + \frac{4C}{N}\left\|D^2W\right\|_\infty^2 \mathbb{E}\left[\left|\bar{X}\right|^2\right] + \frac{\sigma^2}{2}dN \\ &\leq -c\mathbb{E}[H(\bar{X}_t,\bar{V}_t)] + \frac{C}{N}\mathbb{E}[H(\bar{X}_t,\bar{V}_t)] + \frac{\sigma^2}{2}dN \\ &\leq -c\mathbb{E}[H(\bar{X}_t,\bar{V}_t)] + \frac{\sigma^2}{2}dN\end{aligned}$$

by taking N large enough and relabelling constants. This closes our inequality and gives that for all t

$$\begin{aligned}\mathbb{E}[H(\bar{X}_t,\bar{V}_t)] &\leq e^{-ct}\left(\mathbb{E}[\bar{X}(0),\bar{V}(0)] - \frac{2c}{\sigma^2 dN}\right) + \frac{2c}{\sigma^2 dN} \\ &\leq \max\left\{\mathbb{E}[H(\bar{X}(0)),H(\bar{V}(0))],\frac{2c}{\sigma^2 dN}\right\}\end{aligned}$$

□

Acknowledgements

The author would like to express deep thanks to François Bolley for help received and to Laboratoire de Probabilités, Statistique et Modélisation in Paris for kind hosting. We would also like to thank reviewers for detailed and careful reading and Centre for Mathematical Sciences in Cambridge for facilities provided.

References

[1] D. Bakry and Michel Émery. Diffusions hypercontractives. In *Séminaire de probabilités, XIX, 1983/84*, volume 1123 of *Lecture Notes in Math.*, pages 177–206. Springer, Berlin, 1985.

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

[3] M. H. Duong and J. Tugaut. Stationary solutions of the Vlasov-Fokker-Planck equation: existence, characterization and phase-transition. *Appl. Math. Lett.*, 52:38–45, 2016.

[4] F. Malrieu. Logarithmic Sobolev inequalities for some nonlinear PDE's. *Stochastic Process. Appl.*, 95(1):109–132, 2001.

[5] P. Monmarché. Generalized Γ calculus and application to interacting particles on a graph. *ArXiv e-prints*, October 2015.

[6] P. Monmarché. Ergodicity and propagation of chaos for mean field kinetic particles. *ArXiv e-prints*, March 2016.

[7] Filippo Santambrogio. *Optimal Transport for Applied Mathematicians*, volume 87 of *Progress in Nonlinear Differential Equations and Theier Applications*. Birkh auser, Basel, 2015. Calculus of Variations, PDEs, and Modeling.

[8] Julian Tugaut. Convergence in Wasserstein distance for self-stabilizing diffusion evolving in a double-well landscape. *C. R. Math. Acad. Sci. Paris*, 356(6):657–660, 2018.

[9] Cédric Villani. Hypocoercivity. *Mem. Amer. Math. Soc.*, 202(950):iv+141, 2009.

[10] Cédric Villani. *Optimal transport*, volume 338 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, 2009. Old and new.