Inequalities for generalized entropy and optimal transportation

Dario Cordero-Erausquin, Wilfrid Gangbo,* Christian Houdré†

April 10, 2003

Abstract

A new concept of *Fisher-information* is introduced through a cost function. That concept is used to obtain extensions and variants of transport and logarithmic Sobolev inequalities for general entropy functionals and transport costs. Our proofs rely on optimal mass transport from the Monge-Kantorovich theory. They express the convexity of entropy functionals with respect to suitably chosen paths on the set of probability measures.

Contents

1 Introduction

The purpose of these notes is to further explore the connections between optimal mass transport and Sobolev type functional inequalities. More precisely, the aim of the paper is two-fold.

*WG gratefully acknowledges the support provided by NSF grants DMS-99-70520, DMS-00-74037 and DMS-02-00267.

†CH gratefully acknowledges the support provided by the NSF grant DMS-98-03239.

On one hand we introduce generalized notions of entropy and of Fisher information, already present in previous notes of Gangbo and Houdré [15]. This new concept of *Fisher-information* occurs naturally when one studies quasilinear parabolic–elliptic equations (1.1), an important class of equations in partial differential equations (see [2], [21], [22] and, more recently, in [1]). On the other hand, we show that the simple approach given by Cordero-Erausquin in [13] for proving logarithmic Sobolev, transport and interpolation inequalities extends to this new setting. Our approach is purely analytical and uses the Monge–Kantorovich theory.

Since its study by Gross [19], the logarithmic Sobolev inequality (in various forms) has proved useful in several fields of mathematics. In PDE, it provides a control of the entropy production for evolutive dissipative systems. In probability theory, it is a tool to obtain concentration of measure phenomena or to study smoothness properties of Markov processes. In combinatorial theory, it gives estimates on mixing time of randomized algorithms. Since our framework is motivated by classes of non linear PDEs, the version of the logarithmic Sobolev inequality we have in mind is the one linking entropy and Fisher information. Indeed, the entropy appears naturally as an energy functional in the study of the Fokker–Planck equation, and in turn the logarithmic Sobolev inequality is used to study the asymptotic behavior of its solutions. The Fokker–Planck equation is part of an important class of equations modelling dissipative systems, the so–called quasilinear parabolic–elliptic equations which are described via

$$\begin{aligned} \frac{\partial \rho}{\partial t}(t,x) + \operatorname{div}\Big[\rho(t,x)\mathbf{U}_\rho(t,x)\Big] = 0, \qquad (t,x) \in [0,+\infty) \times \Omega \\ \mathbf{U}_\rho := -\nabla c^*\Big(\nabla(F' \circ \rho + V)\Big), \qquad (t,x) \in [0,+\infty) \times \Omega. \end{aligned} \tag{1.1}$$

Here, $\Omega \subset \mathbf{R}^d$ is an open set, $c \in C^1(\mathbf{R}^d)$ is strictly convex with Legendre transform c^*, $F \in C^1(\mathbf{R}^+)$, and the unknown is $t \to \rho(t,\cdot) \in W^{1,1}(\Omega)$. When Ω is bounded, we impose in addition that the boundary condition

$$\mathbf{U}_\rho \cdot \mathbf{n} = 0,$$

holds for $(t,x) \in [0,+\infty) \times \partial\Omega$, where $\mathbf{n}$ is the outward unit normal to $\partial\Omega$.

Quasilinear parabolic–elliptic equations have been studied by several authors, and pioneering existence results were obtained by Alt and Luckhaus [2]. The functional inequalities we obtain in the present paper should prove useful for studying the asymptotic behavior of these systems.

In order to introduce a criterion which is a way of assessing how two probability measures are different from one another, we introduce a *cost function* $c : \mathbf{R}^d \longrightarrow \mathbf{R}^+ := [0,+\infty)$, so that if x and y are two points in $\mathbf{R}^d$, then $c(x-y)$ represents the cost of transporting a unit mass from x to y. Given two probability measures μ and ν on $\mathbf{R}^d$, the minimum cost for transporting μ onto ν is then

$$W_c(\mu,\nu) := \inf_{\gamma \in \Gamma(\mu,\nu)} \int_{\mathbf{R}^d \times \mathbf{R}^d} c(x-y)\, d\gamma(x,y), \tag{1.2}$$

where $\Gamma(\mu,\nu)$ is the set of Borel probability measures with marginals μ and ν, respectively:

$$\mu(B) = \gamma(B \times \mathbf{R}^d), \quad \nu(B) = \gamma(\mathbf{R}^d \times B)$$

for every Borel set $B \subset \mathbf{R}^d$. When μ and ν are absolutely continuous with respect to the Lebesgue measure, i.e., $d\mu = \rho_0\, dx$, and $d\nu = \rho_1\, dx$, we write $\Gamma(\rho_0, \rho_1)$ instead of $\Gamma(\mu, \nu)$. Next, a Borel map $T : \mathbf{R}^d \longrightarrow \mathbf{R}^d$ is said to *push* μ *forward to* ν (or to *transport* μ onto ν) if for every Borel set $B \subset \mathbf{R}^d$

$$\mu(T^{-1}(B)) = \nu(B).$$

In other words, ν is the image of μ by T, and this is written as $\nu := T_{\#}\mu$ (or using a probabilistic notation $\nu := \mu \circ T^{-1}$). Again, when $d\mu = \rho_0\, dx$, and $d\nu = \rho_1\, dx$, $T_{\#}\rho_0 = \rho_1$ is used instead of $T_{\#}\mu = \nu$. A map T pushing μ forward to ν is said to be *c-optimal* if

$$W_c(\mu, \nu) = \int_{\mathbf{R}^d} c(x - T(x))d\mu(x), \tag{1.3}$$

and in this case

$$\int_{\mathbf{R}^d} c(x - T(x))d\mu = \inf_S \int_{\mathbf{R}^d} c(x - S(x))d\mu,$$

where the infimum is taken over all Borel maps $S : \mathbf{R}^d \longrightarrow \mathbf{R}^d$ pushing μ forward to ν. Let us now make some standing assumptions on the cost c which will be used throughout the text:

(H1) $c : \mathbf{R}^d \longrightarrow [0, \infty)$ *is strictly convex, even, and of class* C^1.

Imposing that "c is even" is not important in the present work. This property of c is used only to state nice symmetric results. As a consequence, the Legendre transform c^* is even and ∇c^* is odd. We also conveniently impose that $c(0) = 0$ so that $c^*(0) = 0$. This allows us to avoid carrying the term $c^*(0)$ as an extra additive constant in inequalities such as (1.28).

(H2) $c(0) = 0$.

(H3) $\displaystyle\lim_{|z| \to \infty} \frac{c(z)}{|z|} = +\infty.$

Observe that (H3) is unnecessary when transporting densities with bounded supports.

Cost functions satisfying (H1–H3) include all the radial costs $c(z) = \ell(|z|)$ of class C^1, growing faster than linearly, and such that $\ell(t) \geq \ell(0) = 0$ with ℓ strictly convex. Homogeneous costs given, for instance, by $c(z) = \|z\|_p^p := \sum_{i=1}^d |z_i|^p$, $p > 1$, also satisfy these conditions.

Since by definition c^*, the Legendre transform of c, is given by

$$c^*(y) = \sup_{z \in \mathbf{R}^d} \{y \cdot z - c(z)\}, \tag{1.4}$$

and c is convex, Young's inequality

$$y \cdot z \leq c^*(y) + c(z), \tag{1.5}$$

holds for all $y, z \in \mathbf{R}^d$ and is saturated when $z = \nabla c^*(y)$:

$$y \cdot \nabla c^*(y) = c^*(y) + c(\nabla c^*(y)), \tag{1.6}$$

for all $y \in \mathbf{R}^d$. Moreover, the convexity of c^* implies that for all $y \in \mathbf{R}^d$,

$$y \cdot \nabla c^*(y) \geq c^*(y) - c^*(0) = c^*(y) \geq 0, \tag{1.7}$$

since $c^*(0) = c(0) = 0$.

When the cost is the quadratic one, i.e. $c(z) := |z|^2/2$, a result of Brenier characterizes the optimal map T in (1.3) as the gradient of a convex function [8, 9]. For general strictly convex costs, Caffarelli [10] as well as Gangbo and McCann [16, 17] independently proved that the c–optimal map is unique and takes the form

$$T(x) = x - \nabla c^*(\nabla\theta(x)),$$

where θ is a c–concave function. We refer to [17] for the definition of c–concavity and precise statements. Now, let $\mathcal{P}^a$ be the set of Borel probability densities, i.e.,

$$\mathcal{P}^a := \left\{\rho \in L^1(\mathbf{R}^d) : \rho \geq 0 \text{ and } \int_{\mathbf{R}^d} \rho\, dx = 1\right\}.$$

Let $V : \mathbf{R}^d \longrightarrow \mathbf{R}$ be such that

$$V(b) - V(a) \geq \nabla V(a) \cdot (b - a) + \alpha_0 c(a - b), \tag{1.8}$$

for some $\alpha_0 \in \mathbf{R}$ and all $a, b \in \mathbf{R}^d$. When $c(z) := |z|^2/2$ and V is twice differentiable, (1.8) is equivalent to $\text{Hess}V \geq \alpha_0 I_d$, where I_d stands for the $d \times d$ identity matrix. Let also $F : \mathbf{R}^+ \longrightarrow \mathbf{R}$ be strictly convex.

