

# $\varphi$ -entropies: convexity, coercivity and hypocoercivity for Fokker-Planck and kinetic Fokker-Planck equations

Jean DOLBEAULT

*CEREMADE (CNRS UMR n° 7534),  
PSL research university, Université Paris-Dauphine,  
Place de Lattre de Tassigny, 75775 Paris 16, France  
E-mail: dolbeaul@ceremade.dauphine.fr*

Xingyu LI

*CEREMADE (CNRS UMR n° 7534),  
PSL research university, Université Paris-Dauphine,  
Place de Lattre de Tassigny, 75775 Paris 16, France  
E-mail: li@ceremade.dauphine.fr*

*July 20, 2018*

This paper is devoted to  $\varphi$ -entropies applied to Fokker-Planck and kinetic Fokker-Planck equations in the whole space, with confinement. The so-called  $\varphi$ -entropies are Lyapunov functionals which typically interpolate between Gibbs entropies and  $L^2$  estimates. We review some of their properties in the case of diffusion equations of Fokker-Planck type, give new and simplified proofs, and then adapt these methods to a kinetic Fokker-Planck equation acting on a phase space with positions and velocities. At kinetic level, since the diffusion only acts on the velocity variable, the transport operator plays an essential role in the relaxation process. Here we adopt the  $H^1$  point of view and establish a sharp decay rate. Rather than giving general but quantitatively vague estimates, our goal here is to consider simple cases, benchmark available methods and obtain sharp estimates on a key example. Some  $\varphi$ -entropies give rise to improved entropy – entropy production inequalities and, as a consequence, to faster decay rates for entropy estimates of solutions to non-degenerate diffusion equations. We prove that faster entropy decay also holds at kinetic level away from equilibrium and that optimal decay rates are achieved only in asymptotic regimes.

*Keywords:* Hypocoercivity; linear kinetic equations; Fokker-Planck operator; transport operator; diffusion limit; confinement; spectral gap; Poincaré inequality.

AMS Subject Classification (2010): 82C40; 76P05, 35K65, 35H10, 35P15, 35Q83, 35Q84.

## 1. Introduction

By definition, the  $\varphi$ -entropy of a nonnegative function  $w \in L^1(\mathbb{R}^d, d\gamma)$  is the functional

$$\mathcal{E}[w] := \int_{\mathbb{R}^d} \varphi(w) d\gamma,$$

where  $\varphi$  is a nonnegative convex continuous function on  $\mathbb{R}^+$  such that  $\varphi(1) = 0$  and  $1/\varphi''$  is concave on  $(0, +\infty)$ , *i.e.*,

$$\varphi'' \geq 0, \quad \varphi \geq \varphi(1) = 0 \quad \text{and} \quad (1/\varphi'')'' \leq 0. \quad (1.1)$$

Notice that the last condition means  $2(\varphi''')^2 \leq \varphi'' \varphi^{(iv)}$  a.e. A classical example of such a function  $\varphi$  is given by

$$\varphi_p(w) := \frac{1}{p-1} (w^p - 1 - p(w-1)) \quad p \in (1, 2]$$

where, in the case  $p = 2$ ,  $\varphi_2(w) = (w-1)^2$  and the limit case as  $p \rightarrow 1_+$  is given by the standard Gibbs entropy

$$\varphi_1(w) := w \log w - (w-1).$$

Many results corresponding to the case  $p = 2$  can be obtained, *e.g.*, by spectral methods. The case  $p = 1$  is important in probability theory and statistical physics. Our goal is to emphasize that they share properties which can be put in a common framework. Throughout this paper we shall assume that  $d\gamma$  is a nonnegative bounded measure, which is absolutely continuous with respect to Lebesgue's measure and write

$$d\gamma = e^{-\psi} dx$$

where  $\psi$  is a *potential* such that  $e^{-\psi}$  is in  $L^1(\mathbb{R}^d, dx)$ . Up to the addition of a constant to  $\psi$ , we can assume without loss of generality that  $d\gamma$  is a probability measure. A review of the *main properties of  $\varphi$ -entropies*, new and simplified proofs and key references are given in Section 2.

Without entering the technical details, let us illustrate the use of the  $\varphi$ -entropy in the case of diffusion equations. A typical application of the  $\varphi$ -entropy is the control of the rate of relaxation of the solution to the *Ornstein-Uhlenbeck equation*

$$\frac{\partial w}{\partial t} = \mathbf{L} w := \Delta w - \nabla \psi \cdot \nabla w, \quad (1.2)$$

which is also known as the *backward Kolmogorov equation*. If we solve the equation with a nonnegative initial datum  $w_0$  such that  $\int_{\mathbb{R}^d} w_0 d\gamma = 1$ , then the solution satisfies  $\int_{\mathbb{R}^d} w(t, \cdot) d\gamma = 1$  for any  $t > 0$  and  $\lim_{t \rightarrow +\infty} w(t, \cdot) = 1$ . The Ornstein-Uhlenbeck operator  $\mathbf{L}$  defined on  $L^2(\mathbb{R}^d, d\gamma)$  is indeed self-adjoint and such that

$$- \int_{\mathbb{R}^d} (\mathbf{L} w_1) w_2 d\gamma = \int_{\mathbb{R}^d} \nabla w_1 \cdot \nabla w_2 d\gamma \quad \forall w_1, w_2 \in H^1(\mathbb{R}^d, d\gamma).$$

As a consequence, it is also straightforward to observe that for any solution  $w$  with initial datum  $w_0$  such that  $\mathcal{E}[w_0]$  is finite, then

$$\frac{d}{dt} \mathcal{E}[w] = - \int_{\mathbb{R}^d} \varphi''(w) |\nabla_x w|^2 d\gamma =: -\mathcal{I}[w],$$

where  $\mathcal{I}[w]$  denotes the  $\varphi$ -*Fisher information* functional. If for some  $\Lambda > 0$  we can establish the *entropy - entropy production* inequality

$$\mathcal{I}[w] \geq \Lambda \mathcal{E}[w] \quad \forall w \in H^1(\mathbb{R}^d, d\gamma), \quad (1.3)$$

then we deduce that

$$\mathcal{E}[w(t, \cdot)] \leq \mathcal{E}[w_0] e^{-\Lambda t} \quad \forall t \geq 0,$$

which controls the convergence of  $w$  to 1 as  $t \rightarrow +\infty$ , for instance in  $L^p(\mathbb{R}^d, d\gamma)$  by a generalized *Csiszár-Kullback inequality* if  $\varphi = \varphi_p$ ,  $1 \leq p \leq 2$ . The entropy – entropy production inequality is the Poincaré inequality associated with  $d\gamma$  if  $\varphi = \varphi_2$ , and the logarithmic Sobolev inequality if  $\varphi = \varphi_1$ .

We recall that the study of (1.2) is equivalent to the study of the *Fokker-Planck equation*

$$\frac{\partial u}{\partial t} = \Delta u + \nabla_x \cdot (u \nabla_x \psi). \quad (1.4)$$

A nonnegative solution with initial datum  $u_0 \in L^1(\mathbb{R}^d, dx)$  and  $\int_{\mathbb{R}^d} u_0 dx = M > 0$  has constant mass  $M = \int_{\mathbb{R}^d} u(t, \cdot) dx$  for any  $t > 0$ , and converges towards the unique stationary solution

$$u_\star = M \frac{e^{-\psi}}{\int_{\mathbb{R}^d} e^{-\psi} dx}.$$

Without loss of generality, we shall assume that  $M = 1$ . Then one observes that  $w = u/u_\star$  solves (1.2), which allows to control the rate of convergence of  $u$  to  $u_\star$ . A list of results concerning the solutions of (1.2) and (1.4) is also collected in Section 2.

The third section of this paper is devoted to the extension of  $\varphi$ -entropy methods to kinetic equations. Section 3 of this paper deals with the *kinetic Fokker-Planck equation*, or *Vlasov-Fokker-Planck equation*, that can be written as

$$\frac{\partial f}{\partial t} + v \cdot \nabla_x f - \nabla_x \psi \cdot \nabla_v f = \Delta_v f + \nabla_v \cdot (v f). \quad (1.5)$$

Our basic example corresponds to the case of the harmonic potential  $\psi(x) = |x|^2/2$ . Unless it is explicitly specified, we will only consider this case. Notice that this problem has an explicit Green function whose expression can be found in <sup>23</sup>.

Since (1.5) is linear, we can assume at no cost that  $\|f\|_{L^1(\mathbb{R}^d \times \mathbb{R}^d)} = 1$  and consider the stationary solution

$$f_\star(x, v) = (2\pi)^{-d} e^{-\psi(x)} e^{-\frac{1}{2}|v|^2} = (2\pi)^{-d} e^{-\frac{1}{2}(|x|^2 + |v|^2)} \quad \forall (x, v) \in \mathbb{R}^d \times \mathbb{R}^d.$$

The function

$$g := \frac{f}{f_\star}$$

solves

$$\frac{\partial g}{\partial t} + \mathbb{T}g = \mathbb{L}g \quad (1.6)$$

where the transport operator  $\mathbb{T}$  and the Ornstein-Uhlenbeck operator  $\mathbb{L}$  are defined respectively by

$$\mathbb{T}g := v \cdot \nabla_x g - x \cdot \nabla_v g \quad \text{and} \quad \mathbb{L}g := \Delta_v g - v \cdot \nabla_v g.$$

Let  $d\mu := f_\star dx dv$  be the invariant measure on the phase space  $\mathbb{R}^d \times \mathbb{R}^d$ , so that  $\mathsf{T}$  and  $\mathsf{L}$  are respectively anti-self-adjoint and self-adjoint. The function

$$h := g^{p/2}$$

solves

$$\frac{\partial h}{\partial t} + \mathsf{T}h = \mathsf{L}h + \frac{2-p}{p} \frac{|\nabla_v h|^2}{h}. \quad (1.7)$$

At the kinetic level, we consider the  $\varphi$ -entropy given by

$$\mathcal{E}[g] := \iint_{\mathbb{R}^d \times \mathbb{R}^d} \varphi(g) d\mu.$$

With this notation,  $\mathcal{E}[g] = \iint_{\mathbb{R}^d \times \mathbb{R}^d} \varphi(f/f_\star) d\mu$  so that, with  $f = g f_\star = h^{2/p} f_\star$  we have

$$\mathcal{E}[g] = \iint_{\mathbb{R}^d \times \mathbb{R}^d} h^2 \log \left( \frac{h^2}{\iint_{\mathbb{R}^d \times \mathbb{R}^d} h^2 d\mu} \right) d\mu \quad \text{if } \varphi = \varphi_1,$$

$$\mathcal{E}[g] = \mathcal{E}[h^{2/p}] = \frac{1}{p-1} \left[ \iint_{\mathbb{R}^d \times \mathbb{R}^d} h^2 d\mu - \left( \iint_{\mathbb{R}^d \times \mathbb{R}^d} h^{2/p} d\mu \right)^{p/2} \right] \\ \text{if } \varphi = \varphi_p, p \in (1, 2].$$

The optimal rate of decay of  $\mathcal{E}[g]$  has been established by A. Arnold and J. Erb in [6](#). In the special case of a harmonic potential, their result goes as follows.

**Proposition 1.1.** *Assume that  $\psi(x) = |x|^2/2$  for any  $x \in \mathbb{R}^d$ . Take  $\varphi = \varphi_p$  for some  $p \in [1, 2]$ . To any nonnegative solution  $g \in L^1(\mathbb{R}^d \times \mathbb{R}^d)$  of [\(1.6\)](#) with initial datum  $g$  such that  $\mathcal{E}[g_0] < \infty$ , we can associate a constant  $\mathcal{C} > 0$  for which*

$$\mathcal{E}[g(t, \cdot, \cdot)] \leq \mathcal{C} e^{-t} \quad \forall t \geq 0. \quad (1.8)$$

Moreover the rate  $e^{-t}$  is sharp as  $t \rightarrow +\infty$ .

The striking point of this *hypocoercivity* result is to identify the sharp rate of decay. The rate is specific of the harmonic potential  $\psi(x) = |x|^2/2$ , but it turns out to be useful for the comparison with rates obtained by other methods. Although probably not optimal, a precise estimate of  $\mathcal{C}$  will be given in [Section 3](#), with a simplified proof of [Proposition 1.1](#).

The method is based on the use of a *Fisher information* type functional

$$\mathcal{J}[h] = \frac{1}{2} \int_{\mathbb{R}^d} |\nabla_v h|^2 d\mu + \frac{1}{2} \int_{\mathbb{R}^d} |\nabla_x h|^2 d\mu + \frac{1}{2} \int_{\mathbb{R}^d} |\nabla_x h + \nabla_v h|^2 d\mu \quad (1.9)$$

which involves derivatives in  $x$  and  $v$ . If  $h$  solves [\(1.6\)](#), then the key estimate is to prove that

$$\frac{d}{dt} \mathcal{J}[h(t, \cdot)] \leq -\mathcal{J}[h(t, \cdot)].$$

The result of Proposition 1.1 follows from the entropy – entropy production inequality (2.6) that will be established in Proposition 2.3: since

$$\Lambda \mathcal{E}[g(t, \cdot, \cdot)] = \Lambda \mathcal{E}[h^{2/p}] \leq \mathcal{J}[h],$$

then  $\mathcal{E}[g(t, \cdot, \cdot)]$  has an exponential decay. However, we underline the fact that

$$\frac{d}{dt} \mathcal{E}[g(t, \cdot)] = - \int_{\mathbb{R}^d} |\nabla_v h|^2 d\mu \neq - \mathcal{J}[h(t, \cdot)].$$

At the level of non-degenerate diffusions, a distinctive property of the  $\varphi$ -entropy with  $\varphi = \varphi_p$  and  $p \in (1, 2)$  is that the entropy – entropy production inequality  $\mathcal{I} \geq \Lambda \mathcal{E}$  with an optimal constant  $\Lambda > 0$  can be improved in the sense that there exists a strictly convex function  $F$  on  $\mathbb{R}^+$  with  $F(0) = 0$  and  $F'(0) = 1$  such that  $\mathcal{I} \geq \Lambda F(\mathcal{E})$ . This has been established in <sup>5</sup> and details will be given in Section 2.5. The key issue is to prove that for some function  $\rho$  on  $\mathbb{R}^+$ , which depends on the solution  $w$ , such that  $\rho > \Lambda$  a.e., we have  $\frac{d}{dt} \mathcal{I}[w(t, \cdot)] \leq -\rho(t) \mathcal{I}[w(t, \cdot)]$ . One may wonder if a similar result also holds in the hypocoercive kinetic Fokker-Planck equation. So far, no global improved inequality has been established. What we shall prove is that, if we consider the more general Fisher information functional

$$\mathcal{J}_\lambda[h] = (1 - \lambda) \int_{\mathbb{R}^d} |\nabla_v h|^2 d\mu + (1 - \lambda) \int_{\mathbb{R}^d} |\nabla_x h|^2 d\mu + \lambda \int_{\mathbb{R}^d} |\nabla_x h + \nabla_v h|^2 d\mu, \quad (1.10)$$

then for an appropriate choice of  $\lambda$  (which turns out to be  $t$ -dependent), the rate of decay is faster than  $e^{-t}$  up to a zero-measure set in  $t$ . The precise statement, which is our main result, goes as follows.

