Dimensional improvements of the logarithmic Sobolev, Talagrand and Brascamp-Lieb inequalities

François Bolley, Ivan Gentil, Arnaud Guillin

▶ **To cite this version:**

François Bolley, Ivan Gentil, Arnaud Guillin. Dimensional improvements of the logarithmic Sobolev, Talagrand and Brascamp-Lieb inequalities. Annals of Probability, 2018, 46 (1), pp.261-301. 10.1214/17-AOP1184 . hal-01171361v4

HAL Id: hal-01171361
https://hal.science/hal-01171361v4

Submitted on 12 Mar 2017

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Dimensional improvements of the logarithmic Sobolev, Talagrand and Brascamp-Lieb inequalities

François Bolley*, Ivan Gentil† and Arnaud Guillin‡

March 12, 2017

Abstract

In this work we consider dimensional improvements of the logarithmic Sobolev, Talagrand and Brascamp-Lieb inequalities. For this we use optimal transport methods and the Borell-Brascamp-Lieb inequality. These refinements can be written as a deficit in the classical inequalities. They have the right scale with respect to the dimension. They lead to sharpened concentration properties as well as refined contraction bounds, convergence to equilibrium and short time behavior for the laws of solutions to stochastic differential equations.

Key words: Logarithmic Sobolev inequality, Talagrand inequality, Brascamp-Lieb inequality, Fokker-Planck equations, optimal transport.

Introduction

We shall be concerned with diverse ways of measuring and bounding the distance between probability measures, and the links between them. We will focus on three main inequalities that we now describe.

- A probability measure μ on $\mathbb{R}^n$ satisfies a logarithmic Sobolev inequality (in short LSI) with constant $R > 0$ (see [4] for instance) if for all probability measures ν in $\mathbb{R}^n$, absolutely continuous with respect to μ,

$$H(\nu|\mu) \leq \frac{1}{2R}\, I(\nu|\mu). \tag{1}$$

  Here H and I are the relative entropy and the Fisher information, defined for $f = \frac{d\nu}{d\mu}$ by

$$H(\nu|\mu) = \mathrm{Ent}_\mu(f) = \int f \log f \, d\mu \qquad \text{and} \qquad I(\nu|\mu) = \int \frac{|\nabla f|^2}{f} d\mu. \tag{2}$$

  For I we assume that $\nabla f/f \in L^2(\nu)$.

- A probability measure μ in $\mathbb{R}^n$ satisfies a Talagrand transportation inequality [37] with constant $R > 0$ if for all ν absolutely continuous with respect to μ

$$W_2^2(\nu, \mu) \leq \frac{2}{R} H(\nu|\mu). \tag{3}$$

*Laboratoire de Probabilités et Modèles Aléatoires, Umr Cnrs 7599, Université Pierre et Marie Curie, Paris, France. francois.bolley@upmc.fr

†Univ Lyon, Université Claude Bernard Lyon 1, CNRS UMR 5208, Institut Camille Jordan, 43 blvd. du 11 novembre 1918, F-69622 Villeurbanne cedex, France. gentil@math.univ-lyon1.fr

‡Laboratoire de Mathématiques, Umr Cnrs 6620, Université Blaise Pascal, Clermont-Ferrand, France. guillin@math.univ-bpclermont.fr

Here W_2 is the Monge-Kantorovich-Wasserstein distance; it is defined for μ and ν in $P_2(\mathbb{R}^n)$ by

$$W_2(\mu,\nu) = \inf_{\pi} \left(\iint |y-x|^2 \, d\pi(x,y) \right)^{1/2}$$

where π runs over the set of (coupling) measures on $\mathbb{R}^n \times \mathbb{R}^n$ with respective marginals μ and ν. We let $P_2(\mathbb{R}^n)$ be the space of probability measures μ on $\mathbb{R}^n$ with finite second moment, that is, $\int |x|^2 d\mu(x) < +\infty$ (see [1], [39]).
By the Otto-Villani Theorem [36], the logarithmic Sobolev inequality (1) implies the Talagrand inequality (3) with the same constant (see also [7], [39, Chap. 22]).

- Let μ be a probability measure in $\mathbb{R}^n$ with density e^{-V} where V is a $\mathcal{C}^2$ and strictly convex function. Then the Brascamp-Lieb inequality asserts that for all smooth functions f,

$$\mathrm{Var}_\mu(f) \leq \int \nabla f \cdot \mathrm{Hess}(V)^{-1} \nabla f \, d\mu. \tag{4}$$

Here $\mathrm{Var}_\mu(f) = \int f^2 d\mu - (\int f d\mu)^2$ is the variance of f under the measure μ, see [4, Sect 4.9.1] for instance.

The standard Gaussian measure γ in $\mathbb{R}^n$ with density e^{-V} for $V(x) = |x|^2/2 + n\log(2\pi)/2$, satisfies the three inequalities (1), (3) with $R = 1$ and (4). In fact, in the Gaussian case, the Brascamp-Lieb inequality (4) can be obtained from (1) by linearization, namely by taking $\nu = f\mu$ with f close to 1. Let us note that in this case $\mathrm{Hess}(V) = \mathrm{Id}_n$, the Brascamp-Lieb inequality becomes exactly the Poincaré inequality. Moreover these inequalities are optimal for the Gaussian measure: by direct computation, equality holds in (1) and (3) for translations of γ, that is, for measures $\nu = \exp(a\cdot x - \frac{|a|^2}{2})\gamma$ with $a \in \mathbb{R}^n$; equality holds in (4) for $f(x) = b \cdot x$, $b \in \mathbb{R}^n$ (see [4, Chap. 4 and 5]).

Inequalities (1), (3) and (4) share the significant property of tensorisation, leading to possible constants R independent of the dimension of the space. In other words, if a probability measure μ satisfies one of these three inequalities with constant $R > 0$, then for any $N \in \mathbb{N}^*$, the product measure $\mu^N = \otimes^N \mu$ satisfies the same inequality with the same constant R. This can be interesting in applications to problems set in large or infinite dimensions.
However, for regularity or integrability arguments, one may need more precise forms capturing the precise dependence on the dimension. Such dimension dependent improvements have been observed in the Gaussian case. Namely, the dimensional improvement

$$H(\nu|\gamma) \leq \frac{1}{2} \int |x|^2 d\nu - \frac{n}{2} + \frac{n}{2} \log\left(1 + \frac{1}{n}\Big(I(\nu|\gamma) + n - \int |x|^2 d\nu\Big)\right) \tag{5}$$

of the logarithmic Sobolev inequality (1) has been obtained by D. Bakry and M. Ledoux [5] by self-improvement from the Euclidean logarithmic Sobolev inequality, or by semigroup arguments on the Euclidean heat semigroup (see also [4, Sect. 6.7.1] and the early work [15] by E. Carlen). The dimensional improvement

$$W_2^2(\nu,\gamma) \leq \int |x|^2 d\nu + n - 2n \exp\left(\int \frac{|x|^2}{2n} d\gamma - \frac{1}{2} - \frac{1}{n} H(\nu|\gamma) \right) \tag{6}$$

of the Talagrand inequality (3) has been derived in [3]; the argument is based on local hypercontractivity techniques on an associated Hamilton-Jacobi semigroup and fine properties of the heat semigroup. It has further been observed in [5] that linearizing (5) leads to the dimensional improvement

$$\mathrm{Var}_\gamma(f) \leq \int |\nabla f|^2 \, d\gamma - \frac{1}{2n}\Big(\int (|x|^2 - n) f d\gamma \Big)^2 \tag{7}$$

of the Brascamp-Lieb (or Poincaré) inequality (4) for the Gaussian measure (see also [4, Sect. 6.7.1]). On the other hand, by a spectral analysis of the Ornstein-Uhlenbeck semigroup, the bound

$$\mathrm{Var}_\gamma(f) \leq \frac{1}{2} \int |\nabla f|^2 \, d\gamma + \frac{1}{2} \Big| \int \nabla f d\gamma \Big|^2 \tag{8}$$

has been established in [30, Sect. 6.2]. By the Cauchy-Schwarz inequality, it improves upon (4). Naturally, both inequalities (7) and (8) are optimal, and equality holds for $f(x) = a \cdot x$; equality also holds for $f(x) = |x|^2$, in fact for the first two Hermite polynomials. The above proofs of (5), (6) and (8) are very specific to the Gaussian case and can not be extended to other measures.

These dimensional improvements can also be written as a deficit in the classical non dimensional versions (1), (3), (4) of the inequalities: namely, for the logarithmic Sobolev (LSI in short) and Talagrand (Tal in short) inequalities, lower bounds on the quantities

$$\delta_{LSI}(\nu|\mu) := \frac{1}{2} I(\nu|\mu) - R\, H(\nu|\mu) \qquad \text{and} \qquad \delta_{Tal}(\nu|\mu) := H(\nu|\mu) - \frac{R}{2}\, W_2^2(\nu,\mu).$$

The problem of dimensional refinements of standard functional inequalities has been recently considered in an intensive manner. Via the development of refined optimal transportation tools, beautiful results for the Gaussian isoperimetric inequality were obtained by Figalli-Maggi-Pratelli [25] (see also R. Eldan [20] or [23] for convex cones). Further recent results have been established on deficit in the logarithmic Sobolev inequality in the Gaussian case by Figalli-Maggi-Pratelli [26], Indrei-Marcon [32] and Bobkov & al [8]. In particular [8] rediscovers (5) and extends earlier results obtained in dimension one by Barthe-Kolesnikov [6] on the Talagrand deficit. Fathi-Indrei-Ledoux [22] also considers these deficits, particularly emphasizing the case where ν has additional properties, such as a Poincaré inequality ensuring a better constant in the logarithmic Sobolev inequality. Very recently D. Cordero-Erausquin [17] has studied refinements of the Talagrand and Brascamp-Lieb inequalities via optimal transport tools.
Let us also quote C. Villani [39, p. 605]:

> *There is no well-identified analog of Talagrand inequalities that would take advantage of the finiteness of the dimension to provide sharper concentration inequalities*

as a motivation to investigate further the problem. As we will see there are other striking applications of these dimensional refinements than sole concentration.
Finally recall that the so-called Bakry-Émery criterion (or Γ_2-criterion) ensures that the measure μ with density e^{-V} satisfies the logarithmic Sobolev inequality (1) and Talagrand inequality (3) as soon as the potential V satisfies $\mathrm{Hess}(V) \geq R\,\mathrm{Id}_n$ with $R > 0$, as symmetric matrices. One of the goals of this paper is to extend the above dimensional inequalities under this condition with $R > 0$ or only $\mathrm{Hess}(V) > 0$. For this we shall use multiple tools and we will compare our inequalities with other recent extensions. Applications to concentration inequalities and short and long time behaviour for the laws of solutions to stochastic differential equations are also given.

Plan of the paper and main results

Let μ be a probability measure on $\mathbb{R}^n$ with density e^{-V} where V is $\mathcal{C}^2$.

In Section 1, we propose a method based on the Borell-Brascamp-Lieb inequality to get dimensional logarithmic Sobolev inequalities in the spirit of the works [9, 11] by S. Bobkov and M. Ledoux. The method is based on a general convexity inequality given in Theorem 1.1. For instance, in Corollary 1.4 we shall prove the following : If $\mathrm{Hess}(V) \geq R\,\mathrm{Id}_n$ with $R > 0$, then

$$\mathrm{Ent}_\mu(f^2) \leq n(s - 1 - \log s) + \frac{1}{2R} \int \left|(1-s)\nabla V + 2s\frac{\nabla f}{f}\right|^2 f^2\, d\mu \tag{9}$$

for any $s > 0$ and any function f such that $\int f^2 d\mu = 1$. This improves upon the classical logarithmic Sobolev inequality (1) under the Bakry-Émery condition, which is recovered for $s = 1$.

In Section 2 (Theorem 2.1) we propose a dimensional Talagrand inequality through optimal transportation in the spirit of Barthe-Kolesnikov [6] and D. Cordero-Erausquin [16] or the recent [17] : If $\mathrm{Hess}(V) \geq R\,\mathrm{Id}_n$ with $R > 0$ then

$$\frac{R}{2} W_2^2(\mu,\nu) \leq \nu(V) - \mu(V) + n - n \exp\left[\frac{1}{n}\Big(\nu(V) - \mu(V) - H(\nu|\mu)\Big)\right] \tag{10}$$

for all $\nu \in P_2(\mathbb{R}^n)$. This bound implies the classical Talagrand inequality (3). Let us observe that, using the terminology of the Γ_2-condition, the associated Markov generator $L = \Delta - \nabla V \cdot \nabla$ does not satisfy

a $CD(R,n)$ curvature dimension condition, but only $CD(R,\infty)$. In particular the general dimensional log Sobolev or Talagrand inequalities, obtained on manifolds (see [4]) or on abstract measure spaces (as in [21]) do not hold. In Section 2.1 we show how the dimensional corrective term in our new Talagrand inequality enables to get sharp concentration inequalities.

Inspired by recent results on the equivalence between contraction and $CD(R,n)$ condition in abstract measure spaces (see [1, 21, 14]), in Section 3 we consider applications to refined dimensional contraction properties under $CD(R,\infty)$ (see Proposition 3.3 and Corollary 3.8); we shall see how the dimension improves the asymptotic behaviour for the laws of solutions to stochastic differential equations (in the spirit of [12, 13]). Again the generator $L = \Delta - \nabla V \cdot \nabla$ does not satisfy a $CD(R,n)$ condition, but only $CD(R,\infty)$. The key point here is to take advantage of the contribution of the diffusion term, which includes a dimensional term. We shall also see how the dimension influences the short time smoothing effect, through very simple arguments (see Proposition 3.1).

In section 4 we prove two kinds of dimensional Brascamp-Lieb inequalities, a first one by a L^2 argument, a second one by a linearization argument in the Borell-Brascamp-Lieb inequality. For instance, under the sole assumption $\mathrm{Hess}(V) > 0$, Theorem 4.3 states that

$$\mathrm{Var}_\mu(f) \leq \int \nabla f \cdot \mathrm{Hess}(V)^{-1} \nabla f \, d\mu - \int \frac{(f - \nabla f \cdot \mathrm{Hess}(V)^{-1} \nabla V)^2}{n + \nabla V \cdot \mathrm{Hess}(V)^{-1} \nabla V} d\mu \tag{11}$$

for any smooth function f such that $\int f d\mu = 0$. We shall discuss the optimality of our bounds and compare them with other very recent dimensional refinements of the Brascamp-Lieb inequality.

In the Gaussian case where $\mu = \gamma$, then the logarithmic Sobolev (9) (by optimising over s) and Talagrand (10) inequalities are exactly (5) and (6) respectively, while the Poincaré inequality (11) improves upon (7).