We now introduce the so-called, *free energy functional*

$$H_V^F(\rho) := \int_{\mathbf{R}^d} (F(\rho) + \rho V)dx,$$

which is the sum of the internal energy and the potential energy given respectively by

$$H^F(\rho) := \int_{\mathbf{R}^d} F(\rho)dx, \tag{1.9}$$

$$H_V(\rho) := \int_{\mathbf{R}^d} \rho V dx. \tag{1.10}$$

Eventually, we will work with the triple (F, V, ρ_∞) where $\rho_\infty \in \mathcal{P}^a$ is uniquely determined by

$$F'(\rho_\infty) + V = 0, \tag{1.11}$$

on its support. In light of (1.11), by the strict convexity of F, and unless $\rho \equiv \rho_\infty$, we have

$$F(\rho) + \rho V > F(\rho_\infty) + \rho_\infty V = -F^*(-V), \tag{1.12}$$

on the support of ρ_∞ (since F is only defined on $\mathbf{R}^+$, and in order to properly define F^* when needed, we set throughout $F(t) = +\infty$, for $t < 0$). Thus, $H_V^F(\rho)$ is well defined, although possibly infinite, provided that $F^*(-V)$ is integrable. Furthermore, ρ_∞ is the unique minimizer of H_V^F on that set. In the classical case, when

$$c(z) = |z|^2/2, \quad F(t) = t\log t - t, \quad F(0) = 0, \tag{1.13}$$

and, in view of (1.11),

$$\rho_\infty = e^{-V},$$

we simply denote the *free energy functional* by H and we have $H(\rho) = \int_{\mathbf{R}^d} \rho \log(\rho/\rho_\infty)dx - 1$. As already mentioned, (1.8) then reads as

$$\mathrm{Hess} V \geq \alpha_0 I_d,$$

when V is twice differentiable. To use a terminology similar to the classical one,

$$H_V^F(\rho) - H_V^F(\rho_\infty), \tag{1.14}$$

is called the *relative entropy* of ρ with respect to ρ_∞. Note that under the condition (1.13), the equation (1.1) is just the Fokker–Planck equation

$$\frac{\partial \rho}{\partial t} = \mathrm{div}(\nabla\rho + \rho\nabla V) = \mathrm{div}(\rho\nabla\left(\log\rho + V\right)). \tag{1.15}$$

with ground state $\rho_\infty = e^{-V}$. When its solution ρ is smooth enough, it satisfies

$$\frac{d}{dt}\Big(H(\rho)\Big) = -I(\rho \mid \rho_\infty), \tag{1.16}$$

where $I(\rho \mid \rho_\infty)$ is the relative Fisher information of ρ with respect to ρ_∞, defined by

$$I(\rho \mid \rho_\infty) = \int_{\mathbf{R}^d} |\nabla(\log\rho + V)|^2 \rho dx. \tag{1.17}$$

Similarly, if ρ is a smooth solution of the quasilinear parabolic–elliptic systems (1.1) then it is easily verified that

$$\frac{d}{dt}\Big(H_V^F(\rho)\Big) = -I_{c^*}(\rho \mid \rho_\infty), \tag{1.18}$$

where

$$I_{c^*}(\rho \mid \rho_\infty) := \int_{\mathbf{R}^d} \nabla(F' \circ \rho + V) \cdot \nabla c^*\Big(\nabla(F' \circ \rho + V)\Big)\rho dx. \tag{1.19}$$

Because of the analogy between (1.15) and (1.16) on one hand, and (1.1) and (1.18) on the other hand, we call $I_{c^*}(\rho \mid \rho_\infty)$ the *generalized relative Fisher information* of ρ with respect to ρ_∞, *measured against the cost* c^*. Note that by (1.7), $I_{c^*}(\rho \mid \rho_\infty)$ is nonnegative (and possibly infinite).

The idea of finding paths connecting elements of $\mathcal{P}^a$ has, in the present context, its origin in the work of McCann [25]. It will provide a nice interpretation of the inequalities we are interested in. When c is homogeneous of degree $p \geq 1$, these paths will turn out to be geodesics for the metric $W_c^{1/p}$. To avoid technical difficulties when studying properties of H_V^F, we keep our focus on path connections for $\rho_0, \rho_1 \in W^{1,\infty}(\Omega)$, ($\Omega$ open bounded and convex), such that $\inf_\Omega \rho_0 > 0$. If T is the c-optimal map that pushes ρ_0 forward to ρ_1, we define the interpolant measures

$$\mu_t := \Big((1-t)\mathbf{id} + tT\Big)_{\#}(\rho_0 dx). \tag{1.20}$$

For the functional H^F in (1.9) to have interesting properties along $t \to \mu_t$ we assume that
(H4) $F \in C^2(0, +\infty) \cap C([0, +\infty))$, and $F(0) = 0$.
The assumption $F(0) = 0$ is made to ensure that

$$\int_{\mathbf{R}^d} F(\rho) dx = \int_{\{\rho>0\}} F(\rho) dx.$$

(H5) $t \longrightarrow t^d F(t^{-d})$ *is convex and nonincreasing on* $(0, +\infty)$,
In the classical context, (H5) appears in McCann [25] as a condition that ensures "displacement convexity" of H^F and is used to prove (1.21) below.

It is important to have in mind the following examples of functions F, related to the so-called *Rényi entropy functionals* in information theory:

$$F_m(\rho) := \frac{1}{m-1}(\rho^m - \rho).$$

In that case, (H4-H5) is satisfied if and only if $m \geq 1 - 1/n$. The case $m = 1$, defined in the limit as $F_1(t) = t \log t$, was already considered above. For a given $m \geq 1 - 1/n$, the corresponding ρ_∞ in (1.11) is given by

$$\rho_\infty(x) := \Big(\sigma + \frac{1-m}{m} V(x)\Big)_+^{-\frac{1}{1-m}}.$$

In the classical case ($V(x) = |x|^2/2$) this function is sometimes called the *Barenblatt profile.* Note that it is compactly supported when $m > 1$ and positive of polynomial decay when $m < 1$.

It shown in [1] and [26] that when c satisfies (H1–H3), and F satisfies (H4-H5) then

$$H^F(\rho_1) - H^F(\rho_0) \geq \int_{\mathbf{R}^d} (T - \mathbf{id}) \cdot \nabla(A \circ \rho_0)\, dx = \int_{\mathbf{R}^d} (T - \mathbf{id}).\nabla(F' \circ \rho_0)\, \rho_0\, dx, \tag{1.21}$$

where $A(0) = 0$, and $A(t) := tF'(t) - F(t)$, for $t > 0$. A simple proof of (1.21) under different boundary conditions will be given in Appendix A, Proposition 5.1. Cordero-Erausquin [14] has noticed that (with stronger regularity assumptions on the cost c) the μ_t's (1.20) are absolutely continuous with respect to Lebesgue measure. Hence, it is natural to define the displacement interpolant densities

$$\rho_t := [(1-t)\mathbf{id} + tT]_\# \rho_0. \tag{1.22}$$

We emphasize that in the proofs of the current work, we will make no use of the fact the μ_t's are absolutely continuous with respect to the Lebesgue measure. We rather follow the "direct" approach of [13]. We have mentioned this property of the μ_t's simply to motivate some of our definitions. For instance, formally (1.21) is equivalent to the familiar inequality

$$H^F(\rho_1) - H^F(\rho_0) \geq \Big[\frac{d}{dt} H^F(\rho_t)\Big]_{t=0},$$

expressing that $t \to H^F(\rho_t)$ is convex. We say that H^F is W_c–convex whenever (1.21) holds (when $c(z) = |z|^2/2$, this is McCann's displacement convexity). In Lemma 2.3 when $V \in C^1(\mathbf{R}^d)$, we show that the pointwise inequality (1.8) is satisfied if and only if

$$H_V(\rho_1) - H_V(\rho_0) \geq \int_{\mathbf{R}^d} \nabla V \cdot (T - \mathbf{id})\rho_0 \, dx + \alpha_0 W_c(\rho_0, \rho_1). \tag{1.23}$$

Formally (1.23) says that the following second order Taylor expansion holds

$$H_V(\rho_1) - H_V(\rho_0) \geq \Big[\frac{d}{dt} H_V(\rho_t)\Big]_{t=0} + \alpha_0 W_c(\rho_0, \rho_1).$$

We then say that H_V is W_c–semiconvex whenever (1.23) holds. Combining (1.21) and (1.23) we derive Theorem 2.4 – the central ingredient of this work – which generalizes an inequality of [13] into:

$$H_V^F(\rho_1) - H_V^F(\rho_0) \geq \alpha_0 W_c(\rho_0, \rho_1) + \int_{\mathbf{R}^d} (T - \mathbf{id}) \cdot \nabla(F' \circ \rho_0 - F' \circ \rho_\infty)\rho_0 \, dx. \tag{1.24}$$

In other words, H_V^F is W_c–semiconvex, and is W_c–convex (resp. uniformly W_c–convex) in the particular case $\alpha_0 \geq 0$ (resp. $\alpha_0 > 0$), since (1.24) expresses that

$$H_V^F(\rho_1) - H_V^F(\rho_0) \geq \Big[\frac{d}{dt} H_V^F(\rho_t)\Big]_{t=0} + \alpha_0 W_c(\rho_0, \rho_1).$$

In Section 3, we obtain the generalized transport inequality (1.25), and the generalized "logarithmic" (there is more logarithm...) Sobolev inequality (1.28) as direct consequences of (1.24). Indeed, assume first that the functions ρ and ρ_∞ have bounded supports. Whenever $\alpha_0 > 0$, by substituting the cost function c by $\alpha_0 c$ if necessary, we may assume without loss of generality that $\alpha_0 = 1$ in (1.8) and (1.24). By setting $\rho_0 := \rho_\infty$ and $\rho_1 := \rho$ in (1.24) we obtain that

$$W_c(\rho, \rho_\infty) \leq H_V^F(\rho) - H_V^F(\rho_\infty), \tag{1.25}$$

which is an generalization of the transport inequality. By an approximation argument, we extend (1.25) to the case where the supports of ρ and ρ_∞ are not necessarily bounded.

To obtain a generalized version of the logarithmic Sobolev inequality, we set $\rho_0 := \rho$, $\rho_1 := \rho_\infty$ and again $\alpha_0 = 1$ in (1.24) to deduce that

$$H_V^F(\rho) - H_V^F(\rho_\infty) + W_c(\rho, \rho_\infty) \leq \int_{\mathbf{R}^d} (\mathbf{id} - T) \cdot \nabla(F' \circ \rho - F' \circ \rho_\infty)\rho \, dx. \tag{1.26}$$

Applying Young's inequality (1.5) to the right hand side of (1.26) we conclude that

$$\begin{aligned} H_V^F(\rho) - H_V^F(\rho_\infty) + W_c(\rho, \rho_\infty) &\leq \int_{\mathbf{R}^d} c(x - Tx)\rho \, dx \\ &+ \int_{\mathbf{R}^d} c^*(\nabla(F' \circ \rho - F' \circ \rho_\infty))\rho \, dx. \end{aligned} \tag{1.27}$$

Using (1.7), (1.11) in (1.27) we obtain

$$H_V^F(\rho) - H_V^F(\rho_\infty) \leq \int_{\mathbf{R}^d} c^*(\nabla(F' \circ \rho + V))\rho \, dx \leq I_{c^*}(\rho|\rho_\infty). \tag{1.28}$$

When $F(t) = t \log t - t$, and $c(z) = \lambda|z|^2/2$ ($\lambda > 0$), (1.28) is the classical logarithmic Sobolev inequality, in the form obtained by Bakry and Emery [4]) and the transport inequality (1.25) is then an extension obtained in [5, 7, 28] of the transport inequality studied by Talagrand [31] and Marton [23]. Since for general F, (1.28) may have no logarithmic term, it may seem misleading to refer to it as a "generalized logarithmic Sobolev inequality". This is why we often refer to it as a *generalized entropy-information inequality.* If $c(z) = \lambda|z|^2/2$ with $\lambda > 0$ and F that satisfy (H4-H5), (1.28) allows us to recover generalizations of the logarithmic Sobolev inequality obtained by Arnold, Carrillo, Juengel, Markovich, Toscani and Unterreiter [3, 11] with the Bakry-Emery semi-group method, and by Del Pino and Dolbeault [18] with a method from the calculus of variations. When c is homogeneous of degree $p \geq 2$ and $F(t) = t \log t - t$, (1.25) and (1.28) also recover results of Bobkov and Ledoux [7] obtained there as consequences of the Prékopa–Leindler inequality.