**Theorem 1.1.** *Let  $p \in (1, 2)$  and  $h$  be a solution of (1.7) with initial datum  $h_0 \in L^1 \cap L^p(\mathbb{R}^d, d\gamma)$ ,  $h_0 \not\equiv 1$ , and  $d\gamma$  be the Gaussian probability measure corresponding to the harmonic potential  $\psi(x) = |x|^2/2$ . Then there exists a function  $\lambda : \mathbb{R}^+ \rightarrow [1/2, 1)$  such that  $\lambda(0) = \lim_{t \rightarrow +\infty} \lambda(t) = 1/2$  and a continuous function  $\rho$  on  $\mathbb{R}^+$  such that  $\rho > 1/2$  a.e., for which we have*

$$\frac{d}{dt} \mathcal{J}_{\lambda(t)}[h(t, \cdot)] \leq -2\rho(t) \mathcal{J}_{\lambda(t)}[h(t, \cdot)] \quad \forall t \geq 0.$$

As a consequence, for any  $t \geq 0$  we have the global estimate

$$\mathcal{J}_{\lambda(t)}[h(t, \cdot)] \leq \mathcal{J}_{1/2}[h_0] \exp\left(-2 \int_0^t \rho(s) ds\right).$$

This result is weaker than the result for non-degenerate diffusions. The qualitative issues are easy to understand and to some extent classical in the hypocoercivity theory, but no quantitative estimate of  $\rho$  in terms of  $h$  is known so far. If  $\varphi_p$ -entropies were initially thought as interesting objects which interpolate between the Gibbs entropy and standard  $L^2$  estimates, improved entropy – entropy production inequalities and the result of Theorem 1.1 capture an important feature when  $p \in (1, 2)$ : faster rates of decay for finite values of  $t$ . As  $t \rightarrow +\infty$ , we cannot expect a faster

decay rate, but we gain a pre-factor which is less than 1. See Section 3.4 for more details.

Let us conclude this introduction with a brief review of the literature. Fokker-Planck equations like (1.4) are ubiquitous in various areas of physics ranging from the description of the motion of particles in a gas or a solute to semi-conductor physics, models of stars in astrophysics or models of populations in biology and social sciences, as microscopic dynamics involving Brownian motion are represented at macroscopic scales by diffusion equations. Second order dynamics (in which forces produce acceleration) in random environments obey in many cases to the Langevin equation and at macroscopic scale the corresponding distribution function solves (1.5). A typical example is given by particles having random encounters with some background obstacles, a situation that can be encountered in many areas of physical modeling. It has to be emphasized that (1.4) appears in the diffusion limit of the solutions of (1.5), that is, in the *overdamped regime* in which friction and other forces equilibrate very fast, so that the velocity instantaneously adapts to the forces, which results in first order dynamics. For some general properties of (1.4) and (1.5), a review of stochastic and PDE methods and some entries to applied cases, we refer for instance to <sup>64,62</sup>, among many other books on this topic.

The word “hypocoercivity” is apparently due to T. Gallay and was made popular by C. Villani in <sup>69</sup>. Our computations are based on Villani’s ideas in Section 3 of <sup>69</sup> (also see <sup>71</sup>), but the use of *twisted gradients* involving simultaneously derivatives in  $x$  and  $v$  can be also found in <sup>45</sup> and in earlier works like <sup>46</sup>. It is actually a consequence of Hörmander’s hypoelliptic theory, which covers simultaneously regularization properties and large time behaviour. One can refer for instance to <sup>36,46,43</sup> and, much earlier, to <sup>48</sup>. The seed for such an approach can actually be traced back to Kolmogorov’s computation of Green’s kernel for the kinetic Fokker-Planck equation in <sup>52</sup>, which has been reconsidered by <sup>50</sup> from a more modern point of view and successfully applied, for instance, to the study of the Vlasov-Poisson-Fokker-Planck system in <sup>68,19,20</sup>. In case of the kinetic Fokker-Planck equation, we can refer to <sup>43,45</sup> in the case of a general potential of confinement, and more specifically to <sup>6</sup>. In this last paper, the authors deal with the issue of accurate rates: “while the main theorem in <sup>71</sup> covers a wide class of problems, the price paid is in the estimate for the decay rate, which is off by orders of magnitude.” The result of Proposition 1.1 addresses the issue of the optimal rate in a very simple case. For completion, one also has to mention <sup>55</sup> and <sup>49</sup> for further theoretical and numerical results.

A twin problem of the kinetic Fokker-Planck equation is the linear BGK model, which has no regularizing properties but shares many common features with the kinetic Fokker-Planck equation as soon as we are concerned with rates of convergence. We refer to <sup>44,59</sup> for early contributions, to <sup>29,30,21,57,1</sup> for more recent ones, and especially to <sup>37</sup>. In this last paper, J. Evans studies the linear BGK model and a kinetic Fokker-Planck equation on the torus using the  $\varphi$ -entropies.

In <sup>69</sup>, only the cases  $p = 1$  and  $p = 2$  were considered, but it is well known since

the founding work <sup>9</sup> of Bakry and Emery that intermediate values of  $p$  can then be considered. In the case of  $\varphi$ -entropies associated with non-degenerate diffusions, this idea was invoked on many occasions, for instance in <sup>15,54,24,7,18,17</sup> in relation with spectral estimates or the *carré du champ* methods. For *carré du champ* techniques in kinetic equations, we can refer to <sup>14</sup>, also <sup>58,56</sup>, and finally Remark 6.7 in <sup>8</sup> for an early contribution on  $\varphi$ -entropies. Although  $\varphi$ -entropies are natural in the context of the kinetic Fokker-Planck equation, precise connections were made only quite recently. In <sup>6</sup>, A. Arnold and J. Erb discuss  $\varphi$ -entropies in the context of the kinetic Fokker-Planck equation and prove, among more general results, Proposition 1.1. We can also refer to <sup>1,2,58</sup> for various related results. As far as we know, no result such as Theorem 1.1 has been established yet.

## 2. A review of results on $\varphi$ -entropies

In this section we consider a  $\varphi$ -entropy defined by  $\mathcal{E}[w] := \int_{\mathbb{R}^d} \varphi(w) d\gamma$  where  $d\gamma = e^{-\psi} dx$  is a probability measure and  $\varphi$  satisfies (1.1). Most of the results presented here are known, but they are scattered in the literature. Our purpose here is to collect some essential statements and present simple proofs.

### 2.1. Generalized Csiszár-Kullback-Pinsker inequality

By assumption (1.1), we know that  $\mathcal{E}$  is nonnegative and achieves its minimum at  $w \equiv 1$ . It results from the strict convexity of  $\varphi$  that  $\mathcal{E}[w]$  controls a norm of  $(w - 1)$  under a generic assumption compatible with the expression of  $\varphi_p$ . The classical result of <sup>63,25,53</sup> has been extended in <sup>51,67,22,26</sup>. Here is a statement, with a short proof taken from Section 1.4 of <sup>13</sup>, for completeness.

**Proposition 2.1.** *Let  $p \in [1, 2]$ ,  $w \in L^1 \cap L^p(\mathbb{R}^d, d\gamma)$  be a nonnegative function, and assume that  $\varphi \in C^2(0, +\infty)$  is a nonnegative strictly convex function such that  $\varphi(1) = \varphi'(1) = 0$ . If  $A := \inf_{s \in (0, \infty)} s^{2-p} \varphi''(s) > 0$ , then*

$$\mathcal{E}[w] \geq 2^{-\frac{2}{p}} A \min \left\{ 1, \|w\|_{L^p(\mathbb{R}^d, d\gamma)}^{p-2} \right\} \|w - 1\|_{L^p(\mathbb{R}^d, d\gamma)}^2 .$$

When  $\varphi = \varphi_p$ , we find that  $A = p$ . This inequality has many variants and extensions: it is not limited to  $\mathbb{R}^d$  but also holds on bounded domains or manifolds and the relative  $\varphi$ -entropy  $\int_{\mathbb{R}^d} (\varphi(w_1) - \varphi(w_2) - \varphi'(w_1)(w_2 - w_1)) d\gamma$  can also be used to measure  $\|w_2 - w_1\|_{L^p(\mathbb{R}^d, d\gamma)}^2$ .

**Proof.** Up to the addition of a small constant, we can assume that  $w > 0$  and argue by density. A Taylor expansion at order two shows that

$$\mathcal{E}[w] = \frac{1}{2} \int_{\mathbb{R}^d} \varphi''(\xi) |w - 1|^2 d\gamma \geq \frac{A}{2} \int_{\mathbb{R}^d} \xi^{p-2} |w - 1|^2 d\gamma$$

where  $\xi$  lies between 1 and  $w$ . With  $\alpha = p(2 - p)/2$  and  $h > 0$ , for any measurable

set  $\mathcal{A} \subset \mathbb{R}^d$ , we get

$$\int_{\mathcal{A}} |w-1|^p h^{-\alpha} h^\alpha d\gamma \leq \left( \int_{\mathcal{A}} |w-1|^2 h^{p-2} d\gamma \right)^{p/2} \left( \int_{\mathcal{A}} h^p d\gamma \right)^{(2-p)/2}$$

by Hölder's inequality. We apply this formula to two different sets.

On  $\mathcal{A} = \{x \in \mathbb{R}^d : w(x) > 1\}$ , we use  $\xi^{p-2} \geq w^{p-2}$  and take  $h = w$ :

$$\int_{\{w>1\}} |w-1|^2 \xi^{p-2} d\gamma \geq \left( \int_{\{w>1\}} |w-1|^p d\gamma \right)^{2/p} \|w\|_{L^p(\mathbb{R}^d, d\gamma)}^{p-2}.$$

On  $\mathcal{A} = \{x \in \mathbb{R}^d : w(x) \leq 1\}$ , we use  $\xi^{p-2} \geq 1$  and take  $h = 1$ :

$$\int_{\{w \leq 1\}} |w-1|^2 \xi^{p-2} d\gamma \geq \left( \int_{\{w \leq 1\}} |w-1|^p d\gamma \right)^{2/p}.$$

By adding these two estimates and using with  $r = 2/p \geq 1$  the elementary inequality  $(a+b)^r \leq 2^{r-1}(a^r + b^r)$  for any  $a, b \geq 0$  allows us to conclude the proof.  $\square$

## 2.2. Convexity, tensorization and sub-additivity

Let us turn our attention to (1.3). To start with, we observe that the functional  $w \mapsto \mathcal{I}[w] = \int_{\mathbb{R}^d} \varphi''(w) |\nabla w|^2 d\gamma$  is convex if and only if  $1/\varphi''$  is concave. Now let us consider two probability measures  $d\gamma_1$  and  $d\gamma_2$  defined respectively on  $\mathbb{R}^{d_1}$  and  $\mathbb{R}^{d_2}$ , such that Inequality (1.3) holds with  $\gamma = \gamma_i$ , and  $i = 1, 2$ :

$$\int_{\mathbb{R}^{d_i}} \varphi''(w) |\nabla w|^2 d\gamma_i =: \mathcal{I}_{\gamma_i}[w] \geq \Lambda_i \mathcal{E}_{\gamma_i}[w] \quad \forall w \in H^1(\mathbb{R}^{d_i}, d\gamma_i), \quad (2.1)$$

Here we denote by  $\mathcal{E}_\gamma$  the  $\varphi$ -entropy for functions which are not normalized, that is,

$$\mathcal{E}_\gamma[w] := \int_{\mathbb{R}^d} \varphi(w) d\gamma - \varphi \left( \int_{\mathbb{R}^d} w d\gamma \right).$$

Assuming that  $d\gamma$  is a probability measure, by Jensen's inequality we know that  $w \mapsto \mathcal{E}_\gamma[w]$  is nonnegative because  $\varphi$  is convex. As we shall see below,  $w \mapsto \mathcal{E}_\gamma[w]$  is also convex, which is the key ingredient for *tensorization*. The question at stake is to know if Inequality (1.3) holds on  $\mathbb{R}^{d_1} \times \mathbb{R}^{d_2}$  for the measure  $d\gamma = d\gamma_1 \otimes d\gamma_2$ . Most of the results of Section 2.2 have been stated in [24](#) or are considered as classical. Our contribution here is to give simplified proofs.

**Theorem 2.1.** *Assume that  $\varphi$  satisfies (1.1). If  $d\gamma_1$  and  $d\gamma_2$  are two probability measures on  $\mathbb{R}^{d_1} \times \mathbb{R}^{d_2}$  satisfying (2.1) with positive constants  $\Lambda_1$  and  $\Lambda_2$ , then  $d\gamma_1 \otimes d\gamma_2$  is such that the following inequality holds:*

$$\begin{aligned} \mathcal{I}_{\gamma_1 \otimes \gamma_2}[w] &= \int_{\mathbb{R}^{d_1} \times \mathbb{R}^{d_2}} \varphi''(w) |\nabla w|^2 d\gamma_1 d\gamma_2 \\ &\geq \min\{\Lambda_1, \Lambda_2\} \mathcal{E}_{\gamma_1 \otimes \gamma_2}[w] \quad \forall w \in H^1(\mathbb{R}^{d_1} \times \mathbb{R}^{d_2}, d\gamma). \end{aligned}$$



It is straightforward to notice that the Fisher information is additive

$$\mathcal{I}_{\gamma_1 \otimes \gamma_2}[w] = \int_{\mathbb{R}^{d_2}} \mathcal{I}_{\gamma_1}[w] d\gamma_2 + \int_{\mathbb{R}^{d_1}} \mathcal{I}_{\gamma_2}[w] d\gamma_1,$$

so that the proof of Theorem 2.1 can be reduced to the proof of a *sub-additivity* property of the  $\varphi$ -entropies that goes as follows.