Notation: whenever there is no ambiguity we shall respectively use H, I, W_2, δ_{LSI} and δ_{Tal} for $H(\nu|\mu)$, $I(\nu|\mu), W_2(\nu,\mu), \delta_{LSI}(\nu|\mu)$ and $\delta_{Tal}(\nu|\mu)$. We shall sometimes let $\mathrm{Ent}_{dx}(f) = \int f \log f dx$ and $\mu(f) = \int f d\mu$ and use the same notation for an absolutely continuous measure with respect to Lebesgue measure, and its density.

1 Logarithmic Sobolev inequalities

The Prékopa-Leindler inequality is a reverse form of the Hölder inequality. Let F, G, H be non-negative measurable functions on $\mathbb{R}^n$ satisfying $\int F dx = \int G dx = 1$, and let $s, t \geq 0$ be fixed such that $t + s = 1$. Under the hypothesis

$$H(tx + sy) \geq F(x)^t G(y)^s \tag{12}$$

for any $x, y \in \mathbb{R}^n$, the Prékopa-Leindler inequality ensures that $\int H dx \geq 1$, see [39, Chap. 19] for instance. The Borell-Brascamp-Lieb inequality is a stronger and dimensional form of the Prékopa-Leindler inequality. Assume again $\int F dx = \int G dx = 1$ and in addition that F, G and H are positive; then the Borell-Brascamp-Lieb inequality asserts that $\int H dx \geq 1$ as soon as

$$H(tx + sy) \geq \left(t F(x)^{-1/n} + s G(y)^{-1/n} \right)^{-n} \tag{13}$$

for any $x, y \in \mathbb{R}^n$, instead of the stronger (12) (by convexity); see again [39].

The Prékopa-Leindler inequality in particular implies many geometrical and functional inequalities as logarithmic Sobolev and Brascamp-Lieb inequalities, as observed by S. Bobkov and M. Ledoux in [9, 11] (see also [28] for an application to the modified logarithmic Sobolev inequality). In the coming sections we shall see how the Borell-Brascamp-Lieb inequality implies dimensional form of these inequalities. Following S. Bobkov and M. Ledoux [9, 11] our proofs are based on Taylor expansions when $s \to 0$ or $F \to 0$.

1.1 A general convexity inequality via the Borell-Brascamp-Lieb inequality

Let us first state a general consequence of the Borell-Brascamp-Lieb inequality. It will lead to various dimensional logarithmic Sobolev inequalities.
In the sequel we let ψ^* be the Legendre transform of a function ψ on $\mathbb{R}^n$, defined for $y \in \mathbb{R}^n$ by

$$\psi^*(y) = \sup_{x \in \mathbb{R}^n} \{y \cdot x - \psi(x)\} \in (-\infty, +\infty].$$

If ψ is C^1 and strictly convex satisfying

$$\lim_{|x| \to +\infty} \frac{\psi(x)}{|x|} = +\infty,$$

then (see [38, Sect. 2.1.3 and 2.4.3] for instance) for all $x \in \mathbb{R}^n$, $\psi^*(x) \in \mathbb{R}$ and

$$\psi(x) = \nabla\psi(x) \cdot x - \psi^*(\nabla\psi(x)) \quad \text{and} \quad \nabla\psi^*(\nabla\psi(x)) = x. \tag{14}$$

Theorem 1.1 (Convexity inequality) *Let g, W be $\mathcal{C}^1$ and positive functions on $\mathbb{R}^n$ satsifying the normalization condition $\int g^{-n} dx = \int W^{-n} dx = 1$. Assume moreover that there exists a constant $C > 0$ such that for all $x \in \mathbb{R}^n$,*

$$W(x) \geq \frac{1}{C} \frac{|x|^2}{2}, \tag{H1}$$

$$\frac{1}{C}(|x|^2 + 1) \leq g(x) \leq C(|x|^2 + 1) \quad \text{and} \quad |\nabla g(x)| \leq C(|x| + 1). \tag{H2}$$

Then

$$\int \frac{W^*(\nabla g)}{g^{n+1}} dx \geq 0. \tag{15}$$

If W is a $\mathcal{C}^1$ positive and strictly convex function which satisfies (H1) and $\int W^{-n} dx = 1$, then (15) is an equality for $g = W$.

The same statement can be proved for a larger class of functions g and W. We only state this result with these restrictive hypotheses for simplicity reasons, as this setting will be sufficient for our main application. The rigorous proof is postponed to the Appendix A. The idea is to perform a Taylor expansion of the Borell-Brascamp-Lieb inequality (13) when $s = 1 - t$ goes to 0. Indeed, let $F = g^{-n}$ and $G = W^{-n}$ in (13), hence satisfying $\int F dx = \int G dx = 1$. Then the function H_t defined by

$$H_t(z)^{-1/n} = \inf_{h \in \mathbb{R}^n} \left\{ t g\left(z + \frac{s}{t} h\right) + s W(z - h) \right\} \tag{16}$$

for $z \in \mathbb{R}^n$ satisfies $\int H_t dx \geq 1$. The first-order Taylor expansion of H_t, when $s = 1 - t$ goes to 0, gives

$$H_t(z) = g(z)^{-n} - s\, n\, g(z)^{-n-1}\big(z \cdot \nabla g(z) - g(z)\big) + s\, n\, \frac{W^*(\nabla g(z))}{g^{n+1}(z)} + o(s).$$

Since

$$\int g^{-n-1}(z \cdot \nabla g - g)\, dx = 0$$

by integration by parts, the Taylor expansion of $\int H_t dx \geq 1$ implies the inequality (15).

Applications of Theorem 1.1 are described in the coming two sections. They are based on the following observation. Let V be a given function and let $W = e^{\frac{V}{n}}$. Then, from the convexity of the exponential function, for any $a \in \mathbb{R}$ and $y \in \mathbb{R}^n$,

$$W^*(y) \leq \frac{1}{n} e^a V^*(n e^{-a} y) + (a - 1) e^a.$$

Combined with Theorem 1.1, this gives the following corollary which is the main tool in our applications:

Corollary 1.2 *Under the hypotheses of Theorem 1.1, let $V = n \log W$. Then for any function a,*

$$\int \frac{1}{g^{n+1}} \big(e^a V^*(n e^{-a} \nabla g) + n(a - 1) e^a\big)\, dx \geq 0. \tag{17}$$

1.2 Euclidean logarithmic Sobolev inequalities

As a warm up, let us first see how to quickly recover the classical Euclidean logarithmic Sobolev inequality, using (17). Let $C : \mathbb{R}^n \to \mathbb{R}^+$ be a strictly convex function such that $\int e^{-C} dx < +\infty$, and let us apply (17) with $V = C + \beta$ and $W = e^{V/n}$; here $\beta = \log \int e^{-C} dx$ so that $\int e^{-V} dx = 1$. Since V is convex and $\int e^{-V} dx < +\infty$, it is classical that V grows at least linearly at infinity, so that W satisfies hypothesis (H1) .
Then let $p > 1$. Let also f be a $\mathcal{C}^1$ positive function such that $\int f^p dx = 1$ and $g = f^{-p/n}$ satisfies (H2), and let $a = -\frac{p}{n} \log f + u$ where u is a real constant. Then $V^* = C^* - \beta$ and (17) can be written as

$$\forall u \in \mathbb{R}, \quad \int f^p \log(f^p)\, dx \leq n(u-1) - \beta + \int C^*\left(-pe^{-u}\frac{\nabla f}{f}\right) f^p\, dx. \tag{18}$$

We can optimise over u in $\mathbb{R}$ in the following case. Suppose that there exists $q > 1$ such that C is q-homogeneous, that is, $C(\lambda x) = \lambda^q C(x)$ for any $\lambda \geq 0$ and x in $\mathbb{R}^n$. Then C^* is p-homogeneous with $1/p+1/q = 1$, and in particular above $C^*(-pe^{-u}\nabla f/f) = p^p e^{-pu} f^{-p} C^*(-\nabla f)$. Thus inequality (18) gives

$$\int f^p \log(f^p)\, dx \leq n(u-1) - \beta + e^{-pu} p^p \int C^*(-\nabla f)\, dx \tag{19}$$

for any function f such that $\int f^p dx = 1$ and $f^{-n/p}$ satisfies (H2). Now, let f be a $\mathcal{C}^1$ non negative and compactly supported function and for $\varepsilon > 0$ let $f_\varepsilon(x) = C_\varepsilon(\varepsilon(|x|^2+1)^{-n/p} + f)$, where C_ε is such that $\int (f_\varepsilon)^p dx = 1$. The function $f_\varepsilon^{-n/p}$ satisfies (H2) for any ε. Taking the limit when ε goes to 0, inequality (19) then holds for any $\mathcal{C}^1$ non negative and compactly supported function f such that $\int f^p dx = 1$.
For the optimal $u = p^{-1} \log\left(p^{p+1} \int C^*(-\nabla f) dx/n\right)$, the bound (19) leads to

$$\int f^p \log(f^p)\, dx \leq \frac{n}{p} \log\left(\frac{p^{p+1}}{n e^{p-1}} \frac{\int C^*(-\nabla f) dx}{(\int e^{-C} dx)^{p/n}}\right)$$

for any $\mathcal{C}^1$ non negative and compactly supported function f such that $\int f^p dx = 1$. Of course, the inequality can be extended to a larger class of functions f. Hence, we recover the optimal L^p-Euclidean log Sobolev inequality proved in [19, 27] and in particular, setting $C(x) = |x|^2/2$ and $p = q = 2$, the classical inequality

$$\int f^2 \log(f^2)\, dx \leq \frac{n}{2} \log\left(\frac{2}{n\pi e} \int |\nabla f|^2 dx\right).$$

1.3 Dimensional logarithmic Sobolev inequalities

In this section we consider a probability measure μ with density e^{-V} and the function $W = e^{V/n}$, and a positive function f such that $\int f^2\, d\mu = 1$. We assume again that V is convex ; then $W = e^{V/n}$ satisfies hypothesis (H1) since $\int e^{-V} dx = 1$.
Corollary 1.2 applied with $g = e^{V/n} f^{-2/n}$ (assuming that g satisfies hypothesis (H2)) and $a = \frac{V}{n} - \frac{2}{n} \log f + u$ with $u \in \mathbb{R}$ gives

$$\int \left(V^*\left(e^{-u}\nabla V - 2e^{-u}\frac{\nabla f}{f}\right) + V - \log(f^2) + n(u-1)\right) f^2 e^{-V} dx \geq 0.$$

Corollary 1.3 *Let $d\mu(x) = e^{-V(x)} dx$ be a probability measure with V a convex function and let f be a $\mathcal{C}^1$ positive function such that $\int f^2\, d\mu = 1$ and such that $g = e^{V/n} f^{-2/n}$ satisfies hypothesis (H2). Then for any $s > 0$*

$$\mathrm{Ent}_\mu(f^2) \leq \int \left[V^*\left(s\nabla V - 2s\frac{\nabla f}{f}\right) + V\right] f^2 d\mu - n(1 + \log s). \tag{20}$$

For $s = 1$, inequality (20) simplifies as

$$\mathrm{Ent}_\mu(f^2) \leq \int \left[V^*\left(\nabla V - 2\frac{\nabla f}{f}\right) + V - n\right] f^2 d\mu, \qquad \int f^2 d\mu = 1.$$

In particular, for $V = \frac{|x|^2}{2} + \frac{n}{2}\log(2\pi)$, then μ is the standard Gaussian measure γ and we recover the Gaussian logarithmic Sobolev inequality of L. Gross,

$$\mathrm{Ent}_\gamma(f^2) \leq 2\int |\nabla f|^2 d\gamma, \qquad \int f^2\, d\gamma = 1.$$

More generally, let V be a strictly convex function on $\mathbb{R}^n$. Then inequality (20) with $s = 1$, by (14) and integration by parts, leads to the modified logarithmic Sobolev inequality

$$\mathrm{Ent}_\mu(f^2) \leq \int \left[V^*\Big(\nabla V - 2\frac{\nabla f}{f}\Big) + 2x\cdot\frac{\nabla f}{f} - V^*(\nabla V)\right] f^2 d\mu, \qquad \int f^2\, d\mu = 1$$

proved by the second author in [28].

Assuming *uniform convexity* on V we now optimise over the parameter $s > 0$ in Corollary 1.3, to obtain dimensional logarithmic Sobolev inequalities. Suppose that V is $\mathcal{C}^2$ with $\mathrm{Hess}(V) \geq R\,\mathrm{Id}_n$ for $R > 0$. Then, for their inverse matrices, $\mathrm{Hess}(V^*) \leq R^{-1}\,\mathrm{Id}_n$ on $\mathbb{R}^n$. Hence, for any z and by the Taylor expansion at point $\nabla V(x)$,

$$\begin{aligned} V^*(z) + V(x) &\leq V^*(\nabla V(x)) + \nabla V^*(\nabla V(x))\cdot(z - \nabla V(x)) + \frac{1}{2R}|z - \nabla V(x)|^2 + V(x) \\ &= x\cdot z + \frac{1}{2R}|z - \nabla V(x)|^2. \end{aligned}$$

Here we use the relations (14). For $z = s\nabla V - 2s\frac{\nabla f}{f}$ at point x, and by (20), this leads to

$$\mathrm{Ent}_\mu(f^2) \leq -n(1 + \log s) + s\int x\cdot\Big(\nabla V - 2\frac{\nabla f}{f}\Big) f^2 d\mu + \frac{1}{2R}\int \Big|\Big(s\nabla V - 2s\frac{\nabla f}{f}\Big) - \nabla V\Big|^2 f^2 d\mu.$$

By integration by parts and extending to compactly supported functions, as for (19), we finally obtain:

Corollary 1.4 (Dimensional LSI under Γ_2-condition) *Let μ be a probability measure with density e^{-V} where V is $\mathcal{C}^2$ with $\mathrm{Hess}(V) \geq R\,\mathrm{Id}_n$ for $R > 0$. Then*

$$\mathrm{Ent}_\mu(f^2) \leq n(s - 1 - \log s) + \frac{1}{2R}\int \Big|(1-s)\nabla V + 2s\frac{\nabla f}{f}\Big|^2 f^2\, d\mu \tag{21}$$

for any $s > 0$ and any $\mathcal{C}^1$, non-negative and compactly supported function f such that $\int f^2 d\mu = 1$.

The bound can of course be extended to other classes of functions f.
When $s = 1$, we recover the classical logarithmic Sobolev inequality (1) under the Bakry-Émery condition. Let us observe that the right-hand side in (21) can be expanded as $-n\log s$ plus a second order polynomial in s. Hence it admits a unique minimiser $s > 0$, which solves a second order polynomial. The obtained expression is not appealing and we prefer to omit it. In the Gaussian case where $\mu = \gamma$, then the optimisation over s gets even simpler and leads again to the dimensional Gaussian log Sobolev inequality (5).
Moreover, for a general V and as in (33) or (23) below for the Talagrand inequality, the bound (21) can be written as a (not either appealing) deficit in the log Sobolev inequality.

We will see in Section 3.1 that (21) leads to new and sharp short time smoothing on the entropy of solutions to an associated Fokker-Planck equation.