We stress again that the notion W_c-convexity along paths in the set of probability measures is mentioned only for the nice interpretation it provides but is not used in the proofs. We can also mention that for most of the inequalities presented here (in the case $\alpha_0 > 0$) the use of c-optimal maps is not compulsory: one can for instance work with the more classical "Brenier map" (which is the optimal map for the quadratic cost).

The present paper is organized as follows: in Section 2, we state inequality (1.21) obtained in [1] and [26] and we readily derive from it inequality (1.24). This inequality is then used in Section 3 to generalize the transport and logarithmic Sobolev inequalities. In Section 4 we briefly comment on functionals of the form $K(\rho) = H_V^F(\rho) + \int \rho\, W * \rho$, where a non-local term is added to H_V^F and where V satisfies (1.8) with $c(z) := |z|^2/2$. Convexity properties of K were studied in details in a recent work of Carrillo, McCann and Villani [12]. We show, as enquired by these authors, that the method of [13] as extended here, also applies to the situation where H_V^F is replaced by K (c remaining quadratic).

This collaboration was initiated while the first named author was visiting W. Gangbo and R. McCann at the Georgia Institute of Technology. He wishes to express his gratitude to this institution for its hospitality.

2 Uniform displacement convexity of generalized entropy

In this section, Ω is an open convex *bounded* (unless otherwise noted) subset of $\mathbf{R}^d$ and $\mathcal{P}^a(\Omega)$ denotes the subset of $\mathcal{P}^a$ of density functions defined on Ω; that is the set of Borel functions $\rho : \Omega \to [0, +\infty)$ such that $\int_\Omega \rho dx = 1$. F is also assumed to satisfy (H4-H5). We first state an "*energy inequality*" on $\mathcal{P}^a(\Omega)$, a result obtained by Otto [26] for the so-called Tsallis entropy functionals [30, 32], and later generalized by Agueh [1] to a wider class of entropy functionals. This energy inequality is a generalization of the *displacement convexity inequality* proved by McCann [25] when $c(z) = |z|^2/2$; it can informally be stated as

$$\int F(\rho_1)dx - \int F(\rho_0)dx \geq \Big[\frac{d}{dt}\int F(\rho_t)dx\Big]_{t=0}.$$

In fact, this generalized version is instrumental in studying existence of solutions of the quasi-linear elliptic–parabolic-degenerate equations in [1].

Proposition 2.1 (Agueh, Otto) *Let $\rho_0, \rho_1 \in \mathcal{P}^a(\Omega)$ with $\rho_0 \in W^{1,\infty}(\Omega)$ and $\inf_\Omega \rho_0 > 0$. Let also c and F satisfy respectively* (H1–H3) *and* (H4-H5)*. Then*

$$H^F(\rho_1) - H^F(\rho_0) \geq \int_\Omega (T - \mathbf{id}) \cdot \nabla(A \circ \rho_0) dx = \int_\Omega (T - \mathbf{id}).\nabla(F' \circ \rho_0)\, \rho_0\, dx, \tag{2.1}$$

where $A(0) = 0$, and $A(t) := tF'(t) - F(t)$ for $t > 0$, and where T is the c–optimal map such that $T_\# \rho_0 = \rho_1$.

Proof: The proof of Proposition 2.1 can be found in [1]. In Appendix A, we state Proposition 5.1, a variant of Proposition 2.1, for which, we provide a simpler proof than those in [1] and [26]. The generalized version of the logarithmic Sobolev inequality obtained below can also be derived from Proposition 5.1. However, the proof of the transport inequality seems to require rather involved approximation arguments. QED

Remark 2.2 (should one use the c-optimal map?) *What is apparent in [1] and from the proof of Proposition 5.1, is the fact that the energy inequality (2.1) continues to hold if one substitutes the c-optimal map T by any map S such that its differential "dS" (possibly defined in some weak sense) has only real nonnegative eigenvalues and such that $S_\# \rho_0 = \rho_1$. For exemple, one can use the Brenier map ($c(z) = |z|^2/2$) or the (triangular) Knothe map. We refer the reader to Remark 2.5 where we mention some advantages of working with the c-optimal map T instead of any other map S.*

Let $\{\rho_t\}$ be the displacement interpolant density functions introduced in (1.22). Formally

$$\frac{d}{dt} \left[H_V(\rho_t)\right]_{t=0} = \int \nabla V \cdot (T - \mathbf{id}) \rho_0 \, dx. \tag{2.2}$$

Since $\alpha_0 = 0$ in (1.8) means that V is convex, the next lemma can be interpreted as follow: V is convex if and only if $t \to \int V \rho_t \, dx$ is convex.

Lemma 2.3 *Let $V \in C^1(\mathbf{R}^d)$ and let Ω be an open convex subset of $\mathbf{R}^d$. The following assertions are equivalent:*

(i) V satisfies the pointwise inequality (1.8) for all $a, b \in \Omega$.

(ii) For $\rho_0, \rho_1 \in \mathcal{P}^a(\Omega)$ and, T the c–optimal map such that $T_\# \rho_0 = \rho_1$, we have that

$$H_V(\rho_1) - H_V(\rho_0) \geq \int_\Omega \nabla V \cdot (T - \mathbf{id}) \rho_0 dx + \alpha_0 W_c(\rho_0, \rho_1). \tag{2.3}$$

Proof: Assume first that (1.8) holds and let $\rho_0, \rho_1 \in \mathcal{P}^a(\Omega)$. Then

$$V(Tx) - V(x) \geq \nabla V(x) \cdot (Tx - x) + \alpha_0 c(x - T(x)),$$

for ρ_0-almost every $x \in \Omega$. Now, integrate both sides of the above inequality, use the fact that T is c-optimal and that $T_\# \rho_0 = \rho_1$ to obtain (2.3).

Conversely, assume that (2.3) holds for all $\rho_0, \rho_1 \in \mathcal{P}^a(\Omega)$. Let $a, b \in \Omega$. Choose a collection $\{\rho_0^r\} \subset \mathcal{P}^a(\Omega)$ supported in the ball of center a and radius r, and such that $\{\rho_0^r\}$ converges weak-$*$ to the Dirac mass at a. Define $T_o : x \to x + b - a$ and the measures

$$\rho_1^r = T_{o\#}\rho_0^r.$$

If T is another Borel map such that $\rho_1^r = T_\# \rho_0^r$ then, Jensen's inequality yields that

$$\int_\Omega c(x - Tx)\rho_0^r(x)dx \geq c\Big(\int_\Omega (x - Tx)\rho_0^r(x)dx\Big) = c(a-b) = \int_\Omega c(x - T_o x)\rho_0^r(x)dx.$$

This proves that T_o is the c-optimal map that pushes forward ρ_0^r to ρ_1^r. Using ρ_0^r (resp. ρ_1^r) in place of ρ_0 (resp. ρ_1) in (2.3) we have that

$$\int_\Omega (V(x+b-a) - V(x))\rho_0^r(x)dx \geq \int_\Omega \nabla V(x)\cdot(b-a)\rho_0^r(x)dx + \alpha_0 \int_\Omega c(b-a)\rho_0^r(x)dx. \quad (2.4)$$

Letting r go to 0 in (2.4) we obtain (1.8). QED

Our next result expresses the W_c–*semiconvexity* of the generalized entropy functional.

Theorem 2.4 (Evolution of H_V^F along c-optimal transport) *Let c satisfy* (H1–H3) *and let F satisfy* (H4-H5). *Let also $V \in C^1(\mathbf{R}^d)$ satisfies* (1.8), *for some $\alpha_0 \in \mathbf{R}$. Let $\rho_0, \rho_1 \in \mathcal{P}^a(\Omega)$ be such that $\rho_0 \in W^{1,\infty}(\Omega)$ and $\inf_\Omega \rho_0 > 0$. Then if T is the c-optimal transport pushing ρ_0 forward to ρ_1 one has:*

$$H_V^F(\rho_1) - H_V^F(\rho_0) \geq \alpha_0 W_c(\rho_0, \rho_1) + \int_\Omega (T - \mathbf{id})\cdot\nabla(F'(\rho_0) + V)\,\rho_0 dx. \quad (2.5)$$

Proof: Combine Proposition 2.1 and Lemma 2.3. QED

Remark 2.5 (why do we use the c-optimal map? (bis)) *When $\alpha_0 < 0$, and unlike in Proposition 2.1, we don't know how to prove (2.5) without appealing to the c-optimal map T pushing ρ_0 forward to ρ_1. Also, it is convenient to use the map T so that one could interpret (2.5) as a c–displacement convexity as explained in the introduction. Finally, the use of the map T becomes crucial when studying parabolic-elliptic PDEs as in [1].*

Remark 2.6 (Other assumptions) *Using Proposition 5.1, Theorem 2.4 can be restated with slightly different assumptions on ρ_0 and ρ_1. Inequality (2.5) is then valid if ρ_0 and ρ_1 are only assumed to be compactly supported Borel probability densities (no assumption on Ω is required), but we then ask that $\rho_0 \in W^{1,\infty}(\mathbf{R}^d)$.*

Remark 2.7 (Probability densities?) *The results clearly remain valid if instead of requiring that $\int_{\mathbf{R}^d} \rho_i\, dx = 1$ $(i = 0, 1)$ we impose $0 < \int_{\mathbf{R}^d} \rho_0\, dx = \int_{\mathbf{R}^d} \rho_1\, dx < +\infty$.*

3 Generalized transport and entropy–information inequalities

Throughout this section, we assume that $\Omega \subset \mathbf{R}^d$ is open, convex, and possibly $\mathbf{R}^d$. We denote as before by $\mathcal{P}^a(\Omega)$ the set of probability densities on Ω. We assume that $V \in C^1(\mathbf{R}^d)$ and that (1.8) holds with $\alpha_0 = 1$ (for positive α_0, replace, as we already said, the convex function c by $\alpha_0 c$). We assume that F satisfies (H4 − H5) and that there exists $\rho_\infty \in \mathcal{P}^a(\Omega)$, such that $\rho_\infty > 0$ on Ω with also

$$F'(\rho_\infty) + V = 0 \text{ on } \Omega.$$

To ensure that $H_V^F(\rho_\infty)$ is finite, we assume that $F(\rho_\infty) + \rho_\infty V$ is in L^1. Our next result, Corollary 3.1, is a generalization of a transport inequality, due to Talagrand when $c(z) = |z|^2/2$, $F(t) = t \log t - t$ and ρ_∞ is the standard Gaussian density in $\Omega := \mathbf{R}^d$. In Corollary 3.2, we extend the logarithmic Sobolev inequality to general cost functions c.