**Proposition 2.2.** *Assume that  $\varphi$  satisfies (1.1) and consider two probability measures  $d\gamma_1$  and  $d\gamma_2$  on  $\mathbb{R}^{d_1} \times \mathbb{R}^{d_2}$ . Then for any  $w \in L^1(\mathbb{R}^{d_1} \times \mathbb{R}^{d_2}, d\gamma_1 \otimes \gamma_2)$ , we have*

$$\mathcal{E}_{\gamma_1 \otimes \gamma_2}[w] \leq \int_{\mathbb{R}^{d_2}} \mathcal{E}_{\gamma_1}[w] d\gamma_2 + \int_{\mathbb{R}^{d_1}} \mathcal{E}_{\gamma_2}[w] d\gamma_1 \quad \forall w \in L^1(d\gamma_1 \otimes \gamma_2).$$

This last result relies on convexity properties that we are now going to study. As a preliminary step, we establish an inequality of Jensen type.

**Lemma 2.1.** *Let  $w \in L^1(\mathbb{R}^{d_1} \times \mathbb{R}^{d_2}, d\gamma_1 \otimes \gamma_2)$  be a function of two variables  $(x_1, x_2) \in \mathbb{R}^{d_1} \times \mathbb{R}^{d_2}$ . If  $\mathcal{F}_{\gamma_1}$  is a convex functional on  $L^1(d\gamma_1)$  such that*

$$\frac{d}{dt} \int_{\mathbb{R}^{d_2}} \mathcal{F}_{\gamma_1} \left[ t w + (1-t) \int_{\mathbb{R}^{d_2}} w d\gamma_2 \right] d\gamma_2 \Big|_{t=0} = 0, \quad (2.2)$$

then the following inequality holds:

$$\int_{\mathbb{R}^{d_2}} \mathcal{F}_{\gamma_1}[w] d\gamma_2 \geq \mathcal{F}_{\gamma_1} \left[ \int_{\mathbb{R}^{d_2}} w d\gamma_2 \right].$$

**Proof.** Let  $w_t = t w + (1-t) \int_{\mathbb{R}^{d_2}} w d\gamma_2$ . By convexity of  $\mathcal{F}_{\gamma_1}$ ,

$$\mathcal{F}_{\gamma_1}[w_t] \leq t \mathcal{F}_{\gamma_1}[w] + (1-t) \mathcal{F}_{\gamma_1} \left[ \int_{\mathbb{R}^{d_2}} w d\gamma_2 \right].$$

Hence it follows that

$$\mathcal{F}_{\gamma_1}[w_t] - \mathcal{F}_{\gamma_1} \left[ \int_{\mathbb{R}^{d_2}} w d\gamma_2 \right] \leq t \left( \mathcal{F}_{\gamma_1}[w] - \mathcal{F}_{\gamma_1} \left[ \int_{\mathbb{R}^{d_2}} w d\gamma_2 \right] \right),$$

from which we deduce that

$$0 = \frac{d}{dt} \mathcal{F}_{\gamma_1}[w_t] \Big|_{t=0} \leq \mathcal{F}_{\gamma_1}[w] - \mathcal{F}_{\gamma_1} \left[ \int_{\mathbb{R}^{d_2}} w d\gamma_2 \right].$$

Conclusion holds after integrating with respect to  $\gamma_2$ . □

The second observation is the proof of the convexity of  $w \mapsto \mathcal{E}_\gamma[w]$ . The following result is taken from 54.

**Lemma 2.2.** *If  $\varphi$  satisfies (1.1), then  $\mathcal{E}_\gamma$  is convex.*

**Proof.** We give a two steps proof of this result, for completeness.

- Define  $x_t = t y + (1-t)x$ ,  $t \in (0, 1)$ . Since  $1/\varphi''$  is concave,

$$\frac{1}{\varphi''(x_t)} \geq \frac{t}{\varphi''(y)} + \frac{1-t}{\varphi''(x)}. \quad (2.3)$$

The function  $\varphi$  is convex, hence  $\varphi''(x) > 0$  and  $\varphi''(y) > 0$  and so

$$\frac{1}{\varphi''(x_t)} \geq \frac{t}{\varphi''(y)} \quad \text{and} \quad \frac{1}{\varphi''(x_t)} \geq \frac{1-t}{\varphi''(x)}.$$

This means

$$\varphi''(y) \geq t \varphi''(x_t) \quad \text{and} \quad \varphi''(x) \geq (1-t) \varphi''(x_t).$$

We can also rewrite (2.3) as

$$\varphi''(x) \varphi''(y) \geq (t \varphi''(x) + (1-t) \varphi''(y)) \varphi''(x_t).$$

Consider the function

$$F_t(x, y) := t \varphi(y) + (1-t) \varphi(x) - \varphi(x_t)$$

and observe that

$$\text{Hess}(F_t) = \begin{pmatrix} (1-t) \varphi''(x) - (1-t)^2 \varphi''(x_t) & -t(1-t) \varphi''(x_t) \\ -t(1-t) \varphi''(x_t) & t \varphi''(y) - t^2 \varphi''(x_t) \end{pmatrix}$$

is nonnegative since both diagonal terms are nonnegative and the determinant is nonnegative. The matrix  $\text{Hess}(F_t)$  is therefore nonnegative and  $F_t$  is convex.

• We observe that

$$\begin{aligned} t \mathcal{E}_\gamma[w_1] + (1-t) \mathcal{E}_\gamma[w_0] - \mathcal{E}_\gamma[t w_1 + (1-t) w_0] \\ = \int_{\mathbb{R}^d} F_t(w_1, w_0) d\gamma - F_t \left( \int_{\mathbb{R}^d} w_1 d\gamma, \int_{\mathbb{R}^d} w_0 d\gamma \right) \end{aligned}$$

is nonnegative by Jensen's inequality, which proves the result.  $\square$

**Proof of Proposition 2.2.** We claim that  $\mathcal{F}_{\gamma_1} = \mathcal{E}_{\gamma_1}$  satisfies (2.2). Indeed, let us consider  $w_t = t w + (1-t) w_0$  with  $w_0 := \int_{\mathbb{R}^{d_2}} w d\gamma_2$ . A simple computation shows that

$$\frac{d}{dt} \mathcal{F}_{\gamma_1}[w_t] = \int_{\mathbb{R}^{d_1}} \varphi'(w_t) (w - w_0) d\gamma_1 - \varphi' \left( \int_{\mathbb{R}^{d_1}} w_t d\gamma_1 \right) \int_{\mathbb{R}^{d_1}} (w - w_0) d\gamma_1,$$

and, as a consequence at  $t = 0$ ,

$$\frac{d}{dt} \mathcal{F}_{\gamma_1}[w_t]_{t=0} = \int_{\mathbb{R}^{d_1}} \varphi'(w_0) (w - w_0) d\gamma_1 - \varphi' \left( \int_{\mathbb{R}^{d_1}} w_0 d\gamma_1 \right) \int_{\mathbb{R}^{d_1}} (w - w_0) d\gamma_1.$$

Since  $w_0$  does not depend on  $x_2$ , an integration with respect to  $\gamma_2$  concludes the proof of (2.2). From Lemma 2.1, we get

$$\int_{\mathbb{R}^{d_2}} \mathcal{E}_{\gamma_1}[w] d\gamma_2 \geq \mathcal{E}_{\gamma_1} \left[ \int_{\mathbb{R}^{d_2}} w d\gamma_2 \right].$$

By definition of  $\mathcal{E}_{\gamma_1}$ , this means

$$\begin{aligned} \int_{\mathbb{R}^{d_2}} \left[ \int_{\mathbb{R}^{d_1}} \varphi(w) d\gamma_1 - \varphi \left( \int_{\mathbb{R}^{d_1}} w d\gamma_1 \right) \right] d\gamma_2 \\ \geq \int_{\mathbb{R}^{d_1}} \varphi \left( \int_{\mathbb{R}^{d_2}} w d\gamma_2 \right) d\gamma_1 - \varphi \left( \iint_{\mathbb{R}^{d_1} \times \mathbb{R}^{d_2}} w d\gamma_1 \otimes \gamma_2 \right), \end{aligned}$$

from which we deduce

$$\begin{aligned} & \int_{\mathbb{R}^{d_2}} \left[ \int_{\mathbb{R}^{d_1}} \varphi(w) d\gamma_1 - \varphi \left( \int_{\mathbb{R}^{d_1}} w d\gamma_1 \right) \right] d\gamma_2 \\ & \quad + \int_{\mathbb{R}^{d_1}} \left[ \int_{\mathbb{R}^{d_2}} \varphi(w) d\gamma_2 - \varphi \left( \int_{\mathbb{R}^{d_2}} w d\gamma_2 \right) \right] d\gamma_1 \\ & \geq \iint_{\mathbb{R}^{d_1} \times \mathbb{R}^{d_2}} \varphi(w) d\gamma_1 \otimes \gamma_2 - \varphi \left( \iint_{\mathbb{R}^{d_1} \times \mathbb{R}^{d_2}} w d\gamma_1 \otimes \gamma_2 \right). \end{aligned}$$

This ends the proof of Proposition 2.2.  $\square$

**Proof of Theorem 2.1.** The proof is an easy consequence of Proposition 2.2 and of the observation that

$$\begin{aligned} & \min\{\Lambda_1, \Lambda_2\} \mathcal{E}_{\gamma_1 \otimes \gamma_2}[w] \\ & \leq \Lambda_1 \int_{\mathbb{R}^{d_2}} \mathcal{E}_{\gamma_1}[w] d\gamma_2 + \Lambda_2 \int_{\mathbb{R}^{d_1}} \mathcal{E}_{\gamma_2}[w] d\gamma_1 \\ & \leq \iint_{\mathbb{R}^{d_1} \times \mathbb{R}^{d_2}} \varphi''(w) [|\nabla_{x_1} w|^2 + |\nabla_{x_2} w|^2] d\gamma_1 \otimes \gamma_2 \\ & \leq \iint_{\mathbb{R}^{d_1} \times \mathbb{R}^{d_2}} \varphi''(w) |\nabla w|^2 d\gamma_1 \otimes \gamma_2 = \mathcal{I}_{\gamma_1 \otimes \gamma_2}[w]. \end{aligned}$$

$\square$

As a concluding remark, we observe that tensorization is not limited to probability measures on  $\mathbb{R}^d$ . The main interest of such an approach when dealing with  $\mathbb{R}^d$  is that it is enough to establish the inequality when  $d = 1$ . In the case  $d = 1$ , sharp criteria can be found in <sup>11</sup> (also see <sup>10</sup>). There are many related issues that can be traced back to the work of Muckenhoupt, *e.g.*, <sup>60</sup> and Hardy (see <sup>42</sup>).

### 2.3. Entropy – entropy production inequalities: perturbation results

Perturbing the measure in the case of a Poincaré inequality is essentially trivial. In the case of the logarithmic Sobolev inequality, this has been done by Holley and Stroock in <sup>47</sup>. More general entropy functionals have been considered in <sup>67</sup>, which cover all  $\varphi$ -entropies. Also see <sup>3,24</sup>.

Assume that for some probability measure  $d\gamma$  and for some  $\Lambda > 0$ , Inequality (1.3) holds, that is,

$$\Lambda \left[ \int_{\mathbb{R}^d} \varphi(w) d\gamma - \varphi(\bar{w}) \right] \leq \int_{\mathbb{R}^d} \varphi''(w) |\nabla w|^2 d\gamma \quad \forall w \in H^1(d\gamma). \quad (2.4)$$

Here we denote by  $\bar{w}$  the average of  $w$  with respect to  $d\gamma$ :  $\bar{w} := \int_{\mathbb{R}^d} w d\gamma$ . Assume that  $d\mu$  is a measure which is absolutely continuous with respect to  $d\gamma$  and such that

$$e^{-b} d\gamma \leq d\mu \leq e^{-a} d\gamma$$

for some constants  $a, b \in \mathbb{R}$ . The statement below generalizes the one of Lemma 5.2 of [16](#).

**Lemma 2.3.** *Under the above assumption, if  $\varphi$  is a  $C^2$  function such that  $\varphi'' > 0$ , then*

$$e^{a-b} \Lambda \int_{\mathbb{R}^d} [\varphi(w) - \varphi(\tilde{w}) - \varphi'(\tilde{w})(w - \tilde{w})] d\mu \leq \int_{\mathbb{R}^d} \varphi''(w) |\nabla w|^2 d\mu \quad \forall w \in \mathbb{H}^1(d\mu),$$

where  $\tilde{w} := \int_{\mathbb{R}^d} w d\mu / \int_{\mathbb{R}^d} d\mu$ .