2 Talagrand inequalities

The main result of this section is

Theorem 2.1 (Dimensional Talagrand inequality) *Let μ be a probability measure in $P_2(\mathbb{R}^n)$ with density e^{-V} where V is a $\mathcal{C}^2$ function satisfying $\mathrm{Hess}(V) \geq R\,\mathrm{Id}_n$ with $R > 0$. Then for all $\nu \in P_2(\mathbb{R}^n)$*

$$\frac{R}{2}W_2^2(\mu,\nu) \leq \nu(V) - \mu(V) + n - n\exp\Big[\frac{1}{n}\Big(\nu(V) - \mu(V) - H(\nu|\mu)\Big)\Big]. \tag{22}$$

In other words, if $\mathrm{Hess}(V) \geq R\,\mathrm{Id}_n$, then $\nu(V) - \mu(V) - \frac{R}{2}W_2^2(\nu,\mu) > -n$ and

$$\delta_{Tal}(\nu|\mu) \geq \max\Big\{\delta_n\Big(H(\nu|\mu) + \mu(V) - \nu(V))\Big), \Lambda_n\Big(\nu(V) - \mu(V) - \frac{R}{2}W_2^2(\nu,\mu)\Big)\Big\}. \tag{23}$$

Here δ_n and Λ_n are the positive functions respectively defined by $\delta_n(x) = n[e^{-x/n} - 1 + x/n], x \in \mathbb{R}$ and $\Lambda_n(x) = x - n\log(1 + x/n), x > -n$.

The function $\delta_1(x) = e^{-x} - 1 + x$ is positive and convex. It is moreover decreasing on $\mathbb{R}^-$ and increasing on $\mathbb{R}^+$. By a direct computation, $\delta_1(x)$ is bounded from below by $x^2/2$ if $x \leq 0$, x^2/e if $0 \leq x \leq 1$ and x/e if $x > 1$; hence always by $\frac{1}{e}\min(|x|, x^2)$. Then for any $x \in \mathbb{R}$, $\delta_n(x) \geq \frac{1}{e}\min(|x|, \frac{x^2}{n})$.

Since $e^u \geq 1 + u$, the bound (22) implies the classical Talagrand inequality (3) under the condition $\mathrm{Hess}(V) \geq R\,\mathrm{Id}_n$. When μ is the standard Gaussian measure γ on $\mathbb{R}^n$, then $R = 1$ and we recover the dimensional Talagrand inequality (6).

Under a moment condition Theorem 2.1 simplifies as follows:

Corollary 2.2 *Following the same assumptions as in Theorem 2.1, for all ν in $P_2(\mathbb{R}^n)$ with $\nu(V) \leq \mu(V)$,*

$$\delta_{Tal}(\nu|\mu) \geq \delta_n(H(\nu|\mu)) \geq \frac{1}{e}\min\left(H(\nu|\mu), \frac{H(\nu|\mu)^2}{n}\right). \tag{24}$$

Theorem 2.1 will be deduced from the following dimensional HWI-type inequality, applied with $f = 1$ and $\nu = g\mu$. The HWI inequality bounds from above the entropy by the Wasserstein distance and the Fisher information (defined in (2)), in the form

$$H(\nu|\mu) \leq W_2(\nu,\mu)\sqrt{I(\nu|\mu)} - \frac{R}{2}W_2^2(\mu,\nu) \tag{25}$$

for all ν. It has been introduced in [36] and proved in [36] and [16] under the Bakry-Émery condition $\mathrm{Hess}(V) \geq R\,\mathrm{Id}_n, R \in \mathbb{R}$.

Theorem 2.3 (Dimensional HWI inequality) *Let μ be a probability measure on $\mathbb{R}^n$ with density e^{-V} where V is a $\mathcal{C}^2$ function satisfying $\mathrm{Hess}(V) \geq R\,\mathrm{Id}_n$ with $R \in \mathbb{R}$. Let also f, g be smooth functions such that $f\mu$ and $g\mu$ belong to $P_2(\mathbb{R}^n)$. Then*

$$n\exp\left[\frac{1}{n}\Big(H(f\mu|\mu) - H(g\mu|\mu) + \mu(gV) - \mu(fV)\Big)\right] - n \\ \leq \mu(gV) - \mu(fV) + W_2(f\mu, g\mu)\sqrt{I(f\mu|\mu)} - \frac{R}{2}W_2^2(f\mu, g\mu).$$

For $g = 1$ and $\nu = f\mu$, this bound can be written as the dimensional HWI inequality

$$n\exp\left[\frac{1}{n}\Big(H(\nu|\mu) + \mu(V) - \nu(V)\Big)\right] - n \leq \mu(V) - \nu(V) + W_2(\mu,\nu)\sqrt{I(\nu|\mu)} - \frac{R}{2}W_2^2(\mu,\nu). \tag{26}$$

As in (23) for the Talagrand inequality, this can equivalently be written as a deficit in the HWI inequality. It is classical that the HWI inequality (25) implies the logarithmic Sobolev inequality (1) (see [36] for instance). Likewise, from (26), one can obtain a dimension dependent logarithmic Sobolev inequality. We refer to Section 2.5 for further details.

The proof of Theorem 2.3 will be given in Section 2.4.

2.1 An application to concentration

Let us quickly revisit K. Marton's argument for concentration via Talagrand's inequality (as in [39, Chap. 22] for instance) and see how the refined inequality (22) in Theorem 2.1 gives sharpened information for large deviations.

Let $d\mu = e^{-V}dx$ satisfy inequality (22). Let also $A \subset \mathbb{R}^n$, $r > 0$ and $A_r = \{x;\, \forall y \in A,\, |y - x| > r\}$. Let finally $\mu_A = \frac{1_A}{\mu(A)}\mu$ and $\mu_{A_r} = \frac{1_{A_r}}{\mu(A_r)}\mu$ be the restrictions of μ to A and A_r. Then, as W_2 is a distance,

$$r \leq W_2(\mu_A, \mu_{A_r}) \leq W_2(\mu_A, \mu) + W_2(\mu_{A_r}, \mu).$$

First of all

$$W_2(\mu_A,\mu) \le \sqrt{2R^{-1}H(\mu_A|\mu)} = \sqrt{2R^{-1}\log(1/\mu(A))} := c_A$$

by (22), or its weaker form (3). Let now $c_V = \int V d\mu, x_r = H(\mu_{A_r}|\mu) = \log(1/\mu(A_r))$ and $V_r = \int V d\mu_{A_r}$. By (22) again we get, for $r > c_A$,

$$(r-c_A)^2 \le W_2^2(\mu_{A_r},\mu) \le \frac{2}{R}\Big(V_r - c_V + n - n\exp\Big[-\frac{1}{n}(x_r + c_V - V_r)\Big]\Big).$$

Since $x_r = \log(1/\mu(A_r))$ we obtain :

Corollary 2.4 (Concentration inequality) *Following the same assumptions as in Theorem 2.1, let $A \subset \mathbb{R}^n$, $r > 0$ and $A_r = \{x;\, \forall y \in A,\, |y-x| > r\}, c_A = \sqrt{2R^{-1}\log(1/\mu(A))}, c_V = \int V d\mu, V_r = \int V d\mu_{A_r}$. Then for $r > c_A$*

$$\mu(A_r) \le e^{c_V - V_r}\Big[1 + \frac{1}{n}\big(V_r - c_V - \frac{R}{2}(r-c_A)^2\big)\Big]^n.$$

Since $(1+u/n)^n \le e^u$, the bound in Corollary 2.4 implies the classical Gaussian concentration

$$\mu(A_r) \le e^{-\frac{R}{2}(r-c_A)^2}, \qquad r > c_A$$

of the Talagrand inequality (3), see again [39, Chap. 22] for instance.
The bound in Corollary 2.4 captures the behaviour of concentration of the measure μ in a more accurate way: let for instance $V(x) = |x|^2/2 + |x|^p + Z_p$ with $p > 2$ and a normalizing factor Z_p, and A be the Euclidean unit ball in $\mathbb{R}^n$. Then $\mathrm{Hess}(V) \ge \mathrm{Id}_n$, so by Corollary 2.4 with $R = 1$ there exists a constant $C = C(p,n)$ such that for all $r > C$

$$\mu(|x| > r+1) = \mu(A_r) \le \exp\Big[c_V - V_r + n\,\log(1 + V_r/n)\Big].$$

But $V_r \ge r^p + Z_p$, so for all $\varepsilon < 1$ there exists another constant C depending also on ε such that for all $r > C$

$$\mu(|x| > r) \le e^{-(1-\varepsilon)r^p}.$$

This concentration inequality in this precise example can also be obtained by using a L^p-Talagrand inequality or a L^p-log Sobolev inequality; however we have found it interesting to get it by means of the dimension dependence of the classical Talagrand inequality, moreover in a shorter and more straightforward manner.

2.2 Tensorisation and comparison with earlier results

In $\mathbb{R}^n$, let W_1 be the Wasserstein distance between probability measures, for the cost $|y-x|, x, y \in \mathbb{R}^n$. Deficit in the Gaussian Talagrand inequality (for $\mu = \gamma$) and for centered measures ν has been investigated in *one dimension* in [6] and [8], in the form

$$\delta_{Tal}(\nu|\gamma) \ge c \inf_\pi \int_{\mathbb{R}\times\mathbb{R}} \Lambda(|y-x|) d\pi(x,y) \ge c \min\big\{W_1(\nu,\gamma)^2, W_1(\nu,\gamma)\big\}.$$

Here the c's are diverse numerical constants and the infimum runs over couplings π of γ and ν.
This second lower bound has been extended in [22, Th. 5] to *any dimension* n, as

$$\delta_{Tal}(\nu|\gamma) \ge c \min\left(\frac{W_{1,1}(\nu,\gamma)^2}{n}, \frac{W_{1,1}(\nu,\gamma)}{\sqrt{n}}\right) \tag{27}$$

as soon as ν has mean 0; here c is a numerical constant independent of the dimension n, and on $\mathbb{R}^n \times \mathbb{R}^n$

$$W_{1,1}(\mu,\nu) = \inf_\pi \int_{\mathbb{R}^n\times\mathbb{R}^n} \sum_{i=1}^n |y_i - x_i|\, d\pi(x,y).$$

Still under a centering condition, the bound (27) has been improved in [17, Prop. 3] by replacing the quantity $W_{1,1}/\sqrt{n}$ by the larger W_1 Wasserstein distance on $\mathbb{R}^n$, and extended to reference measures μ with density e^{-V} where $\mathrm{Hess}(V) \ge R\,\mathrm{Id}_n$.

In comparison, our bound (23) has the following two advantages : it holds without any centering condition on ν, and gives a lower bound on the deficit in terms of the relative entropy H : this is a strong way of measuring the gap between measures, by the Pinsker inequality for instance (see [39, Chap. 22]), and the relative entropy can be much larger than the weak distance W_2.

As considered in [17] and [22], a natural example is the product measure case when $\mu^N = \otimes^N \mu$ and $\nu^N = \otimes^N \nu$ on $\mathbb{R}^{nN}$ for $N \in \mathbb{N}^*$. Then $\delta_{Tal}(\nu^N|\mu^N) = N\, \delta_{Tal}(\nu|\mu)$ by tensorisation properties of both H and W_2^2. However, the above bound (27) in [17] (so with W_1 instead of $W_{1,1}/\sqrt{n}$) gives a lower bound on $\delta_{Tal}(\nu^N|\mu^N)$ equal to a constant c times

$$\min\left(W_1(\nu^N,\mu^N)^2, W_1(\nu^N,\mu^N)\right) \leq \min\left(W_2(\nu^N,\mu^N)^2, W_2(\nu^N,\mu^N)\right) = \min\left(NW_2(\nu,\mu)^2, \sqrt{N}W_2(\nu,\mu)\right)$$

since $W_1 \leq W_2$. Hence this lower bound has the good order in N at most only for small perturbations ν of the reference measure μ.

In contrast, our bound always has the correct order in N. Indeed, if $V^{(N)} = \oplus^N V$ so that $d\mu^N = e^{-V^{(N)}} dx$ on $\mathbb{R}^{nN}$, then

$$H(\nu^N|\mu^N) + \mu^N(V^{(N)}) - \nu^N(V^{(N)}) = N\left(H(\nu|\mu) + \mu(V) - \nu(V)\right);$$

hence Theorem 2.1 leads to

$$\delta_{Tal}(\nu^N|\mu^N) \geq N\, \delta_n \left(H(\nu|\mu) + \mu(V) - \nu(V)\right),$$

which has the correct order in N.

2.3 Useful facts on optimal transport

In the proof of Theorem 2.3 and in proofs below we shall need the following notation and facts.

If μ is a probability measure on $\mathbb{R}^n$ and $T : \mathbb{R}^n \to \mathbb{R}^n$ a Borel function, we let $T\#\mu$ be the image measure of μ by T, defined by $T\#\mu(h) = \mu(h \circ T)$ for all bounded continuous functions $h : \mathbb{R}^n \to \mathbb{R}$.

Let now μ_0 and μ_1 in $P_2(\mathbb{R}^n)$ be absolutely continuous with respect to Lebesgue measure. Then there exists a convex function φ on $\mathbb{R}^n$ such that $\mu_1 = \nabla\varphi\#\mu_0$ (see [38, Th. 2.12] or [39, Th. 10.41] for instance). The map $\nabla\varphi$ is called the Brenier map. Moreover

$$\int |\nabla\varphi(x) - x|^2 d\mu_0(x) = W_2^2(\mu_0,\mu_1).$$

Now, by the Alexandrov Theorem (see [34] or [39, Th. 14.1] for instance), a convex function ψ is almost everywhere twice differentiable: for almost every $x \in \mathbb{R}^n$ there exists a non negative symmetric matrix A such that

$$\psi(x+h) = \psi(x) + \nabla\psi(x)\cdot h + \frac{1}{2}Ah\cdot h + o(|h|^2)$$

as h tends to 0 in $\mathbb{R}^n$. The matrix A is denoted $\mathrm{Hess}(\psi)(x)$ and called the Hessian of ψ in the sense of Alexandrov. The trace of A will be denoted $\Delta\psi(x)$: it coincides with the density of the absolutely continuous part of the distributional Laplacian of ψ, the singular part being a non negative measure.

In fact, in the above notation and by [34, Th. 4.4] or [1, Th. 6.2.12], $\mathrm{Hess}(\varphi)(x)$ is a positive matrix for μ_0-almost every x. Moreover, by [34] (see also [1, Lem. 5.5.3]), the Brenier map solves the Monge-Ampère equation

$$\mu_0(x) = \mu_1(\nabla\varphi(x))\det(\mathrm{Hess}(\varphi)(x)) \tag{28}$$

at μ_0-almost every x in $\mathbb{R}^n$. Here μ_0 and μ_1 are the densities of the measures.

Let now φ^* be the Legendre transform of φ. Then $\mu_0 = \nabla\varphi^*\#\mu_1$ by [38, Th. 2.12] for instance. Moreover $\nabla\varphi^*(\nabla\varphi(x)) = x$ and $\nabla\varphi(\nabla\varphi^*(y)) = y$ for μ_0-almost every x and μ_1-almost every y.