Corollary 3.1 (Transport inequality) *Let c and F satisfy respectively* (H1–H3) *and* (H4-H5)*. Assume that $V \in C^1(\mathbf{R}^d)$ verifies* (1.8) *for $\alpha_0 = 1$, and that $\rho_\infty > 0$ as above. Then for every Borel measurable function $f : \Omega \to [0, +\infty)$ such that $\rho := f\rho_\infty \in \mathcal{P}^a(\Omega)$, we have that*

$$W_c(\rho, \rho_\infty) \leq H_V^F(\rho) - H_V^F(\rho_\infty). \tag{3.1}$$

Proof: Recall that $F'(\rho_\infty)+V = 0$, and so, the proof of Corollary 3.1 would be straightforward if we could set in (2.5), $\rho_0 := \rho_\infty$ and $\rho_1 := \rho$. Unfortunately, Ω may be unbounded, or $\rho_0 := \rho_\infty$ and $\rho_1 := \rho$ may not satisfy the assumptions of Theorem 2.4. We further assume that $H_V^F(\rho) \neq +\infty$ since otherwise there will be nothing to prove. In light of (1.12), this means that $F(\rho) + \rho V$ is integrable.

To apply Theorem 2.4, we first assume without loss of generality that $f \in C(\Omega)$, and that $\inf_\Omega f > 0$. We approximate ρ and ρ_∞ as follows: let $\{\Omega_n\}_{n=1}^\infty$ be a sequence of open, convex, bounded subsets of $\mathbf{R}^d$ such that $\bar{\Omega}_n \subset \Omega_{n+1}$, and $\cup_{n=1}^\infty \Omega_n = \Omega$. Let χ_{Ω_n} be the characteristic function of Ω_n. Define

$$\rho_{0,n} := \rho_\infty \chi_{\Omega_n},$$

and set $p_n := \int_{\Omega_n} \rho_\infty dx$. By Lemma 6.1 there exists a sequence $\{\rho_{1,n}\}_{n=1}^\infty$ such that

$$\int_\Omega \rho_{1,n} dx = \int_\Omega \rho_{0,n} dx, \qquad \text{and} \quad \lim_{n\to+\infty} H_V^F(\rho_{1,n}) = H_V^F(\rho). \tag{3.2}$$

Since $F'(\rho_{0,n}) + V = 0$ on Ω_n, our main result, Theorem 2.4, gives that

$$W_c(\rho_{1,n}, \rho_{0,n}) \leq H_V^F(\rho_{0,n}) - H_V^F(\rho_{1,n}). \tag{3.3}$$

We combine (3.3) and (3.2), and use that W_c is lower semicontinuous [29], to conclude the proof of Corollary 3.1. QED

Corollary 3.2 (Generalized entropy–information inequality) *Assume that c, F satisfy respectively* (H1–H3), (H4-H5), *and that $V \in C^1(\mathbf{R}^d)$ satisfies (1.8) with $\alpha_0 = 1$. Assume $\rho_\infty \in C(\Omega)$ is a positive probability density verifying $F'(\rho_\infty) = -V$ on Ω. Then for every positive $f \in C(\Omega)$ such that $\rho := f\rho_\infty \in \mathcal{P}^a(\Omega) \cap C^1(\Omega)$, we have that*

$$H_V^F(\rho) - H_V^F(\rho_\infty) \le \int_{\mathbf{R}^d} c^*(\nabla[F' \circ \rho + V])\rho \, dx \le I_{c^*}(\rho|\rho_\infty). \tag{3.4}$$

Proof: Note that since (1.7) gives that $c^*(y) \le \nabla c^*(y) \cdot y$ for all $y \in \mathbf{R}^d$, the second inequality in (3.4) is straightforward to obtain. The task is then to establish the first inequality in (3.4) and, to do so, we assume without loss of generality that $\int_{\mathbf{R}^d} c^*(-\nabla[F' \circ \rho + V])\rho \, dx$ is finite. The result would follow easily from Young's inequality(1.5), if we could set $\rho_0 := \rho$ and $\rho_1 := \rho_\infty$ in theorem 2.4. As noted in the previous corollary, ρ or ρ_∞ may have supports which are not bounded, and so we will approximate them as follows: let $\{\Omega_n\}_{n=1}^\infty$ be a sequence of open, convex, bounded subsets of $\mathbf{R}^d$ such that $\bar{\Omega}_n \subset \Omega_{n+1}$, and $\cup_{n=1}^\infty \Omega_n = \Omega$. We denote by χ_{Ω_n} the characteristic function of Ω_n. Define

$$\rho_{1,n} := \rho_\infty \chi_{\Omega_n},$$

and set $p_n := \int_{\Omega_n} \rho_\infty dx$. Lemma 6.1 gives the existence of a sequence of positive functions $\{\rho_{0,n}\}_{n=1}^\infty \subset C^1(\bar{\Omega}_n)$ converging to ρ in $L^1(\Omega)$ as n tends to $+\infty$. Furthermore, $\rho_{0,n} = \rho$ on $\Omega_n \setminus \Omega_1$ and

$$\rho_{0,n} \text{ converges strongly to } \rho \text{ in } W^{1,\infty}(\Omega_1), \tag{3.5}$$

$$\int_\Omega \rho_{0,n} dx = \int_\Omega \rho_{1,n} dx, \qquad \text{and} \quad \lim_{n\to+\infty} H_V^F(\rho_{0,n}) = H_V^F(\rho). \tag{3.6}$$

Denote by T_n the c–optimal such that $T_{n\#}\rho_{0,n} = \rho_{1,n}$. In light of Theorem 2.4, we have that

$$H_V^F(\rho_{1,n}) - H_V^F(\rho_{0,n}) \ge W_c(\rho_{1,n}, \rho_{0,n}) + \int_\Omega (T_n - \mathbf{id}) \cdot \nabla(F'(\rho_{0,n}) + V) \, \rho_{0,n} dx.$$

This, together with Young's inequality (1.5) and, the fact that $W_c(\rho_{1,n}, \rho_{0,n}) = \int_\Omega c(T_n - \mathbf{id})\rho_{0,n} \, dx$ yields

$$H_V^F(\rho_{1,n}) - H_V^F(\rho_{0,n}) \ge -\int_{\Omega_n} c^*\Big(-\nabla(F'(\rho_{0,n}) + V)\Big) \, \rho_{0,n} \, dx. \tag{3.7}$$

Finally, observe that

$$\begin{aligned} &\int_{\Omega_n} c^*\Big(\nabla(F'(\rho_{0,n}) + V)\Big) \, \rho_{0,n} \, dx \\ &= \int_{\Omega_1} \Big[c^*(\nabla(F' \circ \rho_{0,n} + V))\rho_{0,n} - c^*(\nabla(F' \circ \rho + V))\rho\Big] dx + \int_{\Omega_n} c^*\Big(\nabla(F' \circ \rho + V)\Big)\rho dx \\ &\le \int_{\Omega_1} \Big[c^*(\nabla(F' \circ \rho_{0,n} + V))\rho_{0,n} - c^*(\nabla(F' \circ \rho + V))\rho\Big] dx + \int_{\Omega} c^*\Big(\nabla(F' \circ \rho + V)\Big)\rho dx, \end{aligned} \tag{3.8}$$

where, in the last inequality, we have used that c^* is nonnegative. We let n tend to $+\infty$ in (3.7) and (3.8), we use (3.6) together with the fact that ρ_∞ minimizes H_V^F to conclude the proof of the corollary. QED

Remark 3.3 *By standard approximation arguments, one can extend Corollary 3.1 and Corollary 3.2 to a larger class of density functions $\rho \in \mathcal{P}^a(\Omega)$. The reader can also see that the combination of Lemma 2.3 and Proposition 5.1 yields corollary 3.2, for more general domains Ω and functions $\rho \in \mathcal{P}^a(\Omega)$ (see Remark 2.6).*

Let us write Corollary 3.2 when $F(t) = t\log t - t$. If ρ_∞ is a probability density of the form $\rho_\infty = e^{-V}$, then ρ_∞ verifies $F'(\rho_\infty) + V = 0$ and for every non-negative g with $\int g\,\rho_\infty dx = 1$ we have that

$$H_V^F(g\rho_\infty) - H_V^F(\rho_\infty) = \int_{\mathbf{R}^d} g\log(g)\,\rho_\infty\,dx$$

is the entropy of g with respect to ρ_∞. Thus (3.4) with $\rho = f\rho_\infty$ gives:

Corollary 3.4 (Log-Sobolev inequality for c-uniformly convex potentials) *Let ρ_∞ be a probability density of the form:*

$$\rho_\infty(x) = e^{-V(x)}.$$

Assume that c satisfies (H1–H3) *and that $V \in C^1(\mathbf{R}^d)$ satisfies (1.8) with $\alpha_0 = 1$. Then for every smooth compactly supported non-negative function f such that $\int f\,\rho_\infty dx = 1$ we have:*

$$\int_{\mathbf{R}^d} f\log(f)\,\rho_\infty\,dx \le \int_{\mathbf{R}^d} c^*\left(\frac{\nabla f}{f}\right) f\,\rho_\infty\,dx. \tag{3.9}$$

Note that by approximation (3.9) holds when no other assumptions beyond convexity is made on c, and when c is not assumed to be even. In that case one needs to replace the expression $c^*\left(\frac{\nabla f}{f}\right)$ by $c^*\left(-\frac{\nabla f}{f}\right)$.