**Proof.** We start by observing that

$$\begin{aligned} e^b \int_{\mathbb{R}^d} \varphi''(w) |\nabla w|^2 d\mu &\geq \int_{\mathbb{R}^d} \varphi''(w) |\nabla w|^2 d\gamma = \mathcal{I}_\gamma[w] \\ &\geq \Lambda \mathcal{E}_\gamma[w] = \Lambda \left[ \int_{\mathbb{R}^d} \varphi(w) d\gamma - \varphi(\bar{w}) \right] \\ &= \Lambda \int_{\mathbb{R}^d} (\varphi(w) - \varphi(\bar{w}) - \varphi'(\bar{w})(w - \bar{w})) d\gamma. \end{aligned}$$

By convexity of  $\varphi$ , we know that  $\varphi(w) - \varphi(\bar{w}) - \varphi'(\bar{w})(w - \bar{w}) \geq 0$ , so that

$$\begin{aligned} \Lambda \mathcal{E}_\gamma[w] &\geq \Lambda e^a \int_{\mathbb{R}^d} (\varphi(w) - \varphi(\bar{w}) - \varphi'(\bar{w})(w - \bar{w})) d\mu \\ &= \Lambda e^a \int_{\mathbb{R}^d} (\varphi(w) - \varphi(\bar{w}) - \varphi'(\bar{w})(\tilde{w} - \bar{w})) d\mu. \end{aligned}$$

By convexity of  $\varphi$  again,  $\varphi(\bar{w}) + \varphi'(\bar{w})(\tilde{w} - \bar{w}) \leq \varphi(\tilde{w})$ , which shows that

$$\Lambda \mathcal{E}_\gamma[w] \geq \Lambda e^a \int_{\mathbb{R}^d} (\varphi(w) - \varphi(\tilde{w})) d\mu = e^a \Lambda \int_{\mathbb{R}^d} [\varphi(w) - \varphi(\tilde{w}) - \varphi'(\tilde{w})(w - \tilde{w})] d\mu$$

and completes the proof.  $\square$

#### 2.4. Entropy – entropy production inequalities and linear flows

Let us consider the counterpart of the *Ornstein-Uhlenbeck equation* (1.2) on a smooth convex bounded domain  $\Omega$

$$\frac{\partial w}{\partial t} = \mathbb{L} w := \Delta w - \nabla \psi \cdot \nabla w, \quad (2.5)$$

supplemented with homogenous Neumann boundary conditions

$$\nabla w \cdot \nu = 0 \quad \text{on } \partial\Omega,$$

where  $\nu$  denotes a unit outward pointing normal vector orthogonal to  $\partial\Omega$ . Let us consider the measure  $d\gamma = \left( \int_{\Omega} e^{-\psi} dx \right)^{-1} e^{-\psi} dx$ . If  $w$  solves (2.5) with a non-negative initial datum  $w_0$  such that  $\int_{\Omega} w_0 d\gamma = 1$ , then mass is conserved so that  $\int_{\Omega} w(t, \cdot) d\gamma = 1$  for any  $t \geq 0$  and converges to 1 as  $t \rightarrow +\infty$ . The next question is how to measure the rate of convergence using the  $\varphi$ -entropy. For simplicity, let

us assume that  $\varphi = \varphi_p$  for some  $p \in [1, 2]$ . An answer is given by the formal computation of Section 1, adapted to the bounded domain  $\Omega$ . Because of the boundary condition, it is straightforward to check that

$$\frac{d}{dt} \int_{\Omega} \frac{w^p - 1}{p-1} d\gamma = -\frac{4}{p} \int_{\Omega} |\nabla w^{p/2}|^2 d\gamma$$

if  $p > 1$  and a similar results holds when  $p = 1$ . Hence, if for some  $\Lambda > 0$  we can prove that

$$\int_{\Omega} \frac{w^p - 1}{p-1} d\gamma \leq \frac{4}{p\Lambda} \int_{\Omega} |\nabla w^{p/2}|^2 d\gamma \quad \text{for any } w \text{ such that } \int_{\Omega} w d\gamma = 1, \quad (2.6)$$

then we can conclude that  $\int_{\Omega} \frac{w^p - 1}{p-1} d\gamma$  decays like  $e^{-\Lambda t}$ . The main idea of the Bakry-Emery method, or *carré du champ* method, as it is exposed in <sup>9</sup> is that (2.6) can be established using the flow itself, by computing  $\frac{d}{dt} \int_{\Omega} |\nabla z|^2 d\gamma$  with  $z := w^{p/2}$ . Let us sketch the main steps of the proof.

As a preliminary observation, we notice that  $\mathbb{L}$  is self-adjoint in  $L^2(\Omega, d\gamma)$  in the sense that

$$\int_{\Omega} w_1 (\mathbb{L} w_2) d\gamma = - \int_{\Omega} \nabla w_1 \cdot \nabla w_2 d\gamma = \int_{\Omega} (\mathbb{L} w_1) w_2 d\gamma$$

and also that

$$[\nabla, \mathbb{L}] = -\text{Hess } \psi.$$

Using  $w = z^{2/p}$  we deduce from (2.5) that

$$\frac{\partial z}{\partial t} = \mathbb{L} z + \frac{2-p}{p} \frac{|\nabla z|^2}{z}. \quad (2.7)$$

We adopt the convention that  $a \cdot b = \sum_{i=1}^d a_i b_i$  if  $a = (a_i)_{1 \leq i \leq d}$  and  $b = (b_i)_{1 \leq i \leq d}$  are two vectors with values in  $\mathbb{R}^d$ . If  $m = (m_{i,j})_{1 \leq i,j \leq d}$  and  $n = (n_{i,j})_{1 \leq i,j \leq d}$  are two matrices, then  $m : n = \sum_{i,j=1}^d m_{i,j} n_{i,j}$ . Also  $a \otimes b$  denotes the matrix  $(a_i b_j)_{1 \leq i,j \leq d}$ . We shall use  $|a|^2 = a \cdot a$  and  $\|m\|^2 = m : m$  for vectors and matrices respectively. With these notations, let us use (2.7) to compute

$$\begin{aligned} & \frac{1}{2} \frac{d}{dt} \int_{\Omega} |\nabla z|^2 d\gamma = \int_{\Omega} \nabla z \cdot \nabla \left( \mathbb{L} z + \frac{2-p}{p} \frac{|\nabla z|^2}{z} \right) d\gamma \\ &= \int_{\Omega} \nabla z \cdot (\mathbb{L} \nabla z - \text{Hess } \psi \nabla z) d\gamma + \frac{2-p}{p} \int_{\Omega} \nabla z \cdot \left( 2 \text{Hess } z \frac{\nabla z}{z} - \frac{|\nabla z|^2}{z} \nabla z \right) d\gamma \\ &= - \int_{\Omega} \|\text{Hess } z\|^2 d\gamma - \int_{\Omega} \text{Hess } \psi : \nabla z \otimes \nabla z d\gamma + \int_{\partial\Omega} \text{Hess } z : \nabla z \otimes \nu e^{-\psi} d\sigma \\ & \quad + 2 \frac{2-p}{p} \int_{\Omega} \text{Hess } z : \frac{\nabla z \otimes \nabla z}{z} d\gamma - \frac{2-p}{p} \int_{\Omega} \left\| \frac{\nabla z \otimes \nabla z}{z} \right\|^2 d\gamma \\ &= -\frac{2}{p} (p-1) \int_{\Omega} \|\text{Hess } z\|^2 d\gamma - \int_{\Omega} \text{Hess } \psi : \nabla z \otimes \nabla z d\gamma \\ & \quad - \frac{2-p}{p} \int_{\Omega} \left\| \text{Hess } z - \frac{\nabla z \otimes \nabla z}{z} \right\|^2 d\gamma + \int_{\partial\Omega} \text{Hess } z : \nabla z \otimes \nu e^{-\psi} d\sigma. \end{aligned}$$

Here  $d\sigma$  denotes the surface measure induced by Lebesgue's measure on  $\partial\Omega$ . We learn from Grisvard's lemma, see for instance Lemma 5.1 in <sup>39</sup> or <sup>40</sup>, that  $\int_{\partial\Omega} \text{Hess } z : \nabla z \otimes \nu e^{-\psi} d\sigma$  is nonpositive as soon as  $\Omega$  is convex and  $\nabla z \cdot \nu = 0$  on  $\partial\Omega$ . As soon as we know that either

$$\text{Hess } \psi \geq \Lambda_\star \text{Id}$$

for some  $\Lambda_\star > 0$ , or the inequality

$$\frac{2}{p} (p-1) \int_{\Omega} |\nabla X|^2 d\gamma + \int_{\Omega} \text{Hess } \psi : X \otimes X d\gamma \geq \Lambda(p) \int_{\Omega} |X|^2 d\gamma \quad \forall X \in \text{H}^1(\Omega, d\gamma)^d$$

holds for some  $\Lambda(p) > 0$ , which is a weaker assumption for any  $p > 1$ , then we obtain that

$$\frac{d}{dt} \int_{\Omega} |\nabla z|^2 d\gamma \leq -2\Lambda(p) \int_{\Omega} |\nabla z|^2 d\gamma.$$

Of course we know that  $\Lambda(p) \geq \Lambda_\star$ . By convention, we take  $\Lambda(1) = \Lambda_\star$ .

**Proposition 2.3.** *Assume that  $p \in [1, 2]$ ,  $\varphi = \varphi_p$  and, with the above notations,  $\Lambda(p) > 0$ . If  $\Omega$  is a smooth convex bounded domain in  $\mathbb{R}^d$ , then (2.6) holds with  $\Lambda = 2\Lambda(p)$ .*

**Proof.** It is straightforward. In view of the above computations, we know that

$$\frac{d}{dt} \left( \frac{4}{p\Lambda} \int_{\Omega} |\nabla w^{p/2}|^2 d\gamma - \int_{\Omega} \frac{w^p - 1}{p-1} d\gamma \right) \leq 0$$

and  $\lim_{t \rightarrow +\infty} \int_{\Omega} \frac{w^p - 1}{p-1} d\gamma = \lim_{t \rightarrow +\infty} \int_{\Omega} |\nabla w^{p/2}|^2 d\gamma = 0$ . This is enough to conclude that, for any  $t \geq 0$ ,

$$\frac{4}{p\Lambda} \int_{\Omega} |\nabla w^{p/2}|^2 d\gamma - \int_{\Omega} \frac{w^p - 1}{p-1} d\gamma \geq 0. \quad \square$$

We conclude this section with the unbounded case  $\Omega = \mathbb{R}^d$ . For any given  $p \in [1, 2]$ , let us assume that the inequality

$$\frac{2}{p} (p-1) \int_{\mathbb{R}^d} |\nabla X|^2 d\gamma + \int_{\mathbb{R}^d} \text{Hess } \psi : X \otimes X d\gamma \geq \Lambda(p) \int_{\mathbb{R}^d} |X|^2 d\gamma \quad \forall X \in \text{H}^1(\mathbb{R}^d, d\gamma)^d$$

holds for some  $\Lambda(p) > 0$ . For  $p > 1$ , this assumption is a spectral gap condition on a vector valued Schrödinger operator: see for instance <sup>32</sup> for further details. With this assumption in hand, we have the following functional inequality, which interpolates between the logarithmic Sobolev inequality and the Poincaré inequality.

**Corollary 2.1.** *Assume that  $q \in [1, 2)$  and let us consider the probability measure  $d\gamma = e^{-\psi} dx$  on  $\mathbb{R}^d$ . Then with  $\Lambda = \Lambda(2/q)$ , we have*

$$\frac{\|f\|_{\text{L}^2(\mathbb{R}^d, d\gamma)}^2 - \|f\|_{\text{L}^q(\mathbb{R}^d, d\gamma)}^2}{2-q} \leq \frac{1}{\Lambda} \int_{\mathbb{R}^d} |\nabla f|^2 d\gamma \quad \forall f \in \text{H}^1(\mathbb{R}^d, d\gamma). \quad (2.8)$$

**Proof.** By homogeneity, we know from Proposition 2.3 that

$$\int_{\Omega} \frac{w^p - \bar{w}^p}{p-1} d\gamma \leq \frac{2}{p\Lambda(p)} \int_{\Omega} |\nabla w^{p/2}|^2 d\gamma$$

for all  $w$  such that  $f = w^{p/2}$ . Here we take  $p = 2/q$ . The conclusion holds by approximating  $\mathbb{R}^d$  by a growing sequence of bounded convex domains.  $\square$

An equivalent form of (2.8) is

$$\mathcal{I}[w] \geq \Lambda \mathcal{E}[w] \quad \forall w \in H^1(\mathbb{R}^d, d\gamma) \text{ such that } \int_{\mathbb{R}^d} w d\gamma = 1 \quad (2.9)$$

with the notation of Section 1,  $\varphi = \varphi_p$  and  $p = 2/q \in [1, 2]$ .

**Remark 2.1.** The optimality of the constant  $\Lambda = 1$  in (2.8) is easy to obtain when  $\psi(x) = \frac{1}{2}|x|^2$ . With  $q = 1$ , (2.8) is the Gaussian Poincaré inequality

$$\|f - \bar{f}\|_{L^2(\mathbb{R}^d, d\gamma)}^2 \leq \int_{\mathbb{R}^d} |\nabla f|^2 d\gamma \quad \forall f \in H^1(\mathbb{R}^d, d\gamma) \quad \text{with } \bar{f} = \int_{\mathbb{R}^d} f d\gamma,$$

with equality if  $f = f_1$ ,  $f_1(x) = x_1$ . By taking the limit as  $q \rightarrow 2_-$  in (2.8), we recover Gross' logarithmic Sobolev inequality

$$\int_{\mathbb{R}^d} f^2 \log \left( \frac{f^2}{\|f\|_{L^2(\mathbb{R}^d, d\gamma)}^2} \right) d\gamma \leq 2 \int_{\mathbb{R}^d} |\nabla f|^2 d\gamma \quad \forall f \in H^1(\mathbb{R}^d, d\gamma).$$

For any  $q \in [1, 2)$ , the equality case in (2.8) with  $\Lambda = 1$  is achieved by considering  $1 + \varepsilon f_1$  as a test function in the limit as  $\varepsilon \rightarrow 0$ .

From the point of view of the evolution equation, it is easy to see that the equality in (2.6) is achieved asymptotically as  $t \rightarrow +\infty$  by taking  $w = u/u_*$  where  $u$  is the solution of (1.4) given by

$$u(t, x) = u_*(x - x_*(t))$$

with  $x_*(t) = x_0 e^{-t}$  for any fixed  $x_0 \in \mathbb{R}^d$ .

## 2.5. Improved entropy – entropy production inequalities

In the proof of Proposition 2.3, the term  $\int_{\mathbb{R}^d} \|\text{Hess } z - \nabla z \otimes \nabla z / z\|^2 d\gamma$  has been dropped. In some cases, one can recombine the other terms differently and obtain an improved inequality if  $q \in (1, 2)$ . See <sup>5</sup> (and also <sup>4</sup> for a spectral point of view or <sup>27</sup> in the case of the sphere). The boundary term  $\int_{\partial\Omega} \text{Hess } z : \nabla z \otimes \nu e^{-\psi} d\sigma$  may also be of importance, as it is suggested in nonlinear problems by <sup>28</sup>.