Furthermore, by [34, Th. A.1], if $\mathrm{Hess}(\varphi)(x)$ is invertible at x then φ^* is twice differentiable at $\nabla\varphi(x)$, with $\mathrm{Hess}(\varphi^*)(\nabla\varphi(x)) = \left[\mathrm{Hess}(\varphi)(x)\right]^{-1}$. By the remark above, this is the case for μ_0-almost every x.

Finally, the curve $(\mu_s)_{s\in[0,1]}$ defined by $\mu_s = ((1-s)Id + s\nabla\varphi)\#\mu_0$ is a geodesic path in $P_2(\mathbb{R}^n)$ between μ_0 and μ_1, in the sense that

$$W_2(\mu_s,\mu_t) = |t-s|\, W_2(\mu_0,\mu_1)$$

for all $0 \leq s,t \leq 1$. It holds that μ_s is also absolutely continuous with respect to Lebesgue measure, see [34, Prop. 1.3] or [38, Th. 5.9] for instance.

2.4 Proof of Theorem 2.3

Theorem 2.3 is a consequence of the relation

$$H(h\mu|\mu) - \mu(hV) = \mathrm{Ent}_{dx}(he^{-V}) \tag{29}$$

written with $h = f, g$ and of the following lemma.

Lemma 2.5 *Following the same assumptions as in Theorem 2.3, let f, g be two smooth functions such that $f\mu$ and $g\mu$ belong to $P_2(\mathbb{R}^n)$. Let φ be a convex function on $\mathbb{R}^n$ such that $\nabla\varphi\#(f\mu) = g\mu$. Then*

$$\int V\, g\, d\mu - \int V\, f\, d\mu - \int (\nabla\varphi - x)\cdot\nabla f\, d\mu \geq n\exp\Big[\frac{1}{n}\Big(\mathrm{Ent}_{dx}(fe^{-V}) - \mathrm{Ent}_{dx}(ge^{-V})\Big)\Big] - n$$
$$+ \int\int_0^1 (\nabla\varphi(x) - x)\cdot \mathrm{Hess}(V)(x + t(\nabla\varphi(x) - x))(\nabla\varphi(x) - x)(1-t)dt\, f(x)\, d\mu(x).$$

Indeed, if $\mathrm{Hess}(V) \geq R\,\mathrm{Id}_n$, then the last term above is greater than $\frac{R}{2}\int |\nabla\varphi - x|^2 f\, d\mu = \frac{R}{2}W_2^2(f\mu, g\mu)$. Moreover, on the left-hand side,

$$-\int (\nabla\varphi - x)\cdot\nabla f\, d\mu \leq \Big[\int |\nabla\varphi - x|^2\, f\, d\mu\Big]^{1/2}\Big[\int \frac{|\nabla f|^2}{f} d\mu\Big]^{1/2} = W_2(f\mu, g\mu)\sqrt{I(f\mu|\mu)}$$

by the Cauchy-Schwarz inequality. This implies Theorem 2.3.

Proof of Lemma 2.5.
By the Taylor formula,

$$V(\nabla\varphi(x)) - V(x) = \nabla V(x)\cdot(\nabla\varphi(x) - x) + \int_0^1 (\nabla\varphi(x) - x)\cdot\mathrm{Hess}(V)(x + t(\nabla\varphi(x) - x))(\nabla\varphi(x) - x)(1-t)dt$$

for almost every x in $\mathbb{R}^n$. We now integrate with respect to $f\,\mu$ and use the comparison between Alexandrov and distributional Laplacians to deduce that

$$\int \nabla V(x)\cdot(\nabla\varphi(x) - x)\, f(x)\, d\mu(x) \geq \int \big[(\Delta\varphi - n)f + (\nabla\varphi - x)\cdot\nabla f\big]\, d\mu = \int \Delta\varphi\, f\, d\mu - n + \int (\nabla\varphi - x)\cdot\nabla f\, d\mu,$$

as in [16] or [38, Th. 9.17] for instance. This leads to

$$\int V\, g\, d\mu - \int V\, f\, d\mu - \int (\nabla\varphi - x)\cdot\nabla f\, d\mu \geq \int \Delta\varphi\, f\, d\mu - n$$
$$+ \int\int_0^1 (\nabla\varphi(x) - x)\cdot \mathrm{Hess}(V)(x + t(\nabla\varphi(x) - x))(\nabla\varphi(x) - x)(1-t)dt\, f(x)\, d\mu(x). \tag{30}$$

Then Lemma 2.5 is a consequence of the following Lemma.

Lemma 2.6 *Let $\mu_0, \mu_1 \in P_2(\mathbb{R}^n)$ absolutely continuous with respect to Lebesgue measure, with respective densities also denoted μ_0 and μ_1. Let φ be a convex function on $\mathbb{R}^n$ such that $\nabla\varphi\#\mu_0 = \mu_1$. Then*

$$\int \Delta\varphi\, d\mu_0 \geq n\,\exp\left[\frac{\mathrm{Ent}_{dx}(\mu_0) - \mathrm{Ent}_{dx}(\mu_1)}{n}\right]. \tag{31}$$

Proof
◁ Taking logarithms in the Monge-Ampère equation (28) and integrating with respect to μ_0 lead to

$$\mathrm{Ent}_{dx}(\mu_0) = \mathrm{Ent}_{dx}(\mu_1) + \int \log\det(\mathrm{Hess}(\varphi))\, d\mu_0. \tag{32}$$

Now, if for each x the symmetric matrix $\mathrm{Hess}(\varphi)$ has eigenvalues φ_i, then by the Jensen inequality

$$\int \log\det(\mathrm{Hess}(\varphi))\, d\mu_0 = n\,\frac{1}{n}\sum_i \int \log(\varphi_i)\, d\mu_0 \leq n\log\Big(\int \frac{1}{n}\sum_i \varphi_i\, d\mu_0\Big) = n\log\Big(\frac{1}{n}\int \Delta\varphi\, d\mu_0\Big).$$

This concludes the proof. ▷

Remark 2.7 *In the Gaussian case, we have already observed that translations of the Gaussian measure are extremals of the Talagrand inequality. As observed in [16], or as can be observed from the proof above, there are no other extremals. Indeed the Hessian of the map φ has to be constant and equal to the identity matrix for all inequalities to be equalities.*
In fact, if $\mathrm{Hess}(V) \geq R\,\mathrm{Id}_n$*, then equality in the Talagrand inequality implies that the potential is necessarily Gaussian and that extremals are translations of the Gaussian measure.*

2.5 Logarithmic Sobolev inequalities by transport

As observed in [36], the HWI inequality (25) classically implies the logarithmic Sobolev inequality (1) by bounding from above the second order polynomial in W_2 in HWI by its maximum. Likewise, the dimensional HWI inequality (26) is another path towards dimensional logarithmic Sobolev inequalities. Here we obtain :
Let μ have density e^{-V} where V is $\mathcal{C}^2$ and satisfies $\mathrm{Hess}(V) \geq R\,\mathrm{Id}_n$ with $R > 0$. Then

$$H(\nu|\mu) \leq \nu(V) - \mu(V) + n \log\Big(1 + \frac{1}{n}\Big(\frac{I(\nu|\mu)}{2R} + \mu(V) - \nu(V)\Big)\Big)$$

for all ν. Equivalently, in terms of deficit,

$$\delta_{LSI}(\nu|\mu) \geq R \max\Big\{\delta_n\Big(\nu(V) - \mu(V) - H(\nu|\mu)\Big), \Lambda_n\Big(\frac{I(\nu|\mu)}{2R} - \nu(V) + \mu(V))\Big)\Big\}. \tag{33}$$

In the Gaussian case, then $R = 1$ and we obtain a bound which is slightly worse than (5), where a $\log(1 + 2u)$ term is replaced by the larger $2\log(1 + u)$.
At this point, let us observe that still in the Gaussian case a dimensional HWI has been derived in [8, Th. 1.1]. It is also observed by the authors that the HWI inequality in [8] does not seem to imply (5). We could not compare the HWI in [8] to our bound (26) in full generality. However, if $\nu(|x|^2) = n = \gamma(|x|^2)$ then they can respectively be written as

$$2h \leq x - y + \log(1 + x) \qquad \text{and} \qquad h \leq \log(1 + x - y/2)$$

for $x = W_2\sqrt{I}/n, y = W_2^2/n$ and $h = H/n$; hence our bound is at least significantly more precise in the common range $I \gg W_2 \sim 1$: indeed then $x \gg y \sim 1$ in this range, so that comparing the two right-hand sides amounts to $x \gg \log(1 + x)$.

As remarked in [8, 22] it is also possible to get refined logarithmic Sobolev inequalities by combining the HWI and Talagrand inequalities. Here, if $\mathrm{Hess}(V) \geq R\,\mathrm{Id}_n$ with $R > 0$, then (26) can be written as

$$H + \delta_n(-h) \leq W_2\sqrt{I} - \frac{R}{2}W_2^2 \tag{34}$$

where $h = H + \mu(V) - \nu(V)$. Moreover $H = \frac{R}{2}W_2^2 + \delta_{Tal}$, so

$$\frac{\delta_{Tal} + \delta_n(-h)}{W_2} \leq \sqrt{I} - RW_2.$$

Then, by (34) again and Theorem 2.1,

$$\begin{aligned}\delta_{LSI} = \frac{1}{2}I - R\,H \geq R\,\delta_n(-h) + \frac{1}{2}\left(\sqrt{I} - R\,W_2\right)^2 &\geq R\delta_n(-h) + \frac{1}{2}\frac{(\delta_{Tal} + \delta_n(-h))^2}{W_2^2} \\ &\geq R\delta_n(-h) + \frac{1}{2}\frac{(\delta_n(h) + \delta_n(-h))^2}{W_2^2}.\end{aligned}$$

In particular this improves upon the first lower bound in (33). Let us recall that the function δ_n is defined above, after Theorem 2.1.

Refined Gaussian logarithmic Sobolev inequalities have been considered for certain classes of test measures ν : measures ν satisfying lower and upper curvature bounds as in [8] and [32], measures ν satisfying a (weaker) Poincaré inequality as in [22]. Under these additional assumptions on ν, the goal is then to

obtain better constants in the logarithmic Sobolev inequality, mimicking in a sense the phenomenon observed in the Poincaré inequality when considering test functions orthogonal to the first eigenfunctions. In Indrei-Marcon [32], the deficit is controlled by the Wasserstein distance for the class of centered functions with upper and lower bounded curvature. The authors in [8] also give new bounds in terms of conditionally centered vectors. Further improvements are given in [22] in terms of the $W_{1,1}$ distance defined in Section 2.2. Here again our bounds share the advantages of holding without any smoothness, centering, etc. hypothesis on ν, and of having the good dimensional behaviour when considering product measures.

3 Applications to Fokker-Planck equations

Let us now see how our results (or methods) lead to short-time smoothing of the entropy and improved contraction rates for the laws of solutions to stochastic differential equations.

For this, let again V be a C^2 function on $\mathbb{R}^n$ such that $\int e^{-V} = 1$ and $\mathrm{Hess}(V) \geq R\,\mathrm{Id}_n$, with R possibly negative, and satisfying the doubling condition $V(x+y) \leq C(1+V(x)+V(y))$ for a C and all x, y. Let also μ be the probability measure with density e^{-V}. We let u_0 in $P_2(\mathbb{R}^n)$ and consider gradient flow solutions $u = (u_t)_{t\geq 0} \in C([0,+\infty), P_2(\mathbb{R}^n))$ of the Fokker-Planck equation

$$\frac{\partial u_t}{\partial t} = \Delta u_t + \nabla \cdot (u_t \nabla V), \qquad t > 0, x \in \mathbb{R}^n \tag{35}$$

as in [1, Chap. 11.2.1] and [18, Th. 4.20 and 4.21] (see also [33]). Equation (35) holds in the sense of distributions. Moreover, by [1, Th. 11.2.8] or again [18], for any $t > 0$ the solution u_t has a density; for almost every $t > 0$ this density is in $W^{1,1}_{loc}(\mathbb{R}^n)$, with $\nabla u_t/u_t + \nabla V \in L^2(u_t)$; finally $t \mapsto I(u_t|\mu) \in L^1_{loc}(]0,+\infty[)$ and

$$\frac{d}{dt} H(u_t|\mu) = -I(u_t|\mu)$$

for almost every $t > 0$. The solution u_t can be seen as the law at time t of the solution $(X_t)_{t\geq 0}$ to the stochastic differential equation

$$dX_t = \sqrt{2}\, dB_t - \nabla V(X_t)\, dt.$$

Here $(B_t)_{t\geq 0}$ is a standard Brownian motion on $\mathbb{R}^n$ and the initial datum X_0 has law u_0.

Moreover, the interpretation of (35) as the gradient flow of $H(\cdot|\mu)$ on the space $P_2(\mathbb{R}^n)$ has enabled to obtain the following short-time and contraction properties (see [1, Th. 11.2.1] and [39, Chap. 24]). Let u and v be solutions to (35). Then

$$H(u_t|\mu) \leq \frac{W_2^2(u_0,\mu)}{2t} e^{2\max\{-R,0\}\, t}, \qquad t > 0 \tag{36}$$

and

$$W_2(u_t, v_t) \leq e^{-Rt}\, W_2(u_0, v_0), \qquad t \geq 0. \tag{37}$$

In particular, if $R > 0$, then u_t converges to the steady state μ as

$$W_2(u_t, \mu) \leq e^{-Rt}\, W_2(u_0, \mu), \qquad t \geq 0. \tag{38}$$

The purpose of this section is to improve these three properties by means of the tools and inequalities in the above sections.

3.1 Short-time smoothing of the entropy

In the Gaussian case where μ is the standard Gaussian measure γ, the solution to (35) is given by the Mehler formula (see [4, Sect. 2.7.1]). In particular the fundamental solution, with initial datum u_0 the Dirac mass at 0, is at time $t > 0$ the Gaussian measure with variance $\sigma_t^2 = 1 - e^{-2t}$:

$$u_t(x) = (2\pi\sigma_t^2)^{-n/2} e^{-x^2/(2\sigma_t^2)}, \qquad z \in \mathbb{R}^n.$$

Its relative entropy can be computed as

$$H(u_t|\gamma) = \int_{\mathbb{R}^n} u_t(x) \log \frac{u_t(x)}{\gamma(x)} dx = -\frac{n}{2}\left[e^{-2t} + \log(1 - e^{-2t})\right].$$

Of course this is coherent with (36), with $R = 1$, since

$$-\frac{n}{2}\big[e^{-2t} + \log(1 - e^{-2t})\big] \leq \frac{n}{2t} = \frac{W_2^2(u_0, \mu)}{2t}$$

by direct computation. In fact, for $t \sim 0$ one can observe that

$$H(u_t|\gamma) \sim \frac{n}{2} \log \frac{1}{t}.$$

On the other hand, let u be a solution to (35), still in the Gaussian case, and with initial datum u_0 such that $u_0(|x|^2) = n = \gamma(|x|^2)$. Then $u_t(|x|^2) = n$ for all t since

$$\frac{d}{dt} \int |x|^2 \, du_t = 2n - 2 \int |x|^2 \, du_t. \tag{39}$$

In particular, in the notation $H(t) = H(u_t|\gamma)/n$ and $I(t) = I(u_t|\gamma)/n$, the dimensional Gaussian logarithmic Sobolev inequality (5) simplifies as $2H \leq \log(1 + I)$. Hence

$$H'(t) = -I(t) \leq 1 - e^{2H(t)}, \qquad \text{for } a.e.\ t > 0.$$

By the change of variable $x(t) = e^{-2h(t)}$ this integrates into

$$x(t)e^{2t} \geq x(0) + e^{2t} - 1 \geq e^{2t} - 1.$$

In other words

$$H(u_t|\gamma) \leq -\frac{n}{2} \log(1 - e^{-2t}), \qquad t > 0$$

which gives the same short-time behaviour.