It is interesting to note that when $\mathrm{Hess}V \ge \lambda Id$ ($\lambda > 0$), (1.8) is satisfied with $c(z) := \lambda|z|^2/2$ and the log-Sobolev inequality (3.9) becomes, after setting $f = g^2$,

$$\int_{\mathbf{R}^d} g^2\log(g^2)\,\rho_\infty\,dx \le \frac{2}{\lambda}\int_{\mathbf{R}^d} |\nabla g|^2\,\rho_\infty\,dx, \tag{3.10}$$

for every smooth compactly supported g with $\int g^2\,\rho_\infty = 1$. This is the classical Bakry–Emery log-Sobolev inequality [4]. A more general inequality due to Bobkov and Ledoux can also be recovered.

Example 3.5 (A logarithmic Sobolev inequality of Bobkov and Ledoux) *The previous inequality (3.4), in the form (3.9), extends a result of Bobkov and Ledoux [7]. Let $\|\cdot\|$ be a norm on $\mathbf{R}^d$ and V a convex potential uniformly p-convex with respect to $\|\cdot\|$ for some $p \ge 2$, ie: there exists a constant $\delta > 0$ such that for every $x, y \in \mathbf{R}^d$,*

$$V(x) + V(y) - 2V\left(\frac{x+y}{2}\right) \ge \frac{\delta}{p}\|x-y\|^p \tag{3.11}$$

or equivalently $tV(x) + sV(y) - V(tx+sy) \ge (\delta\min(t,s)/p)\|x-y\|^p$, for $t+s=1,\ t,s \ge 0$ (see [7] for the equivalence). The typical situation is when $\|z\| := \|z\|_p$ is the ℓ_p-norm on $\mathbf{R}^d$ and $V(x) := \|x\|_p^p$. Then inequality (3.11) is satisfied with the optimal $\delta := \delta_p := p2^{-p}$. Denote by q the conjugate number of p, $q = p/(p-1)$ and by $\|\cdot\|_*$ the dual norm of $\|\cdot\|$.

If $\rho_\infty := e^{-V}$ is a probability density with V verifying (3.11), Bobkov and Ledoux proved, using the Prékopa–Leindler inequality, that for very positive smooth and compactly supported function f such that $\int_{\mathbf{R}^d} f^q \rho_\infty dx = 1$, one has

$$\int_{\mathbf{R}^d} f^q \log(f^q) \rho_\infty \, dx \le \left(\frac{q}{\delta}\right)^{q-1} \int_{\mathbf{R}^d} \|\nabla f\|_*^q \rho_\infty \, dx. \tag{3.12}$$

We will see that this is a particular case of our logarithmic Sobolev inequality. For $a, b \in \mathbf{R}^d$, as mentioned, condition (3.11) reads as, for $t \le 1/2$,

$$(1-t)\beta(0) + t\beta(1) - \beta(t) \ge \frac{\delta t}{p}\|b\|^p$$

where $\beta(t) := V(a+tb) = V((1-t)a + t(a+b))$. Looking at the first order Taylor expansion at $t=0$ in the previous inequality we obtain

$$\beta(1) - \beta(0) \ge \beta'(0) + \frac{\delta}{p}\|b\|^p.$$

This is equivalent to saying that V satisfies condition (1.8) with $\alpha_0 = 1$ and $c(z) := \frac{\delta}{p}\|z\|^p$. We next apply Corollary 3.4. Inequality (3.4) applied to $\rho = f^q \rho_\infty$ gives,

$$\begin{aligned} \int_{\mathbf{R}^d} f^q \log(f^q) \rho_\infty \, dx &\le \int_{\mathbf{R}^d} c^* \left(\nabla[\log f^q]\right) f^q \rho_\infty \, dx \\ &= \int_{\mathbf{R}^d} c^* \left(\frac{q \nabla f}{f}\right) f^q \rho_\infty \, dx \end{aligned}$$

Since $(\|\cdot\|^p/p)^*(z) = \|z\|_*^q/q$ we have, using the homogeneity of $\|\cdot\|_*^q$:

$$c^*(z) = \frac{1}{\delta^{q-1} q}\|z\|_*^q$$

and thus we recover exactly inequality (3.12).

To conclude this section we would like to comment on consequences of our results and on related problems. We only work out a few applications, others, such as concentration inequalities, HWI inequalities, etc. are also possible.

Transport implies entropy–information

As in Otto–Villani [28] one can use (2.5) to prove that even when V fails to be convex, a transport inequality implies a log-Sobolev inequality provided that V satisfies an appropriate c–semiconvexity property. Our precise statement is the following:

Proposition 3.6 *Assume as in Theorem 2.4 that c and F respectively satisfy* (H1–H3), *and* (H4-H5). *Assume that $V \in C^1(\mathbf{R}^d)$ satisfies* (1.8) *for some $\alpha_0 \le 0$, that $\rho_\infty \in C^1(\mathbf{R}^d)$, and* *that $F'(\rho_\infty) + V = 0$. Eventually, suppose that ρ_∞ satisfies a transport inequality: for every probability density ρ,*

$$W_c(\rho, \rho_\infty) \leq \frac{1}{\beta_1}(H_V^F(\rho) - H_V^F(\rho_\infty)) \tag{3.13}$$

for some $\beta_1 > 0$. If $0 < k < \beta_1 + \alpha_0$ then ρ_∞ satisfies the following log-Sobolev inequality:

$$H_V^F(\rho) - H_V^F(\rho_\infty) \leq \frac{\beta_1}{\beta_1 + \alpha_0 - k} \int_{\mathbf{R}^d} C^*(\nabla[F'(\rho) + V])\rho dx \leq \frac{\beta_1}{\beta_1 + \alpha_0 - k} I_{C^*}(\rho|\rho_\infty),$$

for all smooth densities ρ. Here, $C := kc$. When c is homogeneous of degree $p > 1$, the previous inequality can be replaced by

$$H_V^F(\rho) - H_V^F(\rho_\infty) \leq \frac{qp^{q-1}\beta_1}{(\beta_1 + \alpha_0)^q} \int_{\mathbf{R}^d} c^*(\nabla[F'(\rho) + V])\rho dx \leq \frac{qp^{q-1}\beta_1}{(\beta_1 + \alpha_0)^q} I_{c^*}(\rho|\rho_\infty),$$

Proof: We use Theorem 2.4, where we substitute the convex function c by the convex function C. As in the proof of Corollary 3.2, we use an approximation argument and Young's inequality to conclude that

$$H_V^F(\rho_\infty) - H_V^F(\rho) \geq (\alpha_0 - k)W_c(\rho, \rho_\infty) - \int_{\mathbf{R}^d} C^*(\nabla[F'(\rho) + V])\rho dx \tag{3.14}$$

In fact, we have used an extension of inequality (2.5) to densities with unbounded supports, with $\rho_0 := \rho$ and $\rho_1 := \rho_\infty$. We combine (3.13) and (3.14) to conclude that

$$\Big(1 - \frac{k - \alpha_0}{\beta_1}\Big) H_V^F(\rho) - H_V^F(\rho_\infty) \leq \int_{\mathbf{R}^d} C^*(\nabla[F'(\rho) + V])\rho dx.$$

This, together with the fact that $C^*(z) \leq \nabla C^*(z) \cdot z$, yields the claimed inequality. When c is homogeneous of degree p, the result is obtained by optimizing in k. QED

Poincaré type inequalities

It is well known that the linearization of a log-Sobolev type inequality gives a Poincaré type inequality. This is also true in our general situation. We have in view the quadratic case $c(z) = |z|^2/2$, but we only assume here that c is homogeneous of degree 2. We assume furthermore that the conditions of Theorem 3.2 are satisfied. As before, let $A(t) = tF'(t) - F(t)$. We apply (3.4) with $\rho_\varepsilon = (1 + \varepsilon f)\rho_\infty$ where $\int f\rho_\infty = 0$. It is easily checked that,

$$H_V^F(\rho_\varepsilon) - H_V^F(\rho_\infty) = \frac{\varepsilon^2}{2} \int_{\mathbf{R}^d} f^2 A'(\rho_\infty)\rho_\infty dx + o(\varepsilon^2)$$

and by homogeneity

$$\int_{\mathbf{R}^d} c^*(\nabla[F'(\rho_\varepsilon) + V])\rho_\varepsilon dx = \varepsilon^2 \int_{\mathbf{R}^d} c^*(\nabla[fA'(\rho_\infty)])\rho_\infty dx + o(\varepsilon^2)$$

when $\varepsilon \to 0$. Thus a second order Taylor expansion in (3.4) gives the following Poincaré type inequality: *if f is a smooth compactly supported function such that $\int f\rho_\infty dx = 0$, one has, under the assumptions of Theorem 3.2 (with c homogeneous of degree 2),*

$$\int_{\mathbf{R}^d} f^2 A'(\rho_\infty)\rho_\infty dx \leq 2 \int_{\mathbf{R}^d} c^*(\nabla[fA'(\rho_\infty)])\rho_\infty dx$$

Note that $F(t) = t\log t - t$ gives $A(t) = t$ and one then recovers classical Poincaré inequalities (with $c(z) = |z|^2/2$).

Infimal convolution inequalities

In the classical case, where $F(t) = t\log t - t$ and $c(z) = |z|^2/2$, it was noticed by Bobkov and Götze [6] that transport inequalities are dual version of infimal convolution inequalities. One can check, that this is still true in our context and that such a duality amounts to an inequality between Legendre transforms, at least at the formal level. The arguments we present here are not rigorous although we believe they can be put into a satisfactory abstract framework.