Let us give an example of an improvement, based on <sup>5</sup>, in the special case  $\psi(x) = |x|^2/2$ . Using  $\text{Hess } \psi = \text{Id}$ , after approximating  $\mathbb{R}^d$  by bounded domains, we obtain that

$$\begin{aligned} \frac{1}{2} \frac{d}{dt} \int_{\mathbb{R}^d} |\nabla z|^2 d\gamma + \int_{\mathbb{R}^d} |\nabla z|^2 d\gamma \leq - \int_{\mathbb{R}^d} \left\| \text{Hess } z - \frac{2-p}{p} \frac{\nabla z \otimes \nabla z}{z} \right\|^2 d\gamma \\ - \frac{2}{p} \kappa_p \int_{\mathbb{R}^d} \frac{|\nabla z|^4}{z^2} d\gamma \end{aligned}$$

with  $\kappa_p = (p-1)(2-p)/p$ . A simple Cauchy-Schwarz inequality shows that

$$\left( \int_{\mathbb{R}^d} |\nabla z|^2 d\gamma \right)^2 \leq \int_{\mathbb{R}^d} \frac{|\nabla z|^4}{z^2} d\gamma \int_{\mathbb{R}^d} z^2 d\gamma.$$

With the notations of Section 1, we have  $\int_{\mathbb{R}^d} z^2 d\gamma = \int_{\mathbb{R}^d} w^p d\gamma = 1 + (p-1)\mathcal{E}[w]$  and  $\int_{\mathbb{R}^d} |\nabla z|^2 d\gamma = \frac{p}{4}\mathcal{I}[w]$  so that

$$\frac{1}{2} \frac{d}{dt} \int_{\mathbb{R}^d} |\nabla z|^2 d\gamma + \int_{\mathbb{R}^d} |\nabla z|^2 d\gamma \leq -\frac{2}{p} \kappa_p \frac{\left( \int_{\mathbb{R}^d} |\nabla z|^2 d\gamma \right)^2}{\int_{\mathbb{R}^d} |z|^2 d\gamma}$$

can be rewritten as

$$\frac{d}{dt} \mathcal{I}[w] + 2\mathcal{I}[w] \leq -\kappa_p \frac{\mathcal{I}[w]^2}{1 + (p-1)\mathcal{E}[w]}. \quad (2.10)$$

We recall that we consider here the case  $\varphi = \varphi_p$ ,  $p \in (1, 2)$ , so that  $\kappa_p$  is positive and we can take advantage of (2.10) to obtain an improved version of Corollary 2.1. The following result follows the scheme of Theorem 2 in <sup>5</sup>.

**Proposition 2.4.** *Assume that  $q \in (1, 2)$  and let us consider the Gaussian probability measure  $d\gamma = (2\pi)^{-d/2} e^{-|x|^2/2} dx$ . Then there exists a strictly convex function  $F$  on  $\mathbb{R}^+$  such that  $F(0) = 0$  and  $F'(0) = 1$ , for which*

$$\frac{1}{q} F \left( q \frac{\|f\|_{L^2(\mathbb{R}^d, d\gamma)}^2 - 1}{2 - q} \right) \leq \|\nabla f\|_{L^2(\mathbb{R}^d, d\gamma)}^2$$

for any  $f \in H^1(\mathbb{R}^d, d\gamma)$  such that  $\|f\|_{L^q(\mathbb{R}^d, d\gamma)} = 1$ .

**Proof.** The proof follows the strategy of <sup>5</sup>. Let  $\mathbf{e}(t) := \frac{1}{p-1} \left( \int_{\mathbb{R}^d} z^2 d\gamma - 1 \right)$  where  $z = w^{p/2}$  solves (2.7) with initial datum  $f$ . We deduce from (2.10) that

$$\mathbf{e}'' + 2\mathbf{e}' \geq \frac{\kappa_p |\mathbf{e}'|^2}{1 + (p-1)\mathbf{e}} \geq \frac{\kappa_p |\mathbf{e}'|^2}{1 + \mathbf{e}}.$$

The function  $F(s) := \frac{1}{1-\kappa_p} [1 + s - (1+s)^{\kappa_p}]$  solves  $F' = 1 + \kappa_p \frac{F}{1+s}$  and we can check that (2.10) is equivalent to

$$\frac{d}{dt} \left( \frac{\mathbf{e}' + 2F(\mathbf{e})}{(1+\mathbf{e})^{\kappa_p}} \right) \geq 0.$$

Since  $\lim_{t \rightarrow +\infty} (\mathbf{e}'(t) + 2F(\mathbf{e}(t))) = 0$ , we have shown that  $\mathbf{e}' + 2F(\mathbf{e}) \leq 0$  for any  $t \geq 0$ . This is true in particular at  $t = 0$ , with  $z(t=0, \cdot) = f$ .  $\square$

From the point of view of entropy – production of entropy inequalities, we have obtained that

$$\mathcal{I}[w] \geq 2F(\mathcal{E}[w])$$



where  $F$  is a strictly convex function such that  $F(0) = 0$  and  $F'(0) = 1$ . Using the homogeneity and substituting  $f/\|f\|_{L^q(\mathbb{R}^d, d\gamma)}$  to  $f$ , similar estimates have been used in <sup>5</sup> to prove that

$$\frac{2}{(2-q)^2} \left[ \|f\|_{L^2(\mathbb{R}^d, d\gamma)}^2 - \|f\|_{L^q(\mathbb{R}^d, d\gamma)}^{2(2-q)} \|f\|_{L^2(\mathbb{R}^d, d\gamma)}^{2(q-1)} \right] \leq \|\nabla f\|_{L^2(\mathbb{R}^d, d\gamma)}^2 \quad \forall f \in H^1(\mathbb{R}^d, d\gamma).$$

### 2.6. Interpolation inequalities: comments and extensions

The inequality of Corollary 2.1 appears in many papers. It is proved for the first time by the *carré du champ* method and any  $q \in [1, 2]$  in <sup>9</sup> in the case of a compact manifold, but special cases were known long before. For instance the case  $q = 2$  corresponding to the logarithmic Sobolev inequality can be traced back to <sup>41,38</sup> (also see <sup>72,35</sup> for related issues) but was already known as the Blachmann-Stam inequality <sup>65</sup>: see <sup>70,66</sup> for a more detailed historical account. The case  $q = 1$  when  $\psi(x) = \frac{1}{2}|x|^2$  is known as the Gaussian Poincaré inequality. It appears for instance in <sup>61</sup> but was probably known much earlier in the framework of the theory of Hermite functions. In the case  $q \in (1, 2)$  when  $\psi(x) = \frac{1}{2}|x|^2$ , we may refer to <sup>15</sup> for a proof based on spectral methods, which has been extended in <sup>4</sup> to more general potentials.

One of the technical limitations of the *carré du champ* method is the difficulty of controlling the boundary terms in the various integrations by parts. In the above proof, we used Grisvard's lemma for convex domains. Alternative methods, which will not be exposed here, rely on the properties of Green's functions, or use direct spectral estimates.

Let us list some possible extensions:

- In Corollary 2.1, for any given  $q \in [1, 2]$ , we need that  $\Lambda(p)$  is positive only for  $p = 2/q$ . The condition for  $p = 1$ , which is equivalent to  $\text{Hess } \psi \geq \Lambda(1)\text{Id}$  with  $\Lambda(1) > 0$ , is not required unless  $q = 2$ . For any  $q < 2$ , the positivity condition of  $\Lambda(2/q)$  is a nonlocal condition, which allows  $\psi$  to be a non-uniformly strictly convex potential: see <sup>32</sup> for details.
- The case of unbounded convex domains can be considered. Reciprocally, according to <sup>7</sup>, the case of a bounded convex domain  $\Omega$  can be deduced from the Euclidean case, by approximating a function  $\psi$  which takes the value  $+\infty$  on  $\Omega^c$  by smooth locally bounded potentials.
- Spectral methods can be used to establish that the family of inequalities of Corollary 2.1 interpolates between the logarithmic Sobolev inequality and the Poincaré inequality: this approach has been made precise in <sup>15,54</sup>, with extensions in <sup>12,4</sup>.
- Exhibiting a whole family of Lyapunov functionals for the same evolution equation needs an explanation that has been given in <sup>31,33</sup>: to each entropy, we associate a notion of distance such that the equation appears as the gradient flow of the entropy.

In the context of linear diffusions and Markov processes,  $\varphi$ -entropies are very natural objects which put the Gibbs entropy and the quadratic form associated

to the Poincaré inequality in a common framework. It is therefore evident to ask the same question in a kinetic framework involving a degenerate diffusion operator coupled to a transport operator. Much less has been done so far and the next section is a contribution to the issue of optimal rates of convergence measured by  $\varphi_p$ -entropies, with a special emphasis on  $p \neq 1, 2$ .

### 3. Sharp rates for the kinetic Fokker-Planck equation

In this section, our goal is to provide a computation of the sharp exponential rate in Proposition 1.1 and establish the improvement of Theorem 1.1 by generalizing the estimate of Proposition 2.4 to the kinetic setting. The method follows the strategy of Section 3 of <sup>69</sup> in case  $p = 2$ , which is sometimes referred to as the  $H^1$  *hypocoercivity method of C. Villani*. This method is also known to cover the case  $p = 1$ . We extend it to any  $p \in [1, 2]$  and compute the precise algebraic expressions, which allows us to identify the sharp rate. Similar computations have been done in <sup>6,2,1,37,58</sup>. According to <sup>23</sup> (see earlier references therein), the Green function associated with (1.5) is a Gaussian kernel, so that integrations by parts can be performed on  $\mathbb{R}^d \times \mathbb{R}^d$  without any special precaution.

#### 3.1. $H^1$ hypocoercive estimates

Using the notation of Section 1, our strategy is to consider the solution  $h = g^{p/2}$  of (1.7), where  $g = f/f_*$ , define

$$\mathcal{J}[h] := \int_{\mathbb{R}^d} |\nabla_v h|^2 d\mu + 2\lambda \int_{\mathbb{R}^d} \nabla_v h \cdot \nabla_x h d\mu + \nu \int_{\mathbb{R}^d} |\nabla_x h|^2 d\mu$$

and adjust the parameters  $\lambda$  and  $\nu$  in order to maximize  $\lambda_* = \lambda_*(\lambda, \nu) > 0$  so that

$$\frac{d}{dt} \mathcal{J}[h(t, \cdot, \cdot)] \leq -\lambda_*(\lambda, \nu) \mathcal{J}[h(t, \cdot, \cdot)].$$

Since (1.6) is linear and preserves positivity, we recall that we can assume that  $g$  is nonnegative and such that  $\|g\|_{L^1(\mathbb{R}^d \times \mathbb{R}^d, d\mu)} = 1$ . Let us define the notations:

$$\begin{aligned} H_{vv} &= \left( \frac{\partial^2 h}{\partial v_i \partial v_j} \right)_{1 \leq i, j \leq d}, & H_{xv} &= \left( \frac{\partial^2 h}{\partial x_i \partial v_j} \right)_{1 \leq i, j \leq d}, \\ M_{vv} &= \left( \frac{\partial \sqrt{h}}{\partial v_i} \frac{\partial \sqrt{h}}{\partial v_j} \right)_{1 \leq i, j \leq d}, & M_{xv} &= \left( \frac{\partial \sqrt{h}}{\partial x_i} \frac{\partial \sqrt{h}}{\partial v_j} \right)_{1 \leq i, j \leq d}. \end{aligned}$$

We start by observing that, up to a few integrations by parts, we obtain the identities

$$\begin{aligned}
 & \frac{1}{2} \frac{d}{dt} \int_{\mathbb{R}^d} |\nabla_v h|^2 d\mu \\
 &= - \int_{\mathbb{R}^d} \nabla_v h \cdot \nabla_v (v \cdot \nabla_x h - x \cdot \nabla_v h) d\mu + \int_{\mathbb{R}^d} \nabla_v h \cdot \nabla_v (\Delta_v h - v \cdot \nabla_v h) d\mu \\
 &\quad + \left(\frac{2}{p} - 1\right) \int_{\mathbb{R}^d} \nabla_v h \cdot \nabla_v \left(\frac{|\nabla_v h|^2}{h}\right) d\mu \\
 &= - \int_{\mathbb{R}^d} \nabla_v h \cdot \nabla_x h d\mu - \left( \int_{\mathbb{R}^d} \|\mathbf{H}_{vv}\|^2 d\mu + \int_{\mathbb{R}^d} |\nabla_v h|^2 d\mu \right) \\
 &\quad + \kappa \int_{\mathbb{R}^d} (\mathbf{H}_{vv} : \mathbf{M}_{vv} - 2 \|\mathbf{M}_{vv}\|^2) d\mu \quad (3.1)
 \end{aligned}$$

with  $\kappa = 8(2-p)/p$ ,

$$\begin{aligned}
 & \frac{1}{2} \frac{d}{dt} \int_{\mathbb{R}^d} |\nabla_x h|^2 d\mu \\
 &= - \int_{\mathbb{R}^d} \nabla_x h \cdot \nabla_x (v \cdot \nabla_x h - x \cdot \nabla_v h) d\mu + \int_{\mathbb{R}^d} \nabla_x h \cdot \nabla_x (\Delta_v h - v \cdot \nabla_v h) d\mu \\
 &\quad + \left(\frac{2}{p} - 1\right) \int_{\mathbb{R}^d} \nabla_x h \cdot \nabla_x \left(\frac{|\nabla_v h|^2}{h}\right) d\mu \\
 &= \int_{\mathbb{R}^d} \nabla_v h \cdot \nabla_x h d\mu - \int_{\mathbb{R}^d} \|\mathbf{H}_{xv}\|^2 d\mu + \kappa \int_{\mathbb{R}^d} (\mathbf{H}_{xv} : \mathbf{M}_{xv} - 2 \|\mathbf{M}_{xv}\|^2) d\mu, \quad (3.2)
 \end{aligned}$$

and

$$\begin{aligned}
 \frac{d}{dt} \int_{\mathbb{R}^d} \nabla_v h \cdot \nabla_x h d\mu &= \int_{\mathbb{R}^d} |\nabla_v h|^2 d\mu - \int_{\mathbb{R}^d} |\nabla_x h|^2 d\mu - \int_{\mathbb{R}^d} \nabla_v h \cdot \nabla_x h d\mu \\
 &\quad - 2 \int_{\mathbb{R}^d} \mathbf{H}_{vv} : \mathbf{H}_{xv} d\mu \\
 &\quad + \kappa \int_{\mathbb{R}^d} (\mathbf{H}_{vv} : \mathbf{M}_{xv} + \mathbf{H}_{xv} : \mathbf{M}_{vv} - 4 \mathbf{M}_{vv} : \mathbf{M}_{xv}) d\mu. \quad (3.3)
 \end{aligned}$$