More generally :

Proposition 3.1 *Let u be a solution to (35) with $\mathrm{Hess}(V) \geq R\,\mathrm{Id}_n$, $R > 0$, and with initial condition u_0 in $P_2(\mathbb{R}^n)$. Let $T > 0$ and assume that $u_t(|\nabla V|^2) \leq M$ for t in $[0, T]$. Then there exists a constant $c > 0$ depending only on n, R and M such that*

$$H(u_t|\mu) \leq \max\Big\{1, \frac{n}{2} \log \frac{c}{t}\Big\}, \qquad t \leq T.$$

Remark 3.2 *The moment assumption $u_t(|\nabla V|^2) \leq M$ for t in $[0, T]$, is not a restrictive condition. It can indeed be checked by time differentiating $u_t(|\nabla V|^2)$ and controlling its non explosion via a Lyapunov type condition on $u_0 e^V$ or on derivatives of V for instance.*
It can also be checked by observing that the Markov semigroup $(P_t)_{t \geq 0}$ with generator $L = \Delta - \nabla V \cdot \nabla$ is such that $\int \phi du_t = \int P_t \phi du_0$ for any test function ϕ. In particular, if Φ is a convex function and if the initial datum has a density also denoted u_0, then

$$\begin{aligned} u_t(|\nabla V|^2) = \int |\nabla V|^2 du_t = \int P_t(|\nabla V|^2) du_0 = \int P_t(|\nabla V|^2) u_0 e^V d\mu \\ \leq \int \Phi(P_t(|\nabla V|^2)) d\mu + \int \Phi^*(u_0 e^V) d\mu \leq \int \Phi(|\nabla V|^2) d\mu + \int \Phi^*(u_0 e^V) d\mu. \end{aligned}$$

Here we use the fact that $t \mapsto \int \Phi(P_t(|\nabla V|^2)) d\mu$ is non increasing since Φ is convex. The moment assumption is then satisfied for all $T > 0$ as soon as the right hand side is finite for a convex function Φ.

Proof
$\lhd$ We shall let c denote diverse positive constants depending only on n, M and R. By Corollary 1.4 applied to the measure $f^2\mu = u_t$, and integration by parts, there holds

$$H(u_t|\mu) \leq n(s - 1 - \log s) + \frac{1 - s^2}{2R} u_t(|\nabla V|^2) + \frac{s(s-1)}{R} u_t(\Delta V) + \frac{s^2}{2R} I(u_t|\mu)$$

for $t > 0$ and $s > 0$. Recall that I has been introduced in (2). Since V is convex, then $\Delta V \geq 0$ and then

$$H(u_t|\mu) \leq -n \log s + c + \frac{s^2}{2R} I(u_t|\mu)$$

for all $s \in]0,1]$ and $t \in]0,T]$.
Now, as far as $H(t) := H(u_t|\mu) \geq 1$, then $I(t) := I(u_t|\mu) \geq 2R$ so that $s = \sqrt{2R/I}$ is smaller than 1. For this s we obtain

$$H \leq c + \frac{n}{2} \log I.$$

Hence

$$H'(t) = -I(t) \leq -e^{2H(t)/n - c}$$

for almost every $t > 0$. As above $x(t) = e^{-2H/n}$ satisfies $x(t) \geq x(0) + ct \geq ct$ by time integration. Written in terms of H, this concludes the proof. ▷

3.2 Refined contraction properties

Let us now see how to make (37) finer. Still by [1, Th. 8. 3. 1] and [18, Th. 4.20 and 4.21], one can write (35) as the continuity equation

$$\frac{\partial u_t}{\partial t} + \nabla \cdot (\xi[u_t] u_t) = 0, \qquad t > 0, x \in \mathbb{R}^n$$

with $\xi[u_t] = -\nabla V - \nabla \log u_t$. Then for almost every $t > 0$

$$-\frac{1}{2}\frac{d}{dt} W_2^2(u_t, v_t) = \int \big(\xi[v_t](\nabla\varphi_t(x)) - \xi[u_t](x)\big) \cdot (\nabla\varphi_t(x) - x)\, u_t(x)\, dx \tag{40}$$

$$\geq \int \Big[\Delta\varphi_t(x) + \Delta\varphi_t^*(\nabla\varphi_t(x)) - 2n + \big(\nabla V(\nabla\varphi_t(x)) - \nabla V(x)\big) \cdot (\nabla\varphi_t(x) - x)\Big]\, u_t(x)\, dx \tag{41}$$

for two solutions u and v. Here φ_t is the convex map such that $v_t = \nabla\varphi_t \# u_t$ and $u_t = \nabla\varphi_t^* \# v_t$ for the Legendre transform φ_t^* of φ_t (see Section 2.3). Equality (40) follows from [39, Th. 23.9] (see also [1, Th. 8.4.7]); its assumptions are satisfied since (and likewise for v)

$$\int_{t_1}^{t_2} \int_{\mathbb{R}^n} |\xi[u_s]|^2 du_s\, ds = \int_{t_1}^{t_2} I(u_s|\mu) ds = H(u_{t_1}|\mu) - H(u_{t_2}|\mu) \leq H(u_{t_1}|\mu)$$

which is finite for any $t_2 > t_1 > 0$, as observed above. Inequality (41) follows from a weak integration by parts, as in [33, Th. 1.5]; there again $\Delta\varphi_t$ is the trace of the Alexandrov Hessian of φ_t.
Now, for given $t > 0$ and u_t-almost every x, the symmetric matrix $\mathrm{Hess}(\varphi_t)(x)$ is positive, as recalled in Section 2.3 : letting $e^{2\lambda_i(x)}$ for $i = 1, \ldots, n$ its n positive eigenvalues , then its inverse matrix $\mathrm{Hess}(\varphi_t^*)(\nabla\varphi_t(x))$ (see again Section 2.3) has eigenvalues $e^{-2\lambda_i(x)}$; hence at point x

$$\Delta\varphi_t + \Delta\varphi_t^*(\nabla\varphi_t) - 2n = \mathrm{tr}\big[\mathrm{Hess}(\varphi_t)\big] + \mathrm{tr}\big[\mathrm{Hess}(\varphi_t^*)(\nabla\varphi_t)\big] - 2n = \sum_i \big(e^{2\lambda_i} + e^{-2\lambda_i} - 2\big) = 4\sum_i \sinh^2(\lambda_i). \tag{42}$$

Hence, by convexity of $\sinh^2$ and the Jensen inequality, and (32),

$$\begin{aligned}
\int \Big[\Delta\varphi_t(x) + \Delta\varphi_t^*(\nabla\varphi_t(x)) - 2n\Big] u_t(x)\, dx &= 4n \frac{1}{n} \sum_i \int \sinh^2(\lambda_i(x))\, u_t(x)\, dx \\
&\geq 4n \sinh^2\left(\frac{1}{n}\sum_i \int \lambda_i(x)\, u_t(x)\, dx\right) \\
&= 4n \sinh^2\left(\frac{1}{2n} \int \log\det \mathrm{Hess}(\varphi_t)(x)\, u_t(x)\, dx\right) \\
&= 4n \sinh^2\left(\frac{\mathrm{Ent}_{dx}(v_t) - \mathrm{Ent}_{dx}(u_t)}{2n}\right).
\end{aligned}$$

Since $\mathrm{Hess}(V) \geq R\,\mathrm{Id}_n$, we obtain

$$-\frac{1}{2}\frac{d}{dt}W_2^2(u_t,v_t) \geq 4n\sinh^2\left(\frac{\mathrm{Ent}_{dx}(v_t)-\mathrm{Ent}_{dx}(u_t)}{2n}\right) + RW_2^2(u_t,v_t). \tag{43}$$

By time integration this ensures the following dimensional contraction property :

Proposition 3.3 *In the above notation, if* $\mathrm{Hess}(V) \geq R\,\mathrm{Id}_n$ *for* $R \in \mathbb{R}$, *then for any solutions to* (35)

$$W_2^2(u_t,v_t) \leq e^{-2Rt}W_2^2(u_0,v_0) - 8n\int_0^t e^{-2R(t-s)}\sinh^2\left(\frac{\mathrm{Ent}_{dx}(v_s)-\mathrm{Ent}_{dx}(u_s)}{2n}\right)ds, \qquad t \geq 0. \tag{44}$$

For the heat equation, namely for $V = 0$, then the associated Markov generator $L = \Delta$ satisfies the $CD(0,n)$ curvature-dimension condition: in particular in this case the bound (44) has been derived in [13] and [14], and is also a consequence of [21]. For $V \neq 0$, then the associated generator $L = \Delta - \nabla V \cdot \nabla$ satisfies a $CD(R,\infty)$ but no $CD(R,n)$ condition: in particular the bound (44) can not be obtained from the works mentioned above.

Remark 3.4 *The above computation can be extended to drifts* $A(x)$ *which are not gradients. In this case the assumption* $\mathrm{Hess}(V) \geq R\,\mathrm{Id}_n$ *should be replaced by the monotonicity condition* $(A(y)-A(x))\cdot(y-x) \geq R\,|y-x|^2$ *for all* x, y *(see [12] for this non-gradient case).*

3.3 A formal gradient flow argument to Proposition 3.3

In this subsection, we provide an alternative formal argument to Proposition 3.3 based on gradient flow.

We begin with the following elementary lemma which gives additional information to [21, Lem. 2.2].

Lemma 3.5 *Let* ψ *be a* C^2 *function on* $[0,1]$. *Then the following properties are equivalent:*

- $\psi'' \geq \psi'^2/n$;
- *for all* r, s *in* $[0,1]$,
$$n - \psi'(r)(s-r) \geq n\,e^{\frac{\psi(r)-\psi(s)}{n}}; \tag{45}$$
- *for all* r, s *in* $[0,1]$,
$$\big(\psi'(s)-\psi'(r)\big)(s-r) \geq 4n\sinh^2\Big(\frac{\psi(s)-\psi(r)}{2n}\Big). \tag{46}$$

Proof
◁ Let indeed $U = e^{-\psi/n}$, so that

$$U'' = -\Big(\psi'' - \frac{\psi'^2}{n}\Big)\frac{U}{n}.$$

Then $\psi'' \geq \psi'^2/n$ if and only if U is concave, hence if and only

$$e^{-\frac{\psi(s)}{n}} = U(s) \leq U(r) + U'(r)(s-r) = e^{-\frac{\psi(r)}{n}} - \frac{\psi'(r)}{n}e^{-\frac{\psi(r)}{n}}(s-r)$$

for all $r, s \in [0,1]$, which is (45) when multiplying both sides by $e^{\psi(r)/n}$.
Adding (45) with the corresponding bound obtained with r, s instead of s, r leads to (46). Conversely, dividing (46) by $(s-r)^2$ and letting s go to r gives $\psi'' \geq \psi'^2/n$ at point r. ▷

Let now μ^0 and μ^1 be absolutely continuous measures in $P_2(\mathbb{R}^n)$, $\nabla\varphi$ their Brenier map and $(\mu^s)_{s\in[0,1]}$ the geodesic between them, as in Section 2.3. Here again we identify the measures with their densities. Let us now recall why the function $\psi : s \mapsto \mathrm{Ent}_{dx}(\mu^s)$ formally satisfies $\psi'' \geq \psi'^2/n$ on [0,1]. For this, recall from Section 2.3 that for μ^0-almost every x the Alexandrov Hessian $\mathrm{Hess}(\varphi)(x)$ is positive, so that the eigenvalues $\theta_i(x)$ of $\mathrm{Hess}(\varphi)(x) - I$ are > -1. Writing (32) with the measures $\mu_0 = \mu^0$ and $\mu_1 = \mu^s$, we obtain

$$\psi(0) = \psi(s) + \int \log\det(I + s(\mathrm{Hess}(\varphi)(x) - I))\,d\mu^0(x) = \psi(s) + \sum_i \int \log(1+s\theta_i(x))\,d\mu^0(x).$$

Hence

$$\psi'(s) = -\sum_i \int \frac{\theta_i}{1+s\theta_i} d\mu^0 \tag{47}$$

and then by the Cauchy-Schwarz inequality

$$\psi''(s) = n\,\frac{1}{n}\sum_i \int \frac{\theta_i^2}{(1+s\theta_i)^2} d\mu^0 \geq n\left(\frac{1}{n}\sum_i \int \frac{\theta_i}{1+s\theta_i} d\mu^0\right)^2 = \frac{1}{n}\psi'(s)^2.$$

Remark 3.6 *Identity (47) can also be formally checked using the continuity equation solved by $(\mu^s)_{s\in[0,1]}$:*

$$\frac{\partial \mu^s}{\partial s} + \nabla\cdot(\mu^s v^s) = 0.$$

Here the vector field v^s satisfies $v^s(x + s(\nabla\varphi(x) - x)) = \nabla\varphi(x) - x$, see e. g. [38, Th. 5.51]. For, and recalling that $\psi(s) = \int \mu^s \log \mu^s\, dx$

$$\psi'(s) = -\int \nabla\cdot(v^s\mu^s)\log\mu^s\, dx = -\int \nabla\cdot v^s\,\mu^s\, dx = -\int \big(\nabla\cdot v^s\big)(x+s(\nabla\varphi(x)-x))\, d\mu^0(x)$$

by integration by parts and since $(x + s(\nabla\varphi(x) - x))\#\mu^0 = \mu^s$. Identity (47) follows since by chain rule

$$\big(\nabla\cdot v^s\big)(x+s(\nabla\varphi(x)-x)) = tr\Big[\big(\mathrm{Hess}(\varphi)(x) - I\big)\big(I + s(\mathrm{Hess}(\varphi)(x) - I)\big)^{-1}\Big] = \sum_i \frac{\theta_i}{1+s\theta_i}.$$

Remark 3.7 *In the above notation, observe that (45) in Lemma 3.5 for $\psi(s) = Ent_{dx}(\mu^s), r = 0$ and $s = 1$ formally leads to (31) in Lemma 2.6. For, in the notation of Remark 3.6 and by integration by parts,*

$$\psi'(0) = \int \nabla\mu^0\cdot v^0\, dx = \int \nabla\mu^0\cdot(\nabla\varphi - x)dx = n - \int \Delta\varphi\, d\mu^0.$$

We can now deduce an alternative formal argument to the bound in Proposition 3.3.