As earlier, we assume that ρ_∞ a Borel probability measure verifying $F'(\rho_\infty) + V = 0$. We consider functions η that are continuous and compactly supported in $\mathbf{R}^d$. We denote by $\mathcal{P}$ the set of Borel probability measures. As in [17], the c-transform of η is given by the infimal convolution

$$\eta^c(y) := \inf_{x\in\mathbf{R}^d} c(y-x) - \eta(x).$$

Introduce the functional

$$\mathcal{G}(\eta) := -\int_{\mathbf{R}^d} \eta^c \rho_\infty dx.$$

One can check that, for a Radon measure ν,

$$\mathcal{G}^*(\nu) := \sup_\eta \int_{\mathbf{R}^d} \eta\, d\nu - \mathcal{G}(\eta) = \begin{cases} W_c(\nu, \rho_\infty) & \text{if } \nu \in \mathcal{P} \\ +\infty & \text{otherwise} \end{cases}$$

Now, for each η there exists $\lambda_\eta \in \mathbf{R}$ such that $\rho_\eta := (F^*)'(\eta - V - \lambda_\eta)$ is a probability ($\rho_\eta \in \mathcal{P}$). Note that $\rho_\infty = (F^*)'(-V)$ is recovered for $\eta \equiv 0$. Introduce the functional

$$\mathcal{F}(\eta) := \lambda_\eta + \int_{\mathbf{R}^d} F^*(\eta - V - \lambda_\eta) dx.$$

Then, the reader can check that for $\rho \in \mathcal{P}^a$ one has formally

$$\mathcal{F}^*(\rho) = H_V^F(\rho).$$

Assume that for every η:

$$\int_{\mathbf{R}^d} \eta^c \rho_\infty\, dx + \int_{\mathbf{R}^d} F^*(\eta - V - \lambda_\eta) dx + H_V^F(\rho_\infty) + \lambda_\eta \le 0. \tag{3.15}$$

Then this implies $\mathcal{F} \le \mathcal{G} - H_V^F(\rho_\infty)$, and thus we can deduce the transport inequality

$$W_c(\rho, \rho_\infty) \le H_V^F(\rho) - H_V^F(\rho_\infty),$$

for every $\rho \in \mathcal{P}^a$. When $F(t) := t\log t - t$, inequality (3.15) becomes the infimal convolution inequality for the measure $\rho_\infty = e^{-V}$ studied by Bobkov and Götze [6]. Inequality (3.15) is then closely related to the Prékopa–Leindler inequality, in particular in the form put forward by Maurey [24] under the name of property (τ). In our general situation, we do not know an adapted Prékopa–Leindler inequality which would make use of (3.15).

4 On a question of Carrillo, McCann and Villani

Throughout this section $V \in C^2(\mathbf{R}^d)$ (confinement potential) and, $W \in C^2(\mathbf{R}^d)$ is even (interaction potential). We assume that the function F as before, satisfies (H4-H5). In [12], Carrillo, McCann and Villani study the PDE

$$\frac{\partial \rho}{\partial t} = \operatorname{div}\left(\rho \nabla [\, F'(\rho) + V + W * \rho\,]\right). \tag{4.1}$$

To this aim they introduced the energy (or entropy) functional

$$\begin{aligned} K(\rho) &:= \int_{\mathbf{R}^d} [F(\rho) + V\rho + \frac{1}{2}(W * \rho)\rho\,] dx \\ &= H^F(\rho) + H_V(\rho) + H_W(\rho) \end{aligned} \tag{4.2}$$

where H^F and H_V are defined as before by (1.9)–(1.10) and $H_W(\rho) := \frac{1}{2}\int_{\mathbf{R}^d}(W*\rho)\rho dx$. Then, accordingly, the information functional here is

$$J(\rho) := \int_{\mathbf{R}^d} \Big|\nabla [\, F'(\rho) + V + W * \rho\,]\Big|^2 \rho\, dx. \tag{4.3}$$

It is well known, as explained in the introduction, that logarithmic Sobolev type inequalities linking K and J provide a control on the rates of convergence to the ground state for solutions of (4.1). In [12], a proof of the logarithmic Sobolev and "HWI" type inequalities was provided by following the strategy of [27, 28] based, roughly speaking, on interpolation along mass transport. In [12], the authors made the following statement: "It will be interesting to see if the argument of [13] can be extended, to provide a simplified proof for the inequalities [...] which will show up in the present work." Here, we positively answer that question. Recall that the Brenier map refers to the unique optimal map (for the quadratic cost $c(z) = |z|^2/2$) pushing forward a probability density onto another. This map T is the gradient of a convex function ϕ, $T = \nabla\phi$, and it can be written as $T(x) = x + \nabla\theta(x)$ by setting $\theta(x) := \phi(x) - |x|^2/2$.

Theorem 4.1 *Assume that V and F given above satisfy in addition the following assumptions:*

$$\mathrm{Hess} V \geq \lambda I_d, \quad \mathrm{Hess} W \geq \mu I_d$$

for some real numbers λ and μ not necessarily nonnegative. Here, I_d is the $d \times d$ identity matrix. Let ρ_0 and ρ_1 be compactly supported probability density functions, and let $T(x) = \nabla\phi(x) = x + \nabla\theta(x)$ be the Brenier map pushing ρ_0 forward to ρ_1. Then,

$$K(\rho_1) - K(\rho_0) \geq \int_{\mathbf{R}^d} \nabla\theta.\nabla[\, F'(\rho_0) + V + W * \rho_0\,]\, \rho_0 dx + \frac{\mu+\lambda}{2}\int_{\mathbf{R}^d} |\nabla\theta|^2 \rho_0 dx - \frac{\mu}{2}|m_1(\rho_1 - \rho_0)|. \tag{4.4}$$

Here $m_1(\rho_1 - \rho_0)$ stands for $\int_{\mathbf{R}^d} x(\rho_1(x) - \rho_0(x))dx$, the difference between the center of mass of ρ_1 and that of ρ_0.

Proof: We could use Proposition 5.1 to conclude that

$$H^F(\rho_1) - H^F(\rho_0) \geq \int_{\mathbf{R}^d} \nabla\theta.\nabla[F'(\rho)]\rho_0\, dx. \tag{4.5}$$

In the present case, where $c(z) = |z|^2/2$, the proof of (4.5) is however simpler, and there is no need to make statements as strong as those appearing in Proposition 5.1. Since T is the gradient of a convex function (the so-called Brenier map) the method used in [13] (to whom we refer for precise definitions) in the case $F(t) = t\log(t) - t$, applies here. Let us recall again the ingredients, which are of course basically the same than those used in Proposition 5.1. The Monge-Ampère equation

$$\rho_0(x) = \rho_1(T(x))\det(I + \mathrm{Hess}\theta) \tag{4.6}$$

holds ρ_0-almost everywhere when Hessθ is understood as the Hessian of θ in the sense of Aleksandrov. Then, condition (H5) combined with the $1/d$ concavity of the determinant on non-negative matrices and (4.6) implies

$$F(\rho_1(T))/\rho_1(T) - F(\rho_0)/\rho_0 \geq -(A'(\rho_0)/\rho_0)\Delta_A\theta,$$

where $\Delta_A\theta := \mathrm{tr}(\mathrm{Hess}\theta)$ and $A(t) := tF'(t) - F(t)$. We conclude by integrating with respect to ρ_0 and noticing that the distributional Laplacian dominates the Laplacian in the sense of Aleksandrov Δ_A (this allows to integrate by parts).

The analogue of Lemma 2.3 is of course straightforward:

$$H_V(\rho_1) - H_V(\rho_0) = \int_{\mathbf{R}^d} [V(x + \nabla\theta(x)) - V(x)]\rho_0\, dx \geq \int_{\mathbf{R}^d} \nabla\theta.\nabla V\, \rho_0\, dx + \frac{\lambda}{2}\int_{\mathbf{R}^d} |\nabla\theta|^2\rho_0\, dx. \tag{4.7}$$

We now deal with the new term H_W. We have

$$\begin{aligned}
H_W(\rho_1) &= \frac{1}{2}\int_{\mathbf{R}^d\times\mathbf{R}^d} W(x-y)\rho_1(x)\rho_1(y)\, dxdy = \frac{1}{2}\int_{\mathbf{R}^d\times\mathbf{R}^d} W(T(x) - T(y))\rho_0(x)\rho_0(y)\, dxdy \\
&\geq \frac{1}{2}\int_{\mathbf{R}^d\times\mathbf{R}^d} \Big[W(x-y) + \nabla W(x-y).(\nabla\theta(x) - \nabla\theta(y)) + \\
&\qquad\qquad \frac{\mu}{2}|\nabla\theta(x) - \nabla\theta(y)|^2\Big]\rho_0(x)\rho_0(y)\, dxdy \\
&= H_W(\rho_0) + \frac{1}{2}\int_{\mathbf{R}^d\times\mathbf{R}^d} \nabla W(x-y).(\nabla\theta(x) - \nabla\theta(y))\, \rho_0(x)\rho_0(y)\, dxdy + \\
&\qquad \frac{\mu}{4}\int_{\mathbf{R}^d\times\mathbf{R}^d} |\nabla\theta(x) - \nabla\theta(y)|^2\rho_0(x)\rho_0(y)\, dxdy.
\end{aligned}$$

Using that ∇W is odd one can readily check that

$$\int_{\mathbf{R}^d\times\mathbf{R}^d} \nabla W(x-y)\cdot(\nabla\theta(x) - \nabla\theta(y))\, \rho_0(x)\rho_0(y)\, dxdy = 2\int_{\mathbf{R}^d} \nabla\theta.\nabla(W*\rho_0)\, \rho_0\, dx.$$

Therefore,

$$H_W(\rho_1) - H_W(\rho_0) \geq \int_{\mathbf{R}^d} \nabla\theta.\nabla(W*\rho_0)\, \rho_0\, dx + \frac{\mu}{4}\int_{\mathbf{R}^d\times\mathbf{R}^d} |\nabla\theta(x) - \nabla\theta(y)|^2\rho_0(x)\rho_0(y)\, dxdy. \tag{4.8}$$

Combining (4.5), (4.7) and (4.8) we have that

$$K(\rho_1) - K(\rho_0) \;\geq\; \int_{\mathbf{R}^d} \nabla\theta . \nabla[\, F'(\rho_0) + V + W * \rho_0 \,]\rho_0 \, dx \; + \tag{4.9}$$

$$\frac{\lambda}{2}\int_{\mathbf{R}^d} |\nabla\theta|^2 \rho_0 dx + \frac{\mu}{4}\int_{\mathbf{R}^d\times\mathbf{R}^d} |\nabla\theta(x) - \nabla\theta(y)|^2 \rho_0(x)\rho_0(y) dxdy. \tag{4.10}$$

Note that

$$\int_{\mathbf{R}^d\times\mathbf{R}^d} |\nabla\theta(x) - \nabla\theta(y)|^2 \rho_0(x)\rho_0(y)\, dxdy = 2\int_{\mathbf{R}^d} |\nabla\theta|^2 \rho_0 \, dx - 2\left|\int_{\mathbf{R}^d} \nabla\theta \, \rho_0 \, dx\right|^2 \tag{4.11}$$

and so, since $T_\# \rho_0 = \rho_1$ we have that

$$\int_{\mathbf{R}^d} \nabla\theta \, \rho_0 \, dx = \int_{\mathbf{R}^d} (T(x) - x)\rho_0(x) \, dx = \int_{\mathbf{R}^d} y\rho_1(y) dy - \int_{\mathbf{R}^d} x\rho_0(x) \, dx. \tag{4.12}$$