Collecting these estimates shows that

$$\begin{aligned}
 & - \frac{1}{2} \frac{d}{dt} \mathcal{J}[h(t, \cdot, \cdot)] \\
 &= - \frac{1}{2} \frac{d}{dt} \left( \int_{\mathbb{R}^d} |\nabla_v h|^2 d\mu + 2\lambda \int_{\mathbb{R}^d} \nabla_v h \cdot \nabla_x h d\mu + \nu \int_{\mathbb{R}^d} |\nabla_x h|^2 d\mu \right) \\
 &= (1-\lambda) \int_{\mathbb{R}^d} |\nabla_v h|^2 d\mu + (1+\lambda-\nu) \int_{\mathbb{R}^d} \nabla_v h \cdot \nabla_x h d\mu + \lambda \int_{\mathbb{R}^d} |\nabla_x h|^2 d\mu \\
 &\quad + \int_{\mathbb{R}^d} \|\mathbf{H}_{vv}\|^2 d\mu - \kappa \int_{\mathbb{R}^d} (\mathbf{H}_{vv} : \mathbf{M}_{vv} - 2 \|\mathbf{M}_{vv}\|^2) d\mu \\
 &+ 2\lambda \int_{\mathbb{R}^d} \mathbf{H}_{vv} : \mathbf{H}_{xv} d\mu - \kappa \lambda \int_{\mathbb{R}^d} (\mathbf{H}_{vv} : \mathbf{M}_{xv} + \mathbf{H}_{xv} : \mathbf{M}_{vv} - 4 \mathbf{M}_{vv} : \mathbf{M}_{xv}) d\mu \\
 &\quad + \nu \int_{\mathbb{R}^d} \|\mathbf{H}_{xv}\|^2 d\mu - \kappa \nu \int_{\mathbb{R}^d} (\mathbf{H}_{xv} : \mathbf{M}_{xv} - 2 \|\mathbf{M}_{xv}\|^2) d\mu
 \end{aligned}$$

where  $\kappa = 8(2-p)/p$ . This can be rewritten as

$$-\frac{1}{2} \frac{d}{dt} \int_{\mathbb{R}^d} X^\perp \cdot \mathfrak{M}_0 X \, d\mu = \int_{\mathbb{R}^d} X^\perp \cdot \mathfrak{M}_1 X \, d\mu + \int_{\mathbb{R}^d} Y^\perp \cdot \mathfrak{M}_2 Y \, d\mu$$

where

$$\mathfrak{M}_0 = \begin{pmatrix} 1 & \lambda \\ \lambda & \nu \end{pmatrix} \otimes \text{Id}_{\mathbb{R}^d}, \quad \mathfrak{M}_1 = \begin{pmatrix} 1-\lambda & \frac{1+\lambda-\nu}{2} \\ \frac{1+\lambda-\nu}{2} & \lambda \end{pmatrix} \otimes \text{Id}_{\mathbb{R}^d}$$

and

$$\mathfrak{M}_2 = \begin{pmatrix} 1 & \lambda & -\frac{\kappa}{2} & -\frac{\kappa\lambda}{2} \\ \lambda & \nu & -\frac{\kappa\lambda}{2} & -\frac{\kappa\nu}{2} \\ -\frac{\kappa}{2} & -\frac{\kappa\lambda}{2} & 2\kappa & 2\kappa\lambda \\ -\frac{\kappa\lambda}{2} & -\frac{\kappa\nu}{2} & 2\kappa\lambda & 2\kappa\nu \end{pmatrix} \otimes \text{Id}_{\mathbb{R}^d \times \mathbb{R}^d}$$

are bloc-matrix valued functions of  $(\lambda, \nu)$ , and

$$X = (\nabla_v h, \nabla_x h), \quad Y = (H_{vv}, H_{xv}, M_{vv}, M_{xv}).$$

The problem is reduced to a problem of linear algebra, namely to maximize

$$\lambda_*(\lambda, \nu) := \min_{X \in \mathbb{R}^{2d}} \frac{X^\perp \cdot \mathfrak{M}_1(\lambda, \nu) X}{X^\perp \cdot \mathfrak{M}_0(\lambda, \nu) X}$$

on the set of parameters  $(\lambda, \nu) \in \mathbb{R}^2$  such that

$$\min_{Y \in \mathbb{R}^{2d} \times \mathbb{R}^{2d}} \frac{Y^\perp \cdot \mathfrak{M}_2 Y}{\|Y\|^2} \geq 0.$$

Here  $X$  and  $Y$  now arbitrary vectors and matrices respectively in  $\mathbb{R}^{2d}$  and  $\mathbb{R}^{2d} \times \mathbb{R}^{2d}$ . Elementary computations show that  $\lambda$  and  $\nu$  must satisfy the condition  $\lambda^2 \leq \nu$  and also that  $\lambda_*(\lambda, \nu)$  achieves its maximum at  $(\lambda, \nu) = (\frac{1}{2}, 1)$ , so that  $\lambda_*(\frac{1}{2}, 1) = \frac{1}{2}$ . For  $(\lambda, \nu) = (\frac{1}{2}, 1)$ ,  $\mathfrak{M}_1(\frac{1}{2}, 1) = \frac{1}{2} \mathfrak{M}_0(\frac{1}{2}, 1)$  and the eigenvalues of  $\mathfrak{M}_2(\frac{1}{2}, 1)$  are given as a function of  $\kappa = 8(2-p)/p$  by

$$\lambda_1(\kappa) := \frac{1}{4} \left( 2\kappa + 1 - \sqrt{5\kappa^2 - 4\kappa + 1} \right), \quad \lambda_2(\kappa) := \frac{3}{4} \left( 2\kappa + 1 - \sqrt{5\kappa^2 - 4\kappa + 1} \right),$$

$$\lambda_3(\kappa) := \frac{1}{4} \left( 2\kappa + 1 + \sqrt{5\kappa^2 - 4\kappa + 1} \right), \quad \lambda_4(\kappa) := \frac{3}{4} \left( 2\kappa + 1 + \sqrt{5\kappa^2 - 4\kappa + 1} \right).$$

In the range  $p \in [1, 2]$ , which means  $\kappa \in [0, 8]$ , they are all nonnegative: see Fig. 1. Since  $\lambda_1(\kappa)$  is the lowest eigenvalue, we have proved the following result.

**Lemma 3.1.** *With the above notations and  $(\lambda, \nu) = (\frac{1}{2}, 1)$ , we have the estimate*

$$\begin{aligned} \int_{\mathbb{R}^d} X^\perp \cdot \mathfrak{M}_1 X \, d\mu + \int_{\mathbb{R}^d} Y^\perp \cdot \mathfrak{M}_2 Y \, d\mu &\geq \frac{1}{2} \int_{\mathbb{R}^d} X^\perp \cdot \mathfrak{M}_0 X \, d\mu \\ &\quad + \frac{1}{4} \left( 2\kappa + 1 - \sqrt{5\kappa^2 - 4\kappa + 1} \right) |Y|^2. \end{aligned}$$

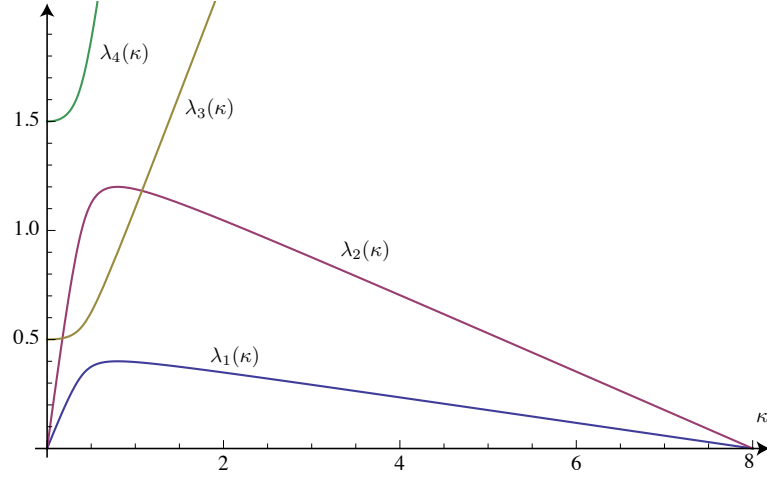


Fig. 1. Plot of the eigenvalues of  $\mathfrak{M}_2(\frac{1}{2}, 1)$  as a function of  $\kappa$ .

### 3.2. Proof of Proposition 1.1

Assume that  $h$  solves (1.7). With  $(\lambda, \nu) = (\frac{1}{2}, 1)$ , we deduce from Lemma 3.1 that  $\mathcal{J}[h]$  is defined by (1.9). Then it satisfies the differential inequality

$$\frac{d}{dt} \mathcal{J}[h(t, \cdot, \cdot)] \leq -\mathcal{J}[h(t, \cdot, \cdot)],$$

from which we deduce that

$$\mathcal{J}[h(t, \cdot, \cdot)] \leq \mathcal{J}[h(0, \cdot, \cdot)] e^{-t} \quad \forall t \geq 0.$$

Using (1.3) with  $d\gamma = \mu dx dv$ ,  $\lambda = 1$  and  $\varphi = \varphi_p$  for any  $p \in [1, 2]$  (also see Remark 2.1), we obtain that

$$\mathcal{E}[h(t, \cdot, \cdot)] \leq \mathcal{J}[h_0] e^{-t} \quad \forall t \geq 0$$

if  $h$  is the solution of (1.7) with initial datum  $h_0$ .

The optimality of the rate is established by considering an initial datum which is a decentred stationary solution. With the notations of Section 1, let

$$f_0(x, v) = f_\star(x - x_0, v - v_0) \quad \forall (x, v) \in \mathbb{R}^d \times \mathbb{R}^d$$

for some  $(x_0, v_0) \neq (0, 0)$ . The reader is invited to check that

$$f(t, x, v) = f_\star(x - x_\star(t), v - v_\star(t))$$

$$\text{with } \begin{cases} x_\star(t) = \left( \cos\left(\frac{\sqrt{3}}{2}t\right) x_0 + \frac{2}{\sqrt{3}} \sin\left(\frac{\sqrt{3}}{2}t\right) (v_0 + \frac{x_0}{2}) \right) e^{-\frac{t}{2}}, \\ v_\star(t) = \left( -\frac{\sqrt{3}}{2} \sin\left(\frac{\sqrt{3}}{2}t\right) (x_0 + \frac{v_0}{2}) + \cos\left(\frac{\sqrt{3}}{2}t\right) v_0 \right) e^{-\frac{t}{2}}, \end{cases} \quad (3.4)$$

solves (1.5). Now let us compute the entropy as  $t \rightarrow +\infty$ : with  $g = f/f_\star$  and  $\varphi = \varphi_p$ , we obtain that, as  $t \rightarrow +\infty$ ,

$$\begin{aligned} \mathcal{E}[g(t, \cdot, \cdot)] &= \iint_{\mathbb{R}^d \times \mathbb{R}^d} \varphi_p(g) d\mu = \frac{p}{2} \iint_{\mathbb{R}^d \times \mathbb{R}^d} |g - 1|^2 d\mu (1 + o(1)) \\ &= \frac{p}{2} (|x_\star(t)|^2 + |v_\star(t)|^2) (1 + o(1)) = O(e^{-t}). \end{aligned}$$

This proves that the rate  $e^{-t}$  of Proposition 1.1 is optimal and completes the proof.  $\square$

Compared to the proof of Proposition 1.1, a refined estimate can be obtained by observing that, in the computation of  $\frac{d}{dt} \int_{\mathbb{R}^d} |\nabla_v h|^2 d\mu$  and  $\frac{d}{dt} \int_{\mathbb{R}^d} |\nabla_x h|^2 d\mu$ , we have

$$\begin{aligned} \|H_{vv}\|^2 - \kappa H_{vv} : M_{vv} + 2\kappa \|M_{vv}\|^2 &\geq 0, \\ \|H_{xv}\|^2 - \kappa H_{xv} : M_{xv} + 2\kappa \|M_{xv}\|^2 &\geq 0, \end{aligned}$$

with  $\kappa = 8(2-p)/p$ . Let us define

$$\begin{aligned} \mathbf{a} &:= e^t \int_{\mathbb{R}^d} |\nabla_v h|^2 d\mu, \quad \mathbf{b} := e^t \int_{\mathbb{R}^d} \nabla_v h \cdot \nabla_x h d\mu, \quad \mathbf{c} := e^t \int_{\mathbb{R}^d} |\nabla_x h|^2 d\mu, \\ &\text{and } \mathbf{j} := \mathbf{a} + \mathbf{b} + \mathbf{c}. \end{aligned}$$

We deduce from (3.1), (3.2) and (3.3) that

$$\frac{d\mathbf{a}}{dt} \leq \mathbf{a} - 2(\mathbf{j} - \mathbf{c}), \quad \frac{d\mathbf{c}}{dt} \leq 2(\mathbf{j} - \mathbf{a}) - \mathbf{c} \quad \text{and} \quad \frac{d\mathbf{j}}{dt} \leq 0$$

while we know by definition of  $\mathbf{a}$ ,  $\mathbf{b}$  and  $\mathbf{c}$  and by the Cauchy-Schwarz estimate that

$$\mathbf{a} \geq 0, \quad \mathbf{c} \geq 0 \quad \text{and} \quad \mathbf{b}^2 \leq \mathbf{a}\mathbf{c}.$$

In terms of  $\mathbf{a}$  and  $\mathbf{c}$ , the inequality  $\mathbf{b}^2 = (\mathbf{a} + \mathbf{c} - \mathbf{j})^2 \leq \mathbf{a}\mathbf{c}$  means that the problem is constrained to the interior of an ellipse, and that  $\mathbf{a} = 0$  if and only if  $\mathbf{c} = \mathbf{j}$ : see Fig. 2. Finally, let us observe that we have the following property.