We begin with the following classical observation in Euclidean space : Let X and Y be two solutions of the Euclidean gradient flow $X_t' = -\nabla U(X_t)$ in $\mathbb{R}^d$, where $U : \mathbb{R}^d \to \mathbb{R}$ is a smooth potential. For $t > 0$ let $U_t(s) = U(X_t + s(Y_t - X_t))$ for $s \in [0, 1]$. Then

$$-\frac{1}{2}\frac{d}{dt}|Y_t - X_t|^2 = (Y_t - X_t)\cdot(\nabla U(Y_t) - \nabla U(X_t)) = U_t'(1) - U_t'(0). \tag{48}$$

Let now u and v two solutions to the Fokker-Planck equation (35), which by [1, Chap. 11.2] and [39, Chap. 23] is the gradient flow of $H(\cdot|\mu)$ on the space $P_2(\mathbb{R}^n)$. For any $t > 0$, let $\nabla\varphi_t$ be the optimal transport map between u_t and v_t, and $(\mu_t^s)_{s\in[0,1]}$ be the geodesic path in $P_2(\mathbb{R}^n)$ between u_t and v_t, as in Section 2.3. Then, formally and by analogy with (48),

$$-\frac{1}{2}\frac{d}{dt}W_2^2(u_t, v_t) = E_t'(1) - E_t'(0) \tag{49}$$

where for given $t > 0$ we let

$$E_t(s) = H(\mu_t^s|\mu) = \mathrm{Ent}_{dx}(\mu_t^s) + \int V\, d\mu_t^s.$$

Indeed, let $\psi : s \mapsto \mathrm{Ent}_{dx}(\mu_t^s)$ for given t and, for each x let the matrix $\mathrm{Hess}(\varphi_t)(x)$ have eigenvalues $e^{2\lambda_i}$. Then, in the above notation $\theta_i = e^{2\lambda_i} - 1$, (47) for $\mu = u_t$ gives

$$\begin{aligned}
\psi'(1) - \psi'(0) &= \int \sum_i \Big[\theta_i - \frac{\theta_i}{1+\theta_i}\Big] du_t \\
&= \int \sum_i \frac{\theta_i^2}{1+\theta_i} du_t \\
&= \int \sum_i \Big[e^{2\lambda_i} + e^{-2\lambda_i} - 2\Big] du_t = \int \Big[\Delta\varphi_t + \Delta\varphi_t^*(\nabla\varphi_t) - 2\,n\Big] du_t
\end{aligned}$$

as in (42). Using moreover the formal derivative

$$\frac{d}{ds}\int V\,d\mu_t^s = \frac{d}{ds}\int V\big(x+s(\nabla\varphi_t(x)-x)\big)du_t(x) = \int \nabla V\big(x+s(\nabla\varphi_t(x)-x)\big)\cdot(\nabla\varphi_t(x)-x)\,du_t(x)$$

for $s=0,1$ we formally recover (41) in (49).
We now use the fact that for given t the function ψ satisfies $\psi'' \geq \psi'^2/n$ on $[0,1]$. Then, by (46) in Lemma 3.5 for $r=0$ and $s=1$ we obtain

$$E_t'(1)-E_t'(0) \geq 4n\sinh^2\Big(\frac{Ent_{dx}(v_t)-Ent_{dx}(u_t)}{2n}\Big) + \int \big(\nabla V(\nabla\varphi_t(x))-\nabla V(x)\big)\cdot\big(\nabla\varphi_t(x)-x\big)\,du_t(x).$$

Since $\int|\nabla\varphi_t(x)-x|^2du_t(x) = W_2^2(u_t,v_t)$ this leads to (43) and then to (44) as soon as $\mathrm{Hess}(V)\geq R\,\mathrm{Id}_n$.

3.4 Improved convergence rates

In this section we consider a solution u to (35) in the Gaussian case where $\mu=\gamma$, and for which we can take $R=1$ above. Let us see how the contraction property (44) can make the convergence estimate (37) more precise.

For simplicity we assume that $u_0(|x|^2)\leq n=\gamma(|x|^2)$. Then $u_t(|x|^2)\leq n$ for all t, by (39). Hence (29) and the Talagrand inequality (22) ensure that $0\leq W_2^2(u_t,\gamma)<2n$ and

$$\frac{\mathrm{Ent}_{dx}(u_t)-\mathrm{Ent}_{dx}(\gamma)}{n} \geq -\log\Big(1-\frac{W_2^2(u_t,\gamma)}{2n}\Big).$$

In particular the right-hand side is non negative. Moreover, for the stationary solution $v_t=v_0=\gamma$, the contraction property (44) with $R=1$, in the form (43), implies

$$-x' \geq \frac{x^2}{1-x}+2\,x$$

where $x(t)=W_2^2(u_t,\gamma)/(2n)\in[0,1)$. Here we use that $\sinh(\log x)=(x-1/x)/2$. In other words $z(t)=1-(1-x(t))^2$ satisfies $z'\leq -2z$. This integrates into $z(t)\leq e^{-2t}z(0)$, that is,

$$x(t)\leq 1-\Big(1-(2x(0)-x(0)^2)e^{-2t}\Big)^{\frac12}. \tag{50}$$

By the lower bound

$$1-(2x(0)-x(0)^2)e^{-2t} \geq (1-x(0)e^{-2t})^2 \tag{51}$$

it implies the classical bound (37). It also improves it: for instance (50) can be written as

$$W_2^2(u_t,\gamma)\leq W_2^2(u_0,\gamma)e^{-2t}\frac{2-x(0)}{1+\Big(1-(2x(0)-x(0)^2)e^{-2t}\Big)^{\frac12}}.$$

Then by (51) we obtain

Corollary 3.8 *In the above notation, let u be a solution to (35) in the Gaussian case, with initial datum u_0 such that $u_0(|x|^2)\leq n$. Then for all $t\geq 0$*

$$W_2^2(u_t,\gamma)\leq W_2^2(u_0,\gamma)e^{-2t}\,\frac{1-W_2^2(u_0,\gamma)/(4n)}{1-W_2^2(u_0,\gamma)e^{-2t}/(4n)}.$$

Observe that the quotient is smaller than 1.

Remark 3.9 *The Gaussian assumption is used here only to ensure uniform convexity of the potential (hence the Talagrand inequality), and that $\int V du_t \leq \int V e^{-V}dx$ as soon as this holds at $t=0$.*

4 Brascamp-Lieb inequalities

It is classical that linearizing a logarithmic Sobolev inequality leads to a Poincaré inequality, which in the Gaussian case is the Brascamp-Lieb inequality. In this section we shall see how to obtain two different dimensional Brascamp-Lieb inequalities: a first one by an improvement of the classical L^2 method, and a second one by linearization in the Borell-Brascamp-Lieb inequality (13).

4.1 Brascamp-Lieb inequality by L^2 method

Proposition 4.1 (Dimensional Brascamp-Lieb inequality I) *Let μ be a probability measure on $\mathbb{R}^n$ with density e^{-V} where V is a $\mathcal{C}^2$ function satisfying $\mathrm{Hess}(V) > 0$. Then*

$$\mathrm{Var}_\mu(f) \leq \int \nabla f \cdot \mathrm{Hess}(V)^{-1} \nabla f \, d\mu - \frac{\left(\int V f \, d\mu - \int V \, d\mu \int f \, d\mu \right)^2}{n - \mathrm{Var}_\mu(V)} \tag{52}$$

for all $\mathcal{C}^1$ compactly supported functions f.

Remark 4.2 *V. H. Nguyen [35] has proven that $\mathrm{Var}_\mu(V) \leq n$ for V convex. We will observe in the proof that even $Var_\mu(V) < n$ as soon as $\mathrm{Hess}(V) > 0$. In fact, it follows from the bound (52) for $f = V$ that $Var_\mu(V) \leq \frac{nI}{n+I} < n$ where $I = \int \nabla V \cdot \mathrm{Hess}(V)^{-1} \nabla V \, d\mu$. In particular, if $R \, \mathrm{Id}_n \leq \mathrm{Hess}(V) \leq S \, \mathrm{Id}_n$, then $I \leq R^{-1} \int |\nabla V|^2 d\mu = R^{-1} \int \Delta V d\mu \leq nS/R$ and $\mathrm{Var}_\mu(V) \leq \frac{nS}{R+S}$. The latter inequality is an equality (to $n/2$) for the Gaussian measure with any variance, for which $R = S$.*

If $\mu = \gamma$ is the standard Gaussian measure then (52) is exactly the dimensional (Poincaré) inequality (7) (and in particular equality holds for $f = |x|^2/2$).

In the non Gaussian case, G. Hargé has derived the following improvement of the Brascamp-Lieb inequality, see [31, Th. 1] : if V is a $\mathcal{C}^2$ function satisfying $R \, \mathrm{Id}_n \leq \mathrm{Hess}(V) \leq S \, \mathrm{Id}_n$ for constants $0 \leq R \leq S$, then

$$\mathrm{Var}_\mu(f) \leq \int \nabla f \cdot \mathrm{Hess}(V)^{-1} \nabla f \, d\mu - \frac{1 + R/S}{n} \left(\int V f \, d\mu - \int V \, d\mu \int f \, d\mu \right)^2 \tag{53}$$

for all f.
We do not know in full generality which of the coefficients $(n - Var_\mu(V))^{-1}$ and $n^{-1}(1 + R/S)$ in the corrective terms of (52) and (53) is the larger.
Besides being equal (to $2n^{-1}$) in the Gaussian case, both coefficients are always larger than n^{-1}. More precisely the coefficient in (52) is always strictly larger than n^{-1} whereas the coefficient in (53) is n^{-1} when $R = 0$ (no uniform convexity) or $S = +\infty$ (no upper bound on $\mathrm{Hess}(V)$): hence at least in these cases our bound is stronger.
The bound (53) has been obtained in [31] by a L^2 argument. We shall see in the appendix that it can be formally recovered by linearization in the Monge-Ampère equation.

Proof of Proposition 4.1. Let ω be in the space $\mathcal{C}_c^\infty$ of $\mathcal{C}^\infty$ and compactly supported functions. Then

$$\int ||\mathrm{Hess}(\omega)||_{HS}^2 \, d\mu - \frac{1}{n} \left(\int \Delta \omega \, d\mu \right)^2 \geq 0$$

by the Cauchy-Schwarz inequality; here $||\mathrm{Hess}(\omega)||_{HS}^2 = \sum_{i,j=1}^n (\partial_{ij} \omega)^2$ is the squared Hilbert-Schmidt norm of the matrix $\mathrm{Hess}(\omega) = (\partial_{ij} \omega)_{i,j}$. In other words

$$\int \nabla \omega \cdot \mathrm{Hess}(V) \nabla \omega \, d\mu \leq \int \left(||\mathrm{Hess}(\omega)||_{HS}^2 + \nabla \omega \cdot \mathrm{Hess}(V) \nabla \omega \right) d\mu - \frac{1}{n} \left(\int \Delta \omega \, d\mu \right)^2. \tag{54}$$

Moreover, by integration by parts,

$$\int \left(||\mathrm{Hess}(\omega)||_{HS}^2 + \nabla \omega \cdot \mathrm{Hess}(V) \nabla \omega \right) d\mu = \int (L\omega)^2 d\mu, \qquad \int \Delta \omega \, d\mu = - \int V L \omega \, d\mu \tag{55}$$

with $L = \Delta - \nabla V \cdot \nabla$, see [4, Section 3.2].

Let now f be a $\mathcal{C}^1$ compactly supported function. Then pointwise

$$2\,\nabla f\cdot\nabla\omega \le \nabla\omega\cdot \mathrm{Hess}(V)\nabla\omega + \nabla f\cdot \mathrm{Hess}(V)^{-1}\nabla f.$$

From these remarks, inequality (54) implies

$$2\int \nabla f\cdot\nabla\omega\, d\mu \le \int \nabla f\cdot \mathrm{Hess}(V)^{-1}\nabla f\, d\mu + \int (L\omega)^2 d\mu - \frac{1}{n}\Big(\int V L\omega\, d\mu\Big)^2.$$

Let now $h=-L\omega$. Then $\int \nabla f\nabla\omega\, d\mu = -\int fL\omega\, d\mu = \int fh\, d\mu$ by integration by parts.
To sum up, we have obtained

$$2\int fh\, d\mu \le \int \nabla f\cdot \mathrm{Hess}(V)^{-1}\nabla f\, d\mu + \int h^2 d\mu - \frac{1}{n}\Big(\int Vh\, d\mu\Big)^2 \tag{56}$$

for any h in $L(\mathcal{C}_c^\infty)$ and any $\mathcal{C}^1$ compactly supported function f.
But, by [31, Lem. 9] for instance, $L(\mathcal{C}_c^\infty)$ is dense (for the $L^2(\mu)$ norm) in the space of functions $h\in L^2(\mu)$ such that $\int h d\mu = 0$. Hence, formula (56) extends to any $h\in L^2(\mu)$ such that $\int h\, d\mu = 0$.
In particular, given a $\mathcal{C}^1$ compactly supported function f such that $\int f\, d\mu = 0$, we can apply (56) to $h = f + a\big(V - \int V\, d\mu\big)$ with $a\in\mathbb{R}$. Observe indeed that $V\in L^2(\mu)$ for $\mu = e^{-V}$ with V convex. We get

$$\int f^2\, d\mu \le \int \nabla f\cdot \mathrm{Hess}(V)^{-1}\nabla f\, d\mu + I\, a^2 - 2a\,\frac{\mathrm{Var}_\mu(V)}{n}\int Vf\, d\mu - \frac{1}{n}\Big(\int Vf\, d\mu\Big)^2$$

for all a, where $I = \mathrm{Var}_\mu(V)(n-\mathrm{Var}_\mu(V))/n$. Necessarily I is positive, that is, $\mathrm{Var}_\mu(V) < n$. Indeed, if I was non positive, then the left-hand side would be $-\infty$ by letting a tend to $\pm\infty$, which is impossible. We finally optimise over a, choosing $a = \int Vf\, d\mu/(n - Var_\mu(V))$. This concludes the proof of Proposition 4.1 for any f such that $\int f\, d\mu = 0$, and then for any f. ▷

4.2 Brascamp-Lieb inequality via the Borell-Brascamp-Lieb inequality

The following result gives an improved version of the Brascamp-Lieb inequality (4) from the Borell-Brascamp-Lieb inequality.