We combine (4.10), (4.10) and (4.12) to conclude the proof of the theorem. QED

With Theorem 4.1 on hand, one can recover several results of [12] such as interpolation (HWI), logarithmic Sobolev, or transport inequalities, just as we obtained this type of inequalities from Theorem 2.4. For the sake of illustration, let us derive in Corollary 4.2 an extension of the logarithmic Sobolev inequality. For that, we assume here that ρ_∞ is the ground state for the equation (4.1). In other words, it is the probability density characterized by

$$F'(\rho_\infty) + V + W * \rho_\infty \; = 0 \; = J(\rho_\infty).$$

Corollary 4.2 (Carrillo-McCann-Villani [12]) *Assume that as in Theorem 4.1*

$$\mathrm{Hess}V \geq \lambda I_d \quad \textit{and} \quad \mathrm{Hess}W \geq \mu I_d$$

for some real numbers λ and μ. Let ρ be a smooth probability density. Assume moreover that either $\mu \geq 0$ or that $\int_{\mathbf{R}^d} x\rho(x)dx = \int_{\mathbf{R}^d} y\rho_\infty(y)dy$. Set $k := \lambda$ if $\mu \geq 0$. When $\mu < 0$ and $\int_{\mathbf{R}^d} x\rho(x)dx = \int_{\mathbf{R}^d} y\rho_\infty dy(y)$, set $k := \lambda + \mu$. Choose k to be either of these values if both $\mu \geq 0$ and $\int_{\mathbf{R}^d} x\rho(x)dx = \int_{\mathbf{R}^d} y\rho_\infty dy(y)$. Then, if $k > 0$, one has

$$K(\rho) - K(\rho_\infty) := K(\rho|\rho_\infty) \leq \frac{1}{2k} J(\rho).$$

Proof: Without loss of generality, we can assume ρ to be compactly supported. Let $\tilde\rho$ be any compactly supported probability density. When $m_1(\rho_1 - \rho_0) = 0$, (4.4) reads as:

$$K(\rho_1) - K(\rho_0) \geq \int_{\mathbf{R}^d} \nabla\theta . \nabla[\, F'(\rho_0) + V + W * \rho_0 \,]\, \rho_0 dx \; + \frac{k}{2}\int_{\mathbf{R}^d} |\nabla\theta|^2 \rho_0 dx. \tag{4.13}$$

When $m_1(\rho_1 - \rho_0) \neq 0$ but $\mu \geq 0$, we use (4.12) and the fact that, by Jensen's inequality, $\int_{\mathbf{R}^d} |\nabla\theta|^2 \rho_0 dx \geq |\int_{\mathbf{R}^d} \nabla\theta\rho_0 dx|^2$, to obtain (4.13). Now, (4.13) and Young's inequality (1.5) with $c(z) := |z|^2/2$ yields that

$$K(\tilde\rho) - K(\rho) \geq -\frac{1}{2k}\int_{\mathbf{R}^d} |\nabla[\, F'(\rho) + V + W * \rho \,]|^2 \, \rho dx.$$

We conclude the proof of the theorem by taking the infimum over all $\tilde\rho$ and, by using that $\inf_{\tilde\rho} K(\tilde\rho) = K(\rho_\infty)$. QED

5 Appendix A: A proof of an energy inequality

Proposition 5.1 *Let ρ_0, ρ_1 be compactly supported probability densities with $\rho_0 \in W^{1,\infty}(\mathbf{R}^d)$. Let c, and F be such that* (H1–H3) *and* (H4-H5) *hold. Then*

$$H^F(\rho_1) - H^F(\rho_0) \geq \int_{\mathbf{R}^d} (T - \mathbf{id}) \cdot \nabla(A(\rho_0))dx = \int_{\mathbf{R}^d} (T - \mathbf{id}).\nabla(F'(\rho_0))\, \rho_0 dx,$$

where $A(0) = 0$, and $A(t) := tF'(t) - F(t) \geq 0$, for $t > 0$, and where T is the c–optimal map such that $T_\# \rho_0 = \rho_1$.

Proof: We first assume that c and c^* are C^2. Then by [14] we know that for ρ_0 almost every x there exists a linear map denoted by dT_x such that

$$\rho_0(x) = \rho_1(T(x)) \det dT_x. \tag{5.14}$$

Furthermore, dT_x has only non-negative real eigenvalues. The map dT_x plays the role of the differential of T and in fact it is the differential of the set valued extension $\partial^c \phi$ of the Borel map $T(x) = x - \nabla c^*(\phi(x)) = \partial^c \phi(x)\, ae.$, in the sense that for almost every x one has

$$\sup_{y \in \partial^c \phi(x+u)} |y - T(x) - dT_x(u)| = o(u).$$

We fix an $x \in \mathbf{R}^d$ where (5.14) holds. Setting

$$\alpha(t) := F\Big(\rho_0(x)/ \det(I + t(dT_x - \mathbf{id}))\Big)\ \det(I + t(dT_x - \mathbf{id}))/\rho_0(x)$$

for $t \in [0, 1]$ and using (5.14) we have

$$F(\rho_1(T(x))/\rho_1(T(x)) - F(\rho_0(x))/\rho_0(x) = \alpha(1) - \alpha(0). \tag{5.15}$$

Since the matrices $\mathbf{id}$ and $dT_x - \mathbf{id}$ commute and $\mathbf{id} + t(dT_x - \mathbf{id})$ has only non-negative real eigenvalues, we can combine the $1/d$ concavity of the determinant (on triangular matrices with non-negative eigenvalues) with the condition (H5) to conclude that α is convex. Thus

$$\alpha(1) - \alpha(0) \geq \alpha'(0) \tag{5.16}$$

Combining (5.15) and (5.16) we have for ρ_0 almost every x,

$$F(\rho_1(T(x))/\rho_1(T(x)) - F(\rho_0(x))/\rho_0(x) \geq -\mathrm{tr}(dT_x - I)A(\rho_0(x))/\rho_0(x).$$

Integrating with respect to ρ_0 we get

$$H^F(\rho_1) - H^F(\rho_0) \geq -\int_{\mathbf{R}^d} \mathrm{tr}(dT_x - I)A(\rho_0(x))\, dx.$$

The conclusion follows from the fact that $\mathrm{tr}(dT - I)$ is dominated by the distributional divergence of $T - I$. This fact is proved in [14]: for non-negative compactly supported test function g we have

$$\int_{\mathbf{R}^d} \mathrm{tr}(dT - I)g \leq -\int_{\mathbf{R}^d} (T - I).\nabla g.$$

This achieves the proof for c and c^* in C^2.

To complete the proof of the proposition in the case of non-smooth cost functions, we approximate c by a sequence $c_n \in C^2(\mathbf{R}^d)$ of strictly convex functions converging to c in $C^1_{loc}(\mathbf{R}^d)$, and such that $c_n^* \in C^2(\mathbf{R}^d)$ is strictly convex. Let Ω be some open bounded set containing the support of ρ_0 and ρ_1. By the above

$$\int_\Omega \Big(F(\rho_1) - F(\rho_0)\Big) dx \geq \int_\Omega (T_n - \mathbf{id}) \cdot \nabla(A(\rho_0)) dx, \tag{5.17}$$

where T_n is the c_n–optimal map such that $T_{n\#}\rho_0 = \rho_1$. In view of Lemma 5.2, $\{T_n\}_{n=1}^\infty$ converges to T in $L^1(\Omega, \rho_0)^d$, and so, letting n tends to $+\infty$ in (5.17) concludes the proof. QED

Lemma 5.2 *Let $c, c_n \in C^1(\mathbf{R}^d)$ be strictly convex functions, such that $\{c_n\}_{n=1}^\infty$ converges to c in $C^1_{loc}(\mathbf{R}^d)$. Assume that $\Omega \subset \mathbf{R}^d$ is a bounded set. Let $\rho_0, \rho_1 \in \mathcal{P}^a(\Omega)$, and let T_n (resp. T) be the unique c_n-optimal map (resp. c-optimal map) that pushes forward ρ_0 to ρ_1. Then, $\{T_n\}_{n=1}^\infty$ converges to T in $L^2(\Omega, \rho_0)^d$.*

Proof: The existence of T, T_n is obtained in [16, 17] and moreover, the measures $\gamma_n := (\mathbf{id} \times T_n)_\# \rho_0$ and $\bar{\gamma} := (\mathbf{id} \times T)_\# \rho_0$ are the unique minimizers of the functionals J_n and J over $\Gamma(\rho_0, \rho_1)$; here

$$J_n(\gamma) := \int_{\mathbf{R}^d \times \mathbf{R}^d} c_n(x - y) d\gamma(x, y), \quad J(\gamma) := \int_{\mathbf{R}^d \times \mathbf{R}^d} c(x - y) d\gamma(x, y).$$

If $\{\gamma_{n_i}\}_{i=1}^\infty$ is a subsequence of $\{\gamma_n\}_{n=1}^\infty$ that converges weak-$*$ to $\tilde{\gamma}$, then $\tilde{\gamma} \in \Gamma(\rho_0, \rho_1)$ and $\tilde{\gamma}$ minimizes J over $\Gamma(\rho_0, \rho_1)$. By the strict convexity of c, $\inf_{\Gamma(\rho_0, \rho_1)} J$ admits a unique minimizer (see [16]), and so $\bar{\gamma} = \tilde{\gamma}$. The sequence $\{\gamma_n\}_{n=1}^\infty$ being precompact for the weak-$*$ topology and every of its subsequences converging weak-$*$ to $\bar{\gamma}$, we deduce that in fact the whole sequence $\{\gamma_n\}_{n=1}^\infty$ converges weak-$*$ to $\bar{\gamma}$. Now set

$$F_1(x, y) := |y|^2, \quad F_2(x, y) := T(x) \cdot y,$$

for $x, y \in \Omega$. Straightforward computations show that

$$\int_\Omega |T_n(x) - T(x)|^2 \rho_0(x) dx = \int_{\mathbf{R}^d \times \mathbf{R}^d} (F_1 - F_2) d\gamma_n + \int_{\mathbf{R}^d \times \mathbf{R}^d} F_1 d\bar{\gamma} - \int_{\mathbf{R}^d \times \mathbf{R}^d} F_2 d\gamma_n. \tag{5.18}$$

We let n go to $+\infty$ in (5.18), use the fact that $\{\gamma_n\}_{n=1}^\infty$ converges weak-$*$ to $\bar{\gamma}$ and that $F_1 = F_2$ $\bar{\gamma}$ almost everywhere to deduce that the expression in the right hand side of (5.18) tends to 0 as n tends $+\infty$. This proves that $\{T_n\}_{n=1}^\infty$ converges to T in $L^2(\Omega, \rho_0)^d$. QED

6 Appendix B: An elementary approximation of density functions

Throughout this section, we assume that $\Omega \subset \mathbf{R}^d$ is nonempty, open, and convex, that F satisfies (H4-H5). We assume that $V \in C^1(\mathbf{R}^d)$ is convex and that there exists $\rho_\infty \in \mathcal{P}^a(\Omega)$ such that $F'(\rho_\infty) + V = 0$ on Ω. Eventually, we assume that $F(\rho_\infty) + \rho_\infty V \in L^1(\Omega)$ so that (1.12) gives that $H_V^F(\rho)$ is well defined for density functions which are absolutely continuous with respect to ρ_∞.