**Lemma 3.2.** *Assume that  $p \in [1, 2]$ ,  $\psi(x) = |x|^2/2$  and let  $h$  be a solution of (1.7) with initial datum  $h_0 \in L^1 \cap L^p(\mathbb{R}^d, d\gamma)$ . With the above notations, if for some  $t_0 > 0$ ,  $\mathbf{a}(t_0) = 0$  and  $\mathbf{j}(t_0) \neq 0$ , then for any  $t > t_0$  with  $t - t_0$  small enough, we have  $\mathbf{a}(t) > 0$ .*

**Proof.** From the equivalence of (1.5) and (1.7), we know that  $h$  is smooth because of the expression of Green's function. By definition of  $\mathbf{b}$  and  $\mathbf{j}$ , we have that  $\mathbf{b}(t_0) = 0$  and  $\mathbf{c}(t_0) = \mathbf{j}(t_0) > 0$ . Since  $\mathbf{a}(t_0) = 0$  means that  $h$  does not depend on  $v$ , we know that  $\frac{d\mathbf{j}}{dt}(t_0) = \mathbf{j}(t_0) > 0$ , hence proving that  $\mathbf{a}(t) > 0$  for  $t - t_0 > 0$ , small, because of the condition  $\mathbf{b}^2 \leq \mathbf{a}\mathbf{c}$  and  $\frac{d\mathbf{c}}{dt} \leq 0$ , which means that  $t \mapsto (\mathbf{a}(t), \mathbf{c}(t))$  is constrained to the interior of the ellipse of Fig. 2.  $\square$

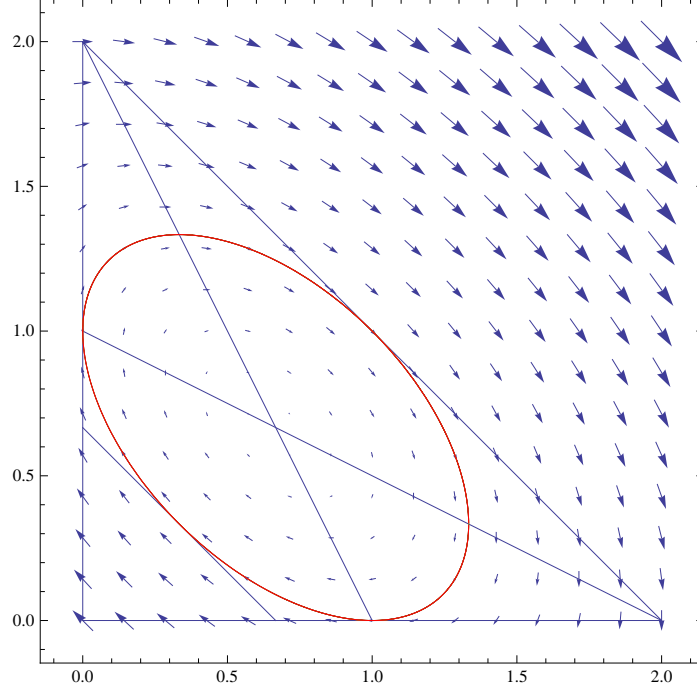


Fig. 2. Plot of the vector field associated with the ODEs  $\frac{da}{dt} = a - 2(j - c)$  and  $\frac{dc}{dt} = 2(j - a) - c$ . The coordinates are  $a/j$  (horizontal axis) and  $c/j$  (vertical axis). The two straight lines intersecting at the center of the ellipse are defined by  $2(j - a) - c = 0$  and  $a - 2j + 2c = 0$ .

### 3.3. Proof of Theorem 1.1

Let us consider the *Fisher information* functional as defined in (1.10). A computation shows that

$$-\frac{1}{2} \frac{d}{dt} \mathcal{J}_{\lambda(t)}[h(t, \cdot)] = X^\perp \cdot \mathfrak{M}_1 X - \frac{1}{2} \lambda'(t) X^\perp \cdot \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} X + Y^\perp \cdot \mathfrak{M}_2 Y$$

where  $\mathfrak{M}_0$ ,  $\mathfrak{M}_1$  and  $\mathfrak{M}_2$  are defined as before, with  $\nu = 1$ , and  $X = (\nabla_v h, \nabla_x h)$ ,  $Y = (H_{vv}, H_{xv}, M_{vv}, M_{xv})$ . We take of course  $\lambda = \lambda(t)$ . We know that

$$Y^\perp \cdot \mathfrak{M}_2 Y \geq \lambda_1(p, \lambda) |Y|^2$$

for some  $\lambda_1(p, \lambda)$  such that  $\lambda_1(p, 1/2) = \frac{1}{4} (2\kappa + 1 - \sqrt{5\kappa^2 - 4\kappa + 1}) > 0$  if  $p \in (1, 2)$ , and  $\kappa = 8(2-p)/p$ . For any  $p \in (1, 2)$ , by continuity we know that  $\lambda_1(p, \lambda) > 0$  if  $\lambda - 1/2 > 0$  is taken small enough. From  $|Y|^2 \geq \|M_{vv}\|^2$  and, by Cauchy-Schwarz,

$$\left( \int_{\mathbb{R}^d} |\nabla_v h|^2 d\mu \right)^2 \leq \int_{\mathbb{R}^d} h^2 d\mu \int_{\mathbb{R}^d} \|M_{vv}\|^2 d\mu \leq c_0 \int_{\mathbb{R}^d} \|M_{vv}\|^2 d\mu$$

where  $c_0 := 1 + (p - 1) \mathcal{E}[h_0^{2/p}]$ , we obtain

$$-\frac{1}{2} \frac{d}{dt} \mathcal{J}_{\lambda(t)}[h(t, \cdot)] \geq X^\perp \cdot \mathfrak{M}_1 X + \frac{1}{2} \lambda'(t) X^\perp \cdot \mathfrak{M}_0 X + \varepsilon X^\perp \cdot \mathfrak{M}_3 X$$

with  $\varepsilon = \lambda_1(p, \lambda) c_0^{-1} \int_{\mathbb{R}^d} |\nabla_v h|^2 d\mu$  and  $\mathfrak{M}_3 = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \otimes \text{Id}_{\mathbb{R}^d}$ . We recall that  $\mathbf{a}$  is defined by  $\mathbf{a} = e^t \int_{\mathbb{R}^d} |\nabla_v h|^2 d\mu$  is positive except for isolated values of  $t > 0$ . Our goal is to find  $\lambda(t)$  and  $\rho(t) > 1/2$  such that

$$X^\perp \cdot \mathfrak{M}_1 X - \frac{1}{2} \lambda'(t) X^\perp \cdot \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} X + \varepsilon X^\perp \cdot \mathfrak{M}_3 X \geq \rho(t) X^\perp \cdot \mathfrak{M}_0 X$$

for any  $X \in \mathbb{R}^{2d}$ .

To establish the existence of  $\rho > 1/2$  a.e., we proceed in several steps.

- If  $\mathbf{a} \geq \mathbf{a}_*$  for some constant  $\mathbf{a}_* > 0$ , then we define  $\varepsilon(t) = \nu e^{-t}$  with  $\nu = \lambda_1(p, \lambda) c_0^{-1} \mathbf{a}_*$ ,  $\lambda(t) = (1 + \varepsilon(t))/2$  and  $\rho(t) = \frac{1}{2} (1 + \nu/(\nu + 3e^t))$ . The same estimate holds on any subinterval of  $\mathbb{R}^+$ .
- If  $\mathbf{a}(t_0) = 0$  for some  $t_0 \geq 0$ , then in a neighborhood of  $(t_0)_+$ , we can solve

$$\frac{d\lambda}{dt} = \nu \varepsilon(t), \quad \lambda(t_0) = \frac{1}{2}.$$

An eigenvalue computation shows that

$$\mathfrak{M}_1 + \frac{1}{2} \nu \varepsilon \mathfrak{M}_0 + \varepsilon \mathfrak{M}_3 \geq \zeta(\varepsilon, \lambda, \nu) \mathfrak{M}_0$$

with

$$\zeta(0, \frac{1}{2}, \nu) = \frac{1}{2}, \quad \frac{\partial \zeta}{\partial \varepsilon}(0, \frac{1}{2}, \nu) = \frac{2 + \sqrt{3} - 2\nu}{3}, \quad \frac{\partial \zeta}{\partial \lambda}(0, \frac{1}{2}, \nu) = -\frac{2}{\sqrt{3}}.$$

We choose an arbitrary  $\nu \in (0, 1 + \sqrt{3}/2)$ . Since  $0 < \lambda(t) - 1/2 = o(\varepsilon(t))$  for  $t - t_0 > 0$ , small enough, this guarantees that  $\rho(t) = \zeta(\varepsilon(t), \lambda(t), \nu)$  satisfies  $\rho(t) > 1/2$  on a neighborhood of  $(t_0)_+$ .

- If  $\zeta(t_0) = 0$  for some  $t_0 > 0$ , then in a neighborhood of  $(t_0)_-$ , we proceed as above with some  $\nu < 0$ .
- If  $(t_n)_{n \in \mathbb{N}}$  is the increasing sequence of points such that  $\mathbf{a}(t_n) = 0$  and if  $\mathbf{a}(t) > 0$  for any  $t \in \mathbb{R}^+$  such that  $t \neq t_n$  for any  $n \in \mathbb{N}$ , we can choose a constant  $\mathbf{a}_*$ , small enough, on any interval  $(t_n, t_{n+1})$  and glue the above solutions to obtain a function  $\rho(t) > 1/2$  on  $(0, t_0)$  and  $\cup_{n \in \mathbb{N}} (t_n, t_{n+1})$ . It is an open question to decide if there is an increasing sequence, finite or infinite, of times  $t_n$  such that  $\mathbf{a}(t_n) = 0$ , or if  $\mathbf{a}(t)$  is positive for any  $t > 0$ . We can of course impose that  $\mathbf{a}(t_0) = 0$  at  $t_0 = 0$  by taking an initial datum  $h_0$  which does not depend on  $v$ . If such a sequence  $(t_n)_{n \in \mathbb{N}}$  exists, then we know that  $\lambda(t_n) = 1/2$  so that we have the remarkable decay estimate

$$\mathcal{J}_{\frac{1}{2}}[h(t_{n+1}, \cdot)] \leq \mathcal{J}_{\frac{1}{2}}[h(t_n, \cdot)] e^{-2 \int_{t_n}^{t_{n+1}} \rho(s) ds} < \mathcal{J}_{\frac{1}{2}}[h(t_n, \cdot)] e^{-(t_{n+1} - t_n)}$$

for any  $p \in (1, 2)$ . As far as  $\mathbf{a}$  is concerned, we expect that it has some oscillatory behaviour as indicated by the vector field in Fig. 2, but since terms involving  $Y$  are neglected, this is so far formal. In any case, we can choose  $\lambda(t)$  such that  $\lim_{t \rightarrow +\infty} \lambda(t) = 1/2$ . This concludes the proof of Theorem 1.1.  $\square$



### 3.4. Concluding remarks

Even if the global rate cannot be improved because it is determined by the large time asymptotics, at any finite time the instantaneous rate of decay is strictly higher in the case of the diffusions studied in Sections 2.4-2.5, or at least higher at almost any time in the case of the kinetic equation, according to Theorem 1.1.

As  $t \rightarrow +\infty$ , Theorem 1.1 provides us with an improved estimate of the leading order term. The exponential decay rate cannot be improved as shown by (3.4), but we prove that there is a constant less than 1 to be taken into account. This observation is reminiscent of what happens for nonlinear diffusions of porous medium or fast diffusion type, which goes as follows. When looking at the relative entropy with respect to the *best matching* (in the sense of relative entropy) profiles, it turns out that there is a delay  $\tau$  compared to the relative entropy with respect to a fixed Barenblatt profile. As a result, we obtain a multiplicative factor  $e^{-\tau}$  corresponding to an improved estimate in an asymptotic expansion as  $t \rightarrow +\infty$  <sup>34</sup>. We have a similar property when we study the large time behavior of the solutions of (1.6) using a  $\varphi_p$ -entropy for any given  $p \in (1, 2)$ .

The key estimate of Theorem 1.1 asserts that

$$\frac{d}{dt} \mathcal{J}_{\lambda(t)}[h(t, \cdot)] \leq -2\rho(t) \mathcal{J}_{\lambda(t)}[h(t, \cdot)] \leq -\mathcal{J}_{\lambda(t)}[h(t, \cdot)]$$

where the last inequality is strict for almost any value of  $t \geq 0$  (unless  $h$  is a stationary solution). Now, let us consider the large time asymptotics and define

$$\tau := \lim_{t \rightarrow +\infty} \left( 2 \int_0^t \rho(s) ds - t \right).$$

We cannot expect that  $\tau = +\infty$  for any initial datum but at least show that  $\tau$  is positive (unless  $h$  is a stationary solution), so that for large values of  $t$  we have

$$\mathcal{J}_{1/2}[h(t, \cdot)] \lesssim e^{-\tau} \mathcal{J}_{1/2}[h_0] e^{-t}. \tag{3.5}$$

For instance, in case of (3.4), one can prove that  $\rho(t) - 1/2$  is of the order of  $e^{-t}$  and  $\tau$  is finite. With  $e^{-\tau} < 1$ , (3.5) is anyway a strict improvement of the usual estimate as  $t \rightarrow +\infty$ .

The improvement of Theorem 1.1 is obtained only for *almost any* time: according to Lemma 3.2, the optimal decay rate could eventually be realized at an increasing sequence of times  $t_n \nearrow +\infty$ , but the solution will then deviate and temporarily regain a faster decay rate. Qualitatively, this comes from the oscillations in the phase space corresponding to the ODE associated with the vector field shown in Fig. 2. Such a pattern is consistent with what is known of the rates measured by *hypocoercive methods* in kinetic equations.

### Acknowledgments

This work has been partially supported by the Projects Kibord and EFI (J.D.) of the French National Research Agency (ANR). The first author (J.D.) thanks J.-P. Bartier and B. Nazaret for fruitful discussions on, respectively, the tensorization

properties of the  $\varphi$ -entropies and various computations based on the Bakry-Emery method, which took place at the occasion of a course taught by C. Mouhot on hypocoercivity methods. Both authors thank an anonymous referee for interesting comments and suggestion which lead to improvements of this article.

© 2018 by the authors. This paper may be reproduced, in its entirety, for non-commercial purposes.