Theorem 4.3 (Dimensional Brascamp-Lieb inequality II) *Let μ be a probability measure on $\mathbb{R}^n$ with density e^{-V} where V is a $\mathcal{C}^2$ function satisfying $\mathrm{Hess}(V) > 0$. Then for any $\mathcal{C}^1$ and compactly supported function f such that $\int f d\mu = 0$,*

$$\mathrm{Var}_\mu(f) \le \int \nabla f\cdot \mathrm{Hess}(V)^{-1}\,\nabla f\, d\mu - \int \frac{(f - \nabla f\cdot \mathrm{Hess}(V)^{-1}\,\nabla V)^2}{n + \nabla V\cdot \mathrm{Hess}(V)^{-1}\,\nabla V} d\mu. \tag{57}$$

Theorem 4.3 is proved in Appendix B.

For the standard Gaussian measure, we obtain

Corollary 4.4 *The Gaussian measure γ satisfies the dimensional Poincaré inequality*

$$\mathrm{Var}_\gamma(f) \le \int |\nabla f|^2\, d\gamma - \int \frac{(f - \nabla f\cdot x)^2}{n+|x|^2} d\gamma \tag{58}$$

for any $\mathcal{C}^1$ and compactly supported function f such that $\int f d\gamma = 0$.

By the Cauchy-Schwarz inequality and integration by part,

$$\int \frac{(f-\nabla f\cdot x)^2}{n+|x|^2} d\gamma \ge \frac{\big(\int \nabla f\cdot x d\gamma\big)^2}{2n} = \frac{\big(\int \Delta f d\gamma\big)^2}{2n} = \frac{\big(\int f|x|^2/2 d\gamma\big)^2}{n - Var_\gamma(|x|^2/2)}.$$

Therefore, for the Gaussian measure, inequality (58) is stronger than (7) mentionned in the introduction (and naturally equality still holds for $f = |x|^2/2$).

4.3 Comparison of Brascamp-Lieb inequalities

Many dimensional Brascamp-Lieb inequalities have recently been proved, and should be compared. We have already compared our inequality (52) with G. Hargé's bound, as the same covariance term appears. Let us now compare (57) with other inequalities. It seems difficult to obtain a global comparison and we are only able to give partial answers or hints.

- The present paper proposes the two inequalities (52) and (57). In the Gaussian case we have already observed that (57)-(58) is stronger than (52). A variant of this argument shows that it is also the case for instance when $V(x) = x^{2a} + \beta, x \in \mathbb{R}$ with $a \in \mathbb{N}^*$ and a normalisation constant β. We believe that it is the case for any V since the additional term in (52) vanishes for functions f for which the one in (57) does not.

  In fact, for a $\mathcal{C}^1$ function f such $\int f e^{-V} = 0$, the additional term in (57) vanishes if and only if there exists $a \in \mathbb{R}^n$ such that $f = a \cdot \nabla V$ (and then $a = \int f(x) x e^{-V(x)}$). For, if $f = \nabla f \, \mathrm{Hess}(V)^{-1} \nabla V$ on $\mathbb{R}^n$, then $g(y) = f(\nabla V^*(y))$ solves $g(y) = \nabla g(y) \cdot y$ on $\mathbb{R}^n$. Hence for fixed $y \in \mathbb{R}^n$ the map $t \mapsto g(ty)/t$ is constant; for $t = 1$ and $t \to 0$ this implies $g(y) = \nabla g(0) \cdot y$. This finally gives f, and conversely. But it is classical that these functions f are exactly those for which equality holds in the Brascamp-Lieb inequality (4). Hence the additional term in (57) can be seen as a (weighted) way of measuring the distance of a function to the optimisers in the Brascamp-Lieb inequality (4).

  Very recently, and under the same hypothesis as in Theorem 4.3, D. Cordero-Erausquin in [17, Prop. 6] proved that

$$Var_\mu(f) \leq \int \nabla f \cdot \mathrm{Hess}(V)^{-1} \nabla f \, d\mu - c\lambda(\mu) \int \mathrm{Hess}(V)^{-1} (\mathrm{Hess}(V) + c\lambda(\mu) \mathrm{Id}_n)^{-1} \nabla f_0 \cdot \nabla f_0 d\mu \quad (59)$$

  for all f satisfying $\int f d\mu = 0$; here $f_0 = f - \int y f(y) d\mu(y) \cdot \nabla V$, c is a numerical constant and $\lambda(\mu)$ is the Poincaré constant of the measure μ. The additional term in (59) vanishes if and only if f_0 is a constant, so also appears here as a distance to the optimisers. A quantitative comparison between (57) and (59) can not easily be performed as in particular a numerical constant appears in (59). After the present work was completed, M. Arnaudon, M. Bonnefont and A. Joulin [2] have derived Brascamp-Lieb inequalities in which the energy has been modified, instead of keeping the original energy and allowing for a remainder term, as here. We could not compare their results with ours.

- We now turn to the Gaussian case when $\mu = \gamma$. We have already observed that (57) is stronger that (52), which is exactly (7). On the other hand, (8) is a purely spectral inequality. We have numerically checked that (57) implies (8) for the Hermite polynomial functions H_k, $k \in \{1, \cdots, 7\}$. We believe that it is the case for all functions, but we do not have a proof of it.

  Let us conclude by mentioning the inequality

$$Var_\gamma(f) \leq 6 \int |\nabla f|^2 d\gamma - 6 \int \frac{(\nabla f \cdot x)^2}{n + |x|^2} d\gamma.$$

  has been proved in [11, Sect. 2]. Their extremal functions have been lost since there is no equality when $f(x) = a \cdot x$ and the constant in front of the energy is larger than in our bounds.

A Proof of Theorem 1.1

Optimality of inequality (15). When W is strictly convex and satisfies (H1), then (14) holds, so that

$$\int \frac{W^*(\nabla W)}{W^{n+1}} \, dx = \int \frac{\nabla W \cdot x - W}{W^{n+1}} \, dx = -\frac{1}{n} \int \nabla(W^{-n}) \cdot x \, dx - 1 = 0.$$

In the last equality, we used an integration by parts, valid from hypothesis (H1) satisfied by W. This gives the equality case in (15) when $g = W$.

Proof of inequality (15). Globally, the proof follows [10], but for completeness we give its main points. It is based on a Taylor expansion of the inequality $\int H_t dx \geq 1$, when $t = 1 - s$ goes to 1, and where H_t

is defined in (16). Equivalently, this inequality can be written as

$$\int t^{-n}\left(\inf_{h\in\mathbb{R}^n}\left\{g\left(\frac{z}{t}-\frac{s}{t}h\right)+\frac{s}{t}W(h)\right\}\right)^{-n}dz\geq 1.$$

Changing variables in the integral by letting $x=z/t$, and letting $u=s/t$, the inequality becomes $\int\varphi_u^{-n}\,dx\geq 1$ for any $u>0$, where for positive u

$$\varphi_u(x)=\inf_{h\in\mathbb{R}^n}\big\{g\big(x-uh\big)+uW(h)\big\}.$$

Since $\int g^{-n}dx=1$, this is

$$\int\frac{\varphi_u^{-n}-g^{-n}}{u}dx\geq 0$$

for any $u>0$. The main goal is now to consider the limit as $u\to 0$, by computing the limit

$$\lim_{u\to 0}\int\frac{\varphi_u^{-n}-g^{-n}}{u}dx. \tag{60}$$

Lemma A.1 *For any $x\in\mathbb{R}^n$,*

$$\lim_{u\to 0^+}\frac{\varphi_u(x)-g(x)}{u}=-W^*(\nabla g(x)).$$

Proof

$\lhd$ For any $x\in\mathbb{R}^n$, from the definition of φ_u, we have for any $h\in\mathbb{R}^n$,

$$\frac{\varphi_u(x)-g(x)}{u}\leq\frac{g(x-uh)-g(x)}{u}+W(h)=-\nabla g(x)\cdot h+W(h)+o(u).$$

It follows that $\limsup_{u\to 0^+}\frac{\varphi_u(x)-g(x)}{u}\leq-\nabla g(x)\cdot h+W(h)$ for any h, and then by taking the infimum over $h\in\mathbb{R}^n$,

$$\limsup_{u\to 0^+}\frac{\varphi_u(x)-g(x)}{u}\leq -W^*(\nabla g(x)).$$

Now, one can observe that

$$\varphi_u(x)=\inf_{h,\ uW(h)\leq g(x)}\big\{g\big(x-uh\big)+uW(h)\big\},$$

so that

$$\frac{g(x)-\varphi_u(x)}{u}=\sup_{h,\ uW(h)\leq g(x)}\left\{\frac{g(x)-g(x-uh)}{u}-W(h)\right\}\leq\sup_{h,\ uW(h)\leq g(x)}\{\nabla g(x)\cdot h+|h|\varepsilon(u|h|)-W(h)\}, \tag{61}$$

where ε is an appropriate function satisfying $\lim_{u\to 0}\varepsilon(u)=0$.
Let now $r=\sup\{u|h|,\ uW(h)\leq g(x)\}$. From the hypothesis (H1),

$$r\leq\sup\left\{u|h|;\ \frac{u|h|^2}{2C}\leq g(x)\right\}\leq D\sqrt{ug(x)}, \tag{62}$$

where D is a constant. Generally, D denotes a constant and can change from line to line. The bound (62) gives

$$\frac{g(x)-\varphi_u(x)}{u}\leq\sup_{h,\ uW(h)\leq g(x)}\Big\{\nabla g(x)\cdot h+|h|\varepsilon(D\sqrt{ug(x)})-W(h)\Big\}.$$

Let now $\eta>0$. Then there exists $u_0>0$ such that $\forall u\in(0,u_0]$, $\varepsilon\left(D\sqrt{ug(x)}\right)\leq\eta$, so that

$$\frac{g(x)-\varphi_u(x)}{u}\leq\sup_{h,\ uW(h)\leq g(x)}\{\nabla g(x)\cdot h+|h|\eta-W(h)\}\leq\sup_{h\in\mathbb{R}^n}\{\nabla g(x)\cdot h+|h|\eta-W(h)\}.$$

By (H1) the supremum is reached, say on a ball of center 0 and radius $R > 0$ independent of $\eta < 1$. Hence

$$\frac{g(x) - \varphi_u(x)}{u} \leq \sup_{h \in \mathbb{R}^n} \{\nabla g(x) \cdot h - W(h)\} + R\eta = W^*(\nabla g(x)) + R\eta.$$

The result follows by taking the superior limit and then letting η go to 0. $\triangleright$

To compute the limit (60), we use the dominated convergence theorem. Since everywhere $\frac{\varphi_u^{-n} - g^{-n}}{u}$ goes to $n\frac{W^*(\nabla g)}{g^{n+1}}$ when $u \to 0$, we only need to give a uniform bound (in u) of the quantity of $\frac{\varphi_u^{-n} - g^{-n}}{u}$.

For any $0 < a \leq b$, the following holds $|a^{-n} - b^{-n}| \leq n|a - b|a^{-1-n}$. Since $0 \leq \varphi_u(x) \leq g(x)$ by definition of φ_u, we can apply this inequality to $a = \varphi_u(x)$ and $b = g(x)$, obtaining

$$\left|\frac{\varphi_u(x)^{-n} - g(x)^{-n}}{u}\right| \leq n \left|\frac{\varphi_u(x) - g(x)}{u}\right| \varphi_u(x)^{-1-n}.$$

Bound on $|\varphi_u(x) - g(x)|/u$:
First, from the equality in (61) and a Taylor expansion,

$$\frac{g(x) - \varphi_u(x)}{u} \leq \sup_{h,\ uW(h) \leq g(x)} \{\mathcal{D}g(x, u|h|)|h| - W(h)\},$$

where $\mathcal{D}g(x, s) = \sup_{|x-y| \leq s} |\nabla g(y)|$. We assume now that $u \in]0, 1]$. Then, from (62), $r \leq D\sqrt{g(x)}$. Hence, by (H1),

$$\frac{g(x) - \varphi_u(x)}{u} \leq \sup_{h,\ uW(h) \leq g(x)} \left\{\mathcal{D}g\left(x, D\sqrt{g(x)}\right)|h| - W(h)\right\} \leq \sup_{h \in \mathbb{R}^n} \left\{\mathcal{D}g\left(x, D\sqrt{g(x)}\right)|h| - \frac{|h|^2}{2C}\right\}.$$

The explicit computation of the infimum gives

$$\frac{g(x) - \varphi_u(x)}{u} \leq D\ \mathcal{D}g\left(x, D\sqrt{g(x)}\right)^2.$$

Then, from the hypothesis (H2), the estimation of $\mathcal{D}g$ gives the bound

$$0 \leq \frac{g(x) - \varphi_u(x)}{u} \leq D(|x|^2 + 1), \qquad u \in]0, 1], x \in \mathbb{R}^n.$$

Bound on $\varphi_u(x)$:
From the hypotheses (H1) and (H2) we have

$$\varphi_u(x) \geq \inf_h \left\{\frac{1}{C}|x - uh|^2 + u\frac{|h|^2}{2C}\right\} + \frac{1}{C}.$$

When $u \in]0, 1]$, the explicit computation of the infimum gives again

$$\varphi_u(x) \geq D(|x|^2 + 1).$$

Finally, we have obtained the upper bound

$$\left|\frac{\varphi_u(x)^{-n} - g(x)^{-n}}{u}\right| \leq D(|x|^2 + 1)^{-n}, \qquad u \in]0, 1], x \in \mathbb{R}^n.$$

The dominated convergence theorem can then be applied. The proof of Theorem 1.1 is then complete.

B Proof of Theorem 4.3

We adapt the argument of [9].
We will assume throughout the proof that V is $\mathcal{C}^3$ with bounded derivatives $\nabla^2 V$ and $\nabla^3 V$, and that there exists $\rho > 0$ such that uniformly in $\mathbb{R}^n$, $\mathrm{Hess}(V) \geq \rho\,\mathrm{Id}_n$. Then the result extends to V as in the Theorem by approximation.

Let f be a $\mathcal{C}^1$ compactly supported function satisfying $\int f d\mu = 0$. We apply the Borell-Brascamp-Lieb inequality (13) for $t = s = 1/2$, $F = \exp(-V)$, $G = \exp(2\delta f - V)/Z_\delta$ ($\delta > 0$) where $Z_\delta = \int \exp(2\delta f) d\mu$, and finally $H = \exp(\phi_\delta - V)$ where

$$\phi_\delta(z) = -n \log \inf_{h \in \mathbb{R}^n} \left\{ Z_\delta^{1/n} \exp\left(-\frac{2\delta}{n} f(z+h) + \frac{V(z+h)}{n} \right) + \exp\left(\frac{V(z-h)}{n} \right) \right\} + n\log(2) + V(z). \quad (63)$$

Then (13) ensures that $\int e^{\phi_\delta} d\mu \geq 1$. The rest of the proof is devoted to a Taylor expansion of $\int \exp(\phi_\delta) d\mu$ as δ goes to 0.