Lemma 6.1 *Assume that $\rho \in \mathcal{P}^a(\Omega) \cap C(\Omega)$, that $\rho > 0$ on Ω, and that $\{p_n\}_{n=1}^\infty$ is a non-decreasing sequence in $[0,1]$ converging to 1 as n tends to $+\infty$. Let $\{\Omega_n\}_{n=1}^\infty$ be a sequence of open, convex, bounded subsets of $\mathbf{R}^d$ such that $\bar{\Omega}_n \subset \Omega_{n+1}$, and $\cup_{n=1}^\infty \Omega_n = \Omega$. Then, there exists a sequence of positive functions $\{\rho_n\}_{n=1}^\infty \subset L^1(\Omega_n) \cap C(\Omega_n)$ converging to ρ in $L^1(\Omega)$ and satisfying the following properties:*

(i) For n large enough, $0 < \inf_{\bar{\Omega}_n} \rho_n$, $\rho_n = \rho$ on $\Omega_n \setminus \Omega_1$ and, $\int_\Omega \rho_n dx = p_n$.

(ii) We have that $||\rho_n - \rho||_{W^{1,\infty}(\Omega_1)}$ tends to 0 as n tends to $+\infty$.

(iii) $H_V^F(\rho_n) = H_V^F(\rho) + o(1)$.

(iv) Furthermore, if $\rho \in C^1(\Omega)$, we can choose ρ_n such that $\rho_n + \frac{1}{\rho_n} \in C^1(\bar{\Omega}_n)$.

Proof: Because Ω is nonempty, relabelling $\{\Omega_n\}_{n=1}^\infty$ if necessary, we may assume that Ω_1 is nonempty. We next choose a function $\varphi \in C_c^\infty(\Omega_1)$ compactly supported inside Ω_1 that is not identically 0 and such that $0 \le \varphi$. Denote by $\chi_{\bar{\Omega}_n}$ the characteristic function of Ω_n and let

$$\rho_n := \rho\chi_{\bar{\Omega}_n} + r_n\varphi, \qquad \text{where } r_n := \frac{p_n - \int_{\Omega_n} \rho dx}{\int_{\Omega_1} \varphi dx}.$$

Note that $\int_\Omega \rho_n dx = p_n$. Because $\varphi \in C_c^\infty(\Omega_1)$, we have that $\rho_n = \rho$ on $\Omega_n \setminus \Omega_1$. Since Ω_n is bounded and $\{r_n\}_{n=1}^\infty$ converges to 0 as n tends to $+\infty$, it is apparent that $0 < \inf_{\bar{\Omega}_n} \rho_n$ for n large enough. This proves (i). Next, (ii) is a direct consequence of the fact that $\rho_n - \rho = r_n\varphi$.

To avoid trivialities, we assume that $H_V^F(\rho)$ is finite. By (1.12), we obtain that $F(\rho) + \rho V \in L^1(\Omega)$. We use again the fact φ is supported inside Ω_1 to deduce that

$$H_V^F(\rho_n) = H_V^F(\rho) + \int_{\Omega_1} (F((1 + r_n\varphi)\rho) - F(\rho) + r_n\rho V\varphi) dx - \int_{\Omega_n^c} (F(\rho) + \rho V) dx. \tag{6.19}$$

The first integrand on the right handside of (6.19) tends to 0 uniformly on Ω_1 as n tends to $+\infty$. The Lebesgue dominated convergence theorem gives that the second integral on the right handside of (6.19) tends to 0 as n tends to $+\infty$. We obtain (iii). The proof of (iv) is easy. QED

Note added in proof. Closely related papers have appeared since the achievements of the present work. We mention for instance:

- M. Agueh, N. Ghoussoub and X. Kang. "Geometric inequalities via a duality between certain quasilinear PDEs and Fokker-Planck equations."

- M. Agueh, N. Ghoussoub and X. Kang. "The mother of most Gaussian and Euclidean inequalities."

- D. Cordero-Erausquin, B. Nazaret and C. Villani. A mass transportation approach to sharp Sobolev and Gagliardo-Nirenberg inequalities.

References

[1] M. Agueh. PhD dissertation. School of Mathematics, Georgia Institute of Technology, 2002.

[2] H.W. Alt and S. Luckhaus. Quasilinear elliptic-parabolic differential equations. *Math. Z.* **183** (1983), 311–341.

[3] A. Arnold, P. Markovich, G. Toscani and A. Unterreiter. On convex Sobolev inequalities and the rate of convergence to equilibrium for Fokker-Planck type equations. To appear in *Commun. PDE.*

[4] D. Bakry & M. Emery. Diffusions hypercontractives. in *Sem. Proba. XIX*, LNM 1123, Springer 1985, pp. 177–206.

[5] G. Blower The Gaussian isoperimetric inequality and transportation. Preprint (1999).

[6] S. Bobkov and F. Götze. Exponential integrability and transportation cost related to logarithmic Sobolev inequalities. *J. Funct. Anal.* **163** (1999), 1–28.

[7] S. Bobkov and M. Ledoux. From Brunn-Minkowski to Brascamp-Lieb and to logarithmic Sobolev inequalities. *Geom. Funct. Anal.* **10** (2000), 1028–1052.

[8] Y. Brenier. Décomposition polaire et réarrangement monotone des champs de vecteurs. *C.R. Acad. Sci. Paris Sér. I Math.* **305** (1987), 805–808.

[9] Y. Brenier. Polar factorization and monotone rearrangement of vector-valued functions. *Comm. Pure Appl. Math.* **44** (1991), 375–417.

[10] L. Caffarelli. Allocation maps with general cost functions. *Partial Differential Equations and Applications*, Lecture Notes in Pure and Applied Math., **177** (1996), 29–35.

steepest

[11] J.A. Carrillo, A. Juengel, P. Markovich, G. Toscani and A. Unterreiter. Entropy production methods for degenerate parabolic problems and generalized Sobolev inequalities. To appear in *Monatschefte für Math.*

[12] J. A. Carrillo, R. J. McCann, C. Villani. Kinetic equilibration rates for granular media and related equations via entropy dissipation and mass transportation estimates. (Preprint).

[13] D. Cordero-Erausquin. Some applications of mass transport to Gaussian type inequalities. *Arch. Rational Mech. Anal.* **161** (2002), 257–269.

[14] D. Cordero-Erausquin. Non-smooth differential properties of optimal transport. *Same collection.*

[15] W. Gangbo and C. Houdré. *Unpublished Research Notes* (1998).

[16] W. Gangbo and R.J. McCann. Optimal maps in Monge's mass transport problem. *C.R. Acad. Sci. Paris Sér. I Math.* **321** (1995), 1653–1658.

[17] W. Gangbo and R.J. McCann. The geometry of optimal transportation. *Acta Math.*, **177** (1996), 113–16.

[18] Del Pino, M., and Dolbeault, J. Best constants for Gagliardo-Nirenberg inequalities and application to nonlinear diffusions. To appear in *J. Math. Pures Appl.*

[19] L. Gross. Logarithmic Sobolev inequalities. *Amer. J. Math.* **97** (1975), 1061–1083.

[20] R. Jordan, D. Kinderleher and F. Otto. The variational formulation of the Fokker–Planck equation. *SIAM Jour. Math Anal.*, **29** (1998), 1–17.

[21] J. Kacur. On a solution of degenerate elliptic-parabolic systems in Orlicz-Sobolev spaces I. *Math. Z.*, **203** (1990), 153–171.

[22] J. Kacur. On a solution of degenerate elliptic-parabolic systems in Orlicz-Sobolev spaces II. *Math. Z.*, **203** (1990), 569–579.

[23] K Marton. A measure concentration inequality for contracting Markov chains. *Geom. Funct. Anal.* **6** (1996), 556–571.

[24] B. Maurey. Some deviation inequalities. *Geom. Funct. Anal.* **1** (1991), 188–197.

[25] R.J. McCann. A convexity principle for interacting gases. *Adv. Math.* **128** (1997), 153–179.

[26] F. Otto. Doubly degenerate diffusion equations as steepest descent. Preprint (1996). http://riemann.iam.uni-bonn.de/˜otto/publications.html .

[27] F. Otto. The geometry of dissipative evolution equation: the porous medium equation. *Comm. Partial Differential Equations* **26** (2001), 101–174.

[28] F. Otto and C. Villani. Generalization of an inequality by Talagrand and links with the logarithmic Sobolev inequality. *J. Funct. Anal.* **173**, (2000) 361–400.

[29] S.T. Rachev and L. Rüschendorf. *Mass Transportation Problems*; Vol. I & II Probability and its Applications. Springer, 1998.

[30] D.A. Stariolo. The Langevin and Fokker-Planck equations in the framework of a generalized statistical mechanics. *Physics Letters A* 185 (1994), 262–264.

[31] M. Talagrand. Transportation cost for Gaussian and other product measures. *Geom. Funct. Anal.* 6 (1996), 587–600.

[32] C. Tsallis. Possible generalization of Boltzmann-Gibbs statistics. *Journal of Statistical Physics* 52 (1998), 479–487.

Dario Cordero-Erausquin
Laboratoire d'Analyse et de Mathématiques Appliquées
CNRS UMR 8050 (Universités de Marne-la-Vallée et de Paris XII)
Université de Marne-la-vallée
77454 Marne la Vallée Cedex 2, France
cordero@math.univ-mlv.fr

Wilfrid Gangbo
School of Mathematics, Georgia Institute of Technology
Atlanta, GA 30332, USA
gangbo@math.gatech.edu

Christian Houdré
Laboratoire d'Analyse et de Mathématiques Appliquées
CNRS UMR 8050 (Universités de Marne-la-Vallée et de Paris XII)
Faculté des Sciences et Technologies de Paris XII
94010 Créteil Cedex, France
and
School of Mathematics, Georgia Institute of Technology
Atlanta, GA 30332, USA
houdre@math.gatech.edu