## References

1. F. ACHLEITNER, A. ARNOLD, AND E. A. CARLEN, *On linear hypocoercive BGK models*, in From particle systems to partial differential equations. III, vol. 162 of Springer Proc. Math. Stat., Springer, [Cham], 2016, pp. 1–37.
2. F. ACHLEITNER, A. ARNOLD, AND D. STÜRZER, *Large-time behavior in non-symmetric Fokker-Planck equations*, Riv. Math. Univ. Parma (N.S.), 6 (2015), pp. 1–68.
3. C. ANÉ, S. BLACHÈRE, D. CHAFAÏ, P. FOUGÈRES, I. GENTIL, F. MALRIEU, C. ROBERTO, AND G. SCHEFFER, *Sur les inégalités de Sobolev logarithmiques*, vol. 10 of Panoramas et Synthèses [Panoramas and Syntheses], Société Mathématique de France, Paris, 2000. With a preface by Dominique Bakry and Michel Ledoux.
4. A. ARNOLD, J.-P. BARTIER, AND J. DOLBEAULT, *Interpolation between logarithmic Sobolev and Poincaré inequalities*, Communications in Mathematical Sciences, 5 (2007), pp. 971–979.
5. A. ARNOLD AND J. DOLBEAULT, *Refined convex Sobolev inequalities*, J. Funct. Anal., 225 (2005), pp. 337–351.
6. A. ARNOLD AND J. ERB, *Sharp entropy decay for hypocoercive and non-symmetric Fokker-Planck equations with linear drift*, [arXiv: 1409.5425](https://arxiv.org/abs/1409.5425), (2014).
7. A. ARNOLD, P. MARKOWICH, G. TOSCANI, AND A. UNTERREITER, *On convex Sobolev inequalities and the rate of convergence to equilibrium for Fokker-Planck type equations*, Comm. Partial Differential Equations, 26 (2001), pp. 43–100.
8. D. BAKRY, P. CATTIAUX, AND A. GUILLIN, *Rate of convergence for ergodic continuous Markov processes: Lyapunov versus Poincaré*, J. Funct. Anal., 254 (2008), pp. 727–759.
9. D. BAKRY AND M. ÉMERY, *Diffusions hypercontractives*, in Séminaire de probabilités, XIX, 1983/84, vol. 1123 of Lecture Notes in Math., Springer, Berlin, 1985, pp. 177–206.
10. F. BARTHE, P. CATTIAUX, AND C. ROBERTO, *Interpolated inequalities between exponential and Gaussian, Orlicz hypercontractivity and isoperimetry*, Rev. Mat. Iberoam., 22 (2006), pp. 993–1067.
11. F. BARTHE AND C. ROBERTO, *Sobolev inequalities for probability measures on the real line*, Studia Math., 159 (2003), pp. 481–497. Dedicated to Professor Aleksander Pełczyński on the occasion of his 70th birthday (Polish).
12. J.-P. BARTIER AND J. DOLBEAULT, *Convex Sobolev inequalities and spectral gap*, C. R. Math. Acad. Sci. Paris, 342 (2006), pp. 307–312.
13. J.-P. BARTIER, J. DOLBEAULT, R. ILNER, AND M. KOWALCZYK, *A qualitative study of linear drift-diffusion equations with time-dependent or degenerate coefficients*, Math. Models Methods Appl. Sci., 17 (2007), pp. 327–362.
14. F. BAUDOIN, *Bakry-Émery meet Villani*, J. Funct. Anal., 273 (2017), pp. 2275–2291.
15. W. BECKNER, *A generalized Poincaré inequality for Gaussian measures*, Proc. Amer. Math. Soc., 105 (1989), pp. 397–400.
16. A. BLANCHET, J. DOLBEAULT, AND M. KOWALCZYK, *Stochastic Stokes’ drift, homogenized functional inequalities, and large time behavior of brownian ratchets*, SIAM Journal on Mathematical Analysis, 41 (2009), pp. 46–76.

17. T. BODINEAU, J. LEBOWITZ, C. MOUHOT, AND C. VILLANI, *Lyapunov functionals for boundary-driven nonlinear drift-diffusion equations*, *Nonlinearity*, 27 (2014), pp. 2111–2132.
18. F. BOLLEY AND I. GENTIL, *Phi-entropy inequalities for diffusion semigroups*, *J. Math. Pures Appl.* (9), 93 (2010), pp. 449–473.
19. F. BOUCHUT, *Existence de solutions régulières globales pour le système de Vlasov-Poisson-Fokker-Planck en dimension trois*, *C. R. Acad. Sci. Paris Sér. I Math.*, 313 (1991), pp. 243–248.
20. ———, *Existence and uniqueness of a global smooth solution for the Vlasov-Poisson-Fokker-Planck system in three dimensions*, *J. Funct. Anal.*, 111 (1993), pp. 239–258.
21. E. BOUIN, J. DOLBEAULT, S. MISCHLER, C. MOUHOT, AND C. SCHMEISER, *Hypocoercivity without confinement*, [hal-01575501](https://hal.archives-ouvertes.fr/hal-01575501) & [arXiv:1708.06180](https://arxiv.org/abs/1708.06180), (2017).
22. M. J. CÁCERES, J. A. CARRILLO, AND J. DOLBEAULT, *Nonlinear stability in  $L^p$  for a confined system of charged particles*, *SIAM J. Math. Anal.*, 34 (2002), pp. 478–494.
23. J. A. CARRILLO, J. DOLBEAULT, P. A. MARKOWICH, AND C. SPARBER, *On the long-time behavior of the quantum Fokker-Planck equation*, *Monatsh. Math.*, 141 (2004), pp. 237–257.
24. D. CHAFAÏ, *Entropies, convexity, and functional inequalities: on  $\Phi$ -entropies and  $\Phi$ -Sobolev inequalities*, *J. Math. Kyoto Univ.*, 44 (2004), pp. 325–363.
25. I. CSISZÁR, *Information-type measures of difference of probability distributions and indirect observations*, *Studia Sci. Math. Hungar.*, 2 (1967), pp. 299–318.
26. I. CSISZAR AND J. KÖRNER, *Information theory: coding theorems for discrete memoryless systems*, Cambridge University Press, 2011.
27. J. DOLBEAULT, M. J. ESTEBAN, M. KOWALCZYK, AND M. LOSS, *Improved interpolation inequalities on the sphere*, *Discrete and Continuous Dynamical Systems Series S (DCDS-S)*, 7 (2014), pp. 695–724.
28. J. DOLBEAULT AND M. KOWALCZYK, *Uniqueness and rigidity in nonlinear elliptic equations, interpolation inequalities, and spectral estimates*, *Annales de la faculté des sciences de Toulouse Sér. 6*, 26 (2017), pp. 949–977.
29. J. DOLBEAULT, C. MOUHOT, AND C. SCHMEISER, *Hypocoercivity for kinetic equations with linear relaxation terms*, *Comptes Rendus Mathématique*, 347 (2009), pp. 511 – 516.
30. ———, *Hypocoercivity for linear kinetic equations conserving mass*, *Trans. Amer. Math. Soc.*, 367 (2015), pp. 3807–3828.
31. J. DOLBEAULT, B. NAZARET, AND G. SAVARÉ, *A new class of transport distances between measures*, *Calculus of Variations and PDE*, (2007).
32. ———, *On the Bakry-Emery criterion for linear diffusions and weighted porous media equations*, *Commun. Math. Sci.*, 6 (2008), pp. 477–494.
33. ———, *From Poincaré to logarithmic Sobolev inequalities: a gradient flow approach*, *SIAM J. Math. Anal.*, 44 (2012), pp. 3186–3216.
34. J. DOLBEAULT AND G. TOSCANI, *Nonlinear diffusions: Extremal properties of Barenblatt profiles, best matching and delays*, *Nonlinear Analysis: Theory, Methods & Applications*, 138 (2016), pp. 31–43.
35. ———, *Stability results for logarithmic Sobolev and Gagliardo-Nirenberg inequalities*, *Int. Math. Res. Not. IMRN*, (2016), pp. 473–498.
36. J.-P. ECKMANN AND M. HAIRER, *Spectral properties of hypoelliptic operators*, *Comm. Math. Phys.*, 235 (2003), pp. 233–253.
37. J. EVANS, *Hypocoercivity in Phi-entropy for the Linear Boltzmann Equation on the Torus*, [arXiv:1702.04168](https://arxiv.org/abs/1702.04168), (2017).
38. P. FEDERBUSH, *Partially alternate derivation of a result of Nelson*, *J. Mathematical*

- Phys., 10 (1969), pp. 50–52.
39. U. GIANAZZA, G. SAVARÉ, AND G. TOSCANI, *The Wasserstein gradient flow of the Fisher information and the quantum drift-diffusion equation*, Arch. Ration. Mech. Anal., 194 (2009), pp. 133–220.
  40. P. GRISVARD, *Elliptic problems in nonsmooth domains*, vol. 24 of Monographs and Studies in Mathematics, Pitman (Advanced Publishing Program), Boston, MA, 1985.
  41. L. GROSS, *Logarithmic Sobolev inequalities*, Amer. J. Math., 97 (1975), pp. 1061–1083.
  42. G. H. HARDY, J. E. LITTLEWOOD, AND G. PÓLYA, *Inequalities*, Cambridge Mathematical Library, Cambridge University Press, Cambridge, 1988. Reprint of the 1952 edition.
  43. B. HELFFER AND F. NIER, *Hypoelliptic estimates and spectral theory for Fokker-Planck operators and Witten Laplacians*, vol. 1862 of Lecture Notes in Mathematics, Springer-Verlag, Berlin, 2005.
  44. F. HÉRAU, *Hypocoercivity and exponential time decay for the linear inhomogeneous relaxation Boltzmann equation*, Asymptot. Anal., 46 (2006), pp. 349–359.
  45. ———, *Short and long time behavior of the Fokker-Planck equation in a confining potential and applications*, J. Funct. Anal., 244 (2007), pp. 95–118.
  46. F. HÉRAU AND F. NIER, *Isotropic hypoellipticity and trend to equilibrium for the Fokker-Planck equation with a high-degree potential*, Arch. Ration. Mech. Anal., 171 (2004), pp. 151–218.
  47. R. HOLLEY AND D. STROOCK, *Logarithmic Sobolev inequalities and stochastic Ising models*, J. Statist. Phys., 46 (1987), pp. 1159–1194.
  48. L. HÖRMANDER, *Hypoelliptic second order differential equations*, Acta Math., 119 (1967), pp. 147–171.
  49. A. IACOBUCCI, S. OLLA, AND G. STOLTZ, *Convergence rates for nonequilibrium Langevin dynamics*, Annales mathématiques du Québec, (2017).
  50. A. IL'IN AND R. KHAS' MINSKII, *On equations of Brownian motion*, Theory of Probability & Its Applications, 9 (1964), pp. 421–444.
  51. J. H. KEMPERMAN, *On the optimum rate of transmitting information*, in Probability and information theory, Springer, 1969, pp. 126–169.
  52. A. KOLMOGOROFF, *Zufällige Bewegungen (zur Theorie der Brownschen Bewegung)*, Ann. of Math. (2), 35 (1934), pp. 116–117.
  53. S. KULLBACK, *On the convergence of discrimination information*, IEEE Trans. Information Theory, IT-14 (1968), pp. 765–766.
  54. R. LATAŁA AND K. OLESZKIEWICZ, *Between Sobolev and Poincaré*, in Geometric aspects of functional analysis, vol. 1745 of Lecture Notes in Math., Springer, Berlin, 2000, pp. 147–168.
  55. P. W. Y. LEE, *Sharp Harnack inequalities for a family of hypoelliptic diffusions*, Journal of Mathematical Physics, 58 (2017), p. 031501.
  56. P. MONMARCHÉ, *On  $\mathcal{H}^1$  and entropic convergence for contractive PDMP*, Electron. J. Probab., 20 (2015), pp. 1–30, Paper No. 128.
  57. ———, *A note on Fisher Information hypocoercive decay for the linear Boltzmann equation*, [arXiv: 1703.10504](https://arxiv.org/abs/1703.10504), (2017).
  58. ———, *Generalized  $\Gamma$  calculus and application to interacting particles on a graph*, Potential Analysis, [Doi: 10.1007/s11118-018-9689-3](https://doi.org/10.1007/s11118-018-9689-3), (2018).
  59. C. MOUHOT AND L. NEUMANN, *Quantitative perturbative study of convergence to equilibrium for collisional kinetic models in the torus*, Nonlinearity, 19 (2006), pp. 969–998.
  60. B. MUCKENHOUT, *Hardy's inequality with weights*, Studia Math., 44 (1972), pp. 31–38. Collection of articles honoring the completion by Antoni Zygmund of 50 years of

scientific activity, I.

61. J. NASH, *Continuity of solutions of parabolic and elliptic equations*, Amer. J. Math., 80 (1958), pp. 931–954.
62. G. A. PAVLIOTIS, *Stochastic processes and applications*, vol. 60 of Texts in Applied Mathematics, Springer, New York, 2014. Diffusion processes, the Fokker-Planck and Langevin equations.
63. M. S. PINSKER, *Information and information stability of random variables and processes*, Translated and edited by Amiel Feinstein, Holden-Day Inc., San Francisco, Calif., 1964.
64. H. RISKEN, *The Fokker-Planck equation*, vol. 18 of Springer Series in Synergetics, Springer-Verlag, Berlin, 1984. Methods of solution and applications.
65. A. J. STAM, *Some inequalities satisfied by the quantities of information of Fisher and Shannon*, Information and Control, 2 (1959), pp. 101–112.
66. G. TOSCANI, *Rényi entropies and nonlinear diffusion equations*, Acta Appl. Math., 132 (2014), pp. 595–604.
67. A. UNTERREITER, A. ARNOLD, P. MARKOWICH, AND G. TOSCANI, *On generalized Csiszár-Kullback inequalities*, Monatsh. Math., 131 (2000), pp. 235–253.
68. H. D. VICTORY, JR. AND B. P. O'DWYER, *On classical solutions of Vlasov-Poisson Fokker-Planck systems*, Indiana Univ. Math. J., 39 (1990), pp. 105–156.
69. C. VILLANI, *Hypocoercive diffusion operators*, in International Congress of Mathematicians. Vol. III, Eur. Math. Soc., Zürich, 2006, pp. 473–498.
70. ———, *Entropy Production and Convergence to Equilibrium*, Springer Berlin Heidelberg, Berlin, Heidelberg, 2008, pp. 1–70.
71. ———, *Hypocoercivity*, Mem. Amer. Math. Soc., 202 (2009), pp. iv+141.
72. F. B. WEISSLER, *Logarithmic Sobolev inequalities for the heat-diffusion semigroup*, Trans. Amer. Math. Soc., 237 (1978), pp. 255–269.