By convexity of V, for any $\delta > 0$ the function in (63) to be minimised is coercive, so indeed admits a (possibly non unique) minimiser, which we first estimate by giving a Taylor expansion as $\delta \to 0$.
For this, let $\delta > 0$ be given and let h_δ be any minimiser. Then

$$Z_\delta^{1/n} \Big(-2\delta \nabla f(z+h_\delta) + \nabla V(z+h_\delta) \Big) \exp\left(-\frac{2\delta}{n} f(z+h_\delta) + \frac{1}{n} V(z+h_\delta) \right) = \nabla V(z-h_\delta) \exp\left(\frac{1}{n} V(z-h_\delta) \right). \quad (64)$$

1. In a first step we prove that $h_\delta = O(\delta)$ uniformly in z: in other words, there exists a constant $C > 0$ such that for any δ small enough and any $z \in \mathbb{R}^n$,

$$|h_\delta| \leq C\,\delta.$$

For this, first, since $\int f d\mu = 0$ and f is compactly supported,

$$Z_\delta^{1/n} = \left(\int e^{2\delta f} d\mu \right)^{1/n} = 1 + \frac{2\delta^2}{n} \int f^2 d\mu + o(\delta^2). \quad (65)$$

Let us now assume that the support of f is included in the ball $\{|x| < R\}$ with $R > 0$. There are two cases, depending on whether $|z + h_\delta| \geq R$ or $|z + h_\delta| \leq R$.

- **First, assume that** $|z + h_\delta| \geq R$**.** Then equation (64) becomes

  $$Z_\delta^{1/n} \nabla V(z+h_\delta) \exp\left(\frac{1}{n} V(z+h_\delta) \right) = \nabla V(z-h_\delta) \exp\left(\frac{1}{n} V(z-h_\delta) \right).$$

  In other words, by (65) and taking the scalar product by h_δ,

  $$\Big(\frac{2\delta^2}{n} \int f^2 d\mu + o(\delta^2) \Big) \nabla\Phi(z+h_\delta) \cdot h_\delta = \nabla\Phi(z-h_\delta) \cdot h_\delta - \nabla\Phi(z+h_\delta) \cdot h_\delta$$

  where $\Phi = \exp(\frac{1}{n} V)$, that is,

  $$-\Big(\frac{2\delta^2}{n} \int f^2 d\mu + o(\delta^2) \Big) \nabla V(z+h_\delta) \cdot h_\delta = e^{-V(z+h_\delta)} \int_{-1}^{1} h_\delta \cdot \mathrm{Hess}(\Phi)(z + t h_\delta) h_\delta dt.$$

  Now $\mathrm{Hess}(V) \geq \rho\,\mathrm{Id}_n$ so

  $$\mathrm{Hess}(\Phi) = \frac{e^V}{n} (\mathrm{Hess}(V) + \frac{1}{n} \nabla V \otimes \nabla V) \geq \frac{e^V}{n} \big(\rho\,\mathrm{Id}_n + \frac{1}{n} \nabla V \otimes \nabla V \big). \quad (66)$$

Hence

$$-\Big(2\delta^2\int f^2 d\mu + o(\delta^2)\Big)\nabla V(z+h_\delta)\cdot h_\delta \geq e^{-V(z+h_\delta)}\int_{-1}^{1}\Big(\rho|h_\delta|^2+\frac{1}{n}|\nabla V(z+th_\delta)\cdot h_\delta|^2\Big)e^{V(z+th_\delta)}dt. \tag{67}$$

In particular, $\nabla V(z+h_\delta)\cdot h_\delta \leq 0$ on the left-hand side for δ small enough, independently of z since the $o(\delta^2)$ comes from Z_δ, see (65), and is uniform in z; hence for any $t\in[-1,1]$

$$V(z+th_\delta)-V(z+h_\delta)\geq (t-1)\nabla V(z+h_\delta)\cdot h_\delta \geq 0$$

by convexity of V. Moreover $\nabla V(z+th_\delta)\cdot h_\delta \leq \nabla V(z+h_\delta)\cdot h_\delta$ again by convexity, whence

$$|\nabla V(z+th_\delta)\cdot h_\delta| \geq |\nabla V(z+h_\delta)\cdot h_\delta|.$$

Collecting all terms, (67) leads to

$$\Big(2\delta^2\int f^2 d\mu + o(\delta^2)\Big)|\nabla V(z+h_\delta)\cdot h_\delta| \geq 2\rho|h_\delta|^2+\frac{2}{n}|\nabla V(z+h_\delta)\cdot h_\delta|^2 \geq 4\sqrt{\frac{\rho}{n}}|h_\delta|\,|\nabla V(z+h_\delta)\cdot h_\delta|.$$

for δ small enough, and where the $o(\delta^2)$ is uniform in z. Hence there exists a constant $A>0$ such that

$$|h_\delta| \leq A\,\delta^2,$$

for any δ small enough and any z, whenever $|z+h_\delta|\geq R$.

- **Assume now that** $|z+h_\delta|\leq R$**.** Let us write equation (64) as

$$\begin{aligned}\nabla\Phi(z+h_\delta)-\nabla\Phi(z-h_\delta) &= \Big[1-Z_\delta^{1/n}\exp\Big(-\frac{2\delta}{n}f(z+h_\delta)\Big)\Big]\nabla\Phi(z+h_\delta)\\ &\quad +2\delta Z_\delta^{1/n}\exp\Big(-\frac{2\delta}{n}f(z+h_\delta)+\frac{1}{n}V(z+h_\delta)\Big)\nabla f(z+h_\delta).\end{aligned}$$

Then f, V and their gradients are continuous and then uniformly bounded on the ball $\{|x|\leq R\}$, so by (65) there exists a constant A such that for all δ small enough and all z with $|z+h_\delta|\leq R$

$$\Big|\nabla\Phi(z+h_\delta)-\nabla\Phi(z-h_\delta)\Big| \leq A\,\delta.$$

Hence, by the Cauchy-Schwarz inequality and the bound $\mathrm{Hess}(\Phi)\geq \frac{\rho}{n}e^V \mathrm{Id}_n$, a consequence of (66),

$$A\delta|h_\delta| \geq \Big(\nabla\Phi(z+h_\delta)-\nabla\Phi(z-h_\delta)\Big)\cdot h_\delta = \int_{-1}^{1} h_\delta\cdot\mathrm{Hess}(\Phi)(z+th_\delta)h_\delta\,dt \geq \frac{2\rho}{n}e^{\min V}|h_\delta|^2.$$

By uniform convexity the function V is indeed bounded from below on $\mathbb{R}^n$, so there exists a constant B such that for all δ small enough and all z with $|z+h_\delta|\leq R$

$$|h_\delta|\leq B\,\delta.$$

All cases being covered, our **first step** is completed.

2. In a second step we perform a first-order Taylor expansion of the equality (64). For z fixed, it gives

$$-\delta\nabla f(z)+\mathrm{Hess}(V)(z)h_\delta-\frac{\delta}{n}f(z)\nabla V(z)+\frac{h_\delta\cdot\nabla V(z)}{n}\nabla V(z)+o_z(\delta)=0, \tag{68}$$

where $o_z(\delta)$ depends on z, δ and h_δ. Since $|h_\delta|\leq C\,\delta$ by the first step, uniformly in z, one deduces from (68) that

$$|o_z(\delta)|\leq A\,\delta^2\,(|\mathrm{Hess}(V)(z)|+|\nabla V(z)|^2+1)$$

for a constant A and for any z.

In the sequel we let $H(z)$ denote positive polynomial functions in $V(z)$, $\nabla V(z)$, etc., independent of δ small and which can change from line to line. The latter inequality can then be written as

$$|o_z(\delta)| \leq \delta^2 H(z). \tag{69}$$

Let now $X = \nabla f \cdot \mathrm{Hess}(V)^{-1}\nabla V$ and $Y = \nabla V \cdot \mathrm{Hess}(V)^{-1}\nabla V$. Taking the scalar product of (68) with $\mathrm{Hess}(V)^{-1}\nabla V$ one gets

$$h_\delta \cdot \nabla V = \delta \frac{X + \frac{fY}{n}}{1 + \frac{Y}{n}} + o_z(\delta)$$

at the point z, where $o_z(\delta)$ satisfies (69) since in particular $\mathrm{Hess}(V)^{-1} \leq \rho^{-1}\,\mathrm{Id}_n$. Then, again by (68),

$$h_\delta = \delta \left[\mathrm{Hess}(V)^{-1}\nabla f + \frac{\mathrm{Hess}(V)^{-1}\nabla V}{n} \frac{f - X}{1 + \frac{Y}{n}} \right] + o_z(\delta)$$

where again $o_z(\delta)$ satisfies (69).

We now compute the second-order Taylor expansion of the function ϕ_δ. First, from the expansion $e^x = 1 + x + x^2/2 + x^3 e^{\theta x}/6$ with $\theta \in (0,1)$, we have at the point z,

$$\phi_\delta = -n \log(1 + \psi_\delta)$$

with

$$\psi_\delta = -\frac{\delta}{n} f - \frac{\delta h_\delta \cdot \nabla f}{n} + \frac{h_\delta \cdot \mathrm{Hess}(V) h_\delta}{2n} + \frac{\delta^2}{n^2} f^2 + \frac{(h_\delta \cdot \nabla V)^2}{2n^2} - \frac{\delta f}{n^2} h_\delta \cdot \nabla V + \frac{\delta^2}{n} \int f^2 d\mu + \bar{o}_z(\delta^2).$$

Here $\bar{o}_z(\delta^2)$ now satisfies

$$|\bar{o}_z(\delta^2)| \leq \delta^3 H(z) \exp\left(\delta^3 K_3(z)\right) \tag{70}$$

with $|K_3(z)| \leq A(|\nabla V| + |\nabla^2 V| + |\nabla^3 V|)$ for an universal constant A.

We now observe that for small δ one has $\phi_\delta(z) \leq n \log 2$ for all z, that is, $\psi_\delta(z) \geq -1/2$. Indeed, for small δ one has

$$Z_\delta^{1/n} \exp\left(-\frac{2\delta}{n} f(x) \right) \geq \frac{1}{2}$$

uniformly in $x \in \mathbb{R}^n$, by (65) and since f is bounded from above. Hence for any $h \in \mathbb{R}^n$

$$\begin{aligned} &Z_\delta^{1/n} \exp\left(-\frac{2\delta}{n} f(z+h) + \frac{V(z+h)}{n} \right) + \exp\left(\frac{V(z-h)}{n} \right) \\ &\quad \geq \frac{1}{2} \exp\left(\frac{V(z+h)}{n} \right) + \exp\left(\frac{V(z-h)}{n} \right) \geq \frac{1}{2} \Big[\exp\left(\frac{V(z+h)}{n} \right) + \exp\left(\frac{V(z-h)}{n} \right) \Big] \geq e^{V(z)/n} \end{aligned}$$

by convexity of $e^{V/n}$. The bound on ϕ_δ follows by its definition (63).

Now from the expansion $(1+x)^{-n} = 1 - nx + n(n+1)x^2/2 - n(n+1)(n+2)x^3(1+\theta x)^{-n-3}/6$ with $\theta \in (0,1)$ and (68), we get

$$(1 + \psi_\delta)^{-n} = 1 + \delta f + \frac{h_\delta \cdot \mathrm{Hess}(V) h_\delta}{2} + \delta^2 \frac{n-1}{2n} f^2 + \frac{(h_\delta \cdot \nabla V)^2}{2n} - \delta^2 \int f^2 d\mu + \bar{o}_z(\delta^2)$$

for a $\bar{o}_z(\delta^2)$ satisfying (70): here we use that $\psi_\delta(z) \geq -1/2$ so that $1 + \theta\psi_\delta \geq 1/2$ in the Taylor expansion, uniformly in z and δ. The above expressions of h_δ and $h_\delta \cdot \nabla V$ finally give

$$(1 + \psi_\delta)^{-n} = 1 + \delta f + \frac{\delta^2}{2} \nabla f \cdot \mathrm{Hess}(V)^{-1} \nabla f - \frac{\delta^2}{2} \frac{(f - X)^2}{n + Y} + \frac{\delta^2}{2} f^2 - \delta^2 \int f^2 d\mu + \bar{o}_z(\delta^2).$$

In conclusion, by integration the second-order Taylor expansion of the Borell-Brascamp-Lieb inequality $\int (1 + \psi_\delta)^{-n} d\mu = \int e^{\phi_\delta} d\mu \geq 1$ implies

$$\int f^2 d\mu \leq \int \nabla f \cdot \mathrm{Hess}(V)^{-1} \nabla f d\mu - \int \frac{(f - X)^2}{n + Y} d\mu$$

for all $\mathcal{C}^1$ compactly supported f such that $\int f d\mu = 0$. Here we use that $\delta^{-2} \int \bar{o}_z(\delta^2) e^{-V(z)} dz \to 0$ as $\delta \to 0$ by (70), since the right-hand side in (70) is in $L^1(e^{-V})$ by our hypotheses on V. By definition of X and Y this concludes the argument.

C Link with G. Hargé's bound (53)

In this Appendix, we observe that G. Hargé's bound (53) can be formally recovered by linearization in the Monge-Ampère equation (28). Let indeed f be a smooth function such that $\int f\,d\mu = 0$, and $\mu_2 = (1+\varepsilon f)\mu$ for $\varepsilon > 0$, and expand the transport map $\nabla\varphi(x)$ sending $\mu_1 = \mu$ onto μ_2 as $x+\varepsilon\nabla\theta_1(x)+\varepsilon^2\nabla\theta_2(x)+o(\varepsilon^2)$. Taking logarithms in (28) with such μ_1 and μ_2 and observing that

$$\log\det(\mathrm{Hess}(\varphi)) = \log\det\Big(I+\varepsilon\mathrm{Hess}(\theta_1)+\varepsilon^2\mathrm{Hess}(\theta_2)+o(\varepsilon^2)\Big) = \varepsilon\Delta\theta_1+\varepsilon^2\Delta\theta_2-\frac{\varepsilon^2}{2}\mathrm{tr}\big[(\mathrm{Hess}(\theta_1))^2\big]+o(\varepsilon^2),$$

a second-order Taylor expansion ensures that $f = -L\theta_1$ in the first-order terms; moreover

$$f^2 = -\nabla\theta_1\cdot\mathrm{Hess}(V)\nabla\theta_1 + 2L\theta_2 + 2\nabla f\cdot\nabla\theta_1 - \mathrm{tr}\big[(\mathrm{Hess}(\theta_1))^2\big]$$

in the second-order terms. Assume now that $\mathrm{Hess}(V) > 0$, and let $M = \mathrm{Hess}(V)^{1/2} > 0$. Then

$$-\nabla\theta_1\cdot\mathrm{Hess}(V)\nabla\theta_1 + 2\nabla f\cdot\nabla\theta_1 = |M^{-1}\nabla f|^2 - |M\nabla\theta_1 - M^{-1}\nabla f|^2$$

so that

$$\int f^2\,d\mu = \int \nabla f\cdot\mathrm{Hess}(V)^{-1}\nabla f\,d\mu - \int\Big(|M\nabla\theta_1 - M^{-1}\nabla f|^2 + \mathrm{tr}\big[(\mathrm{Hess}(\theta_1))^2\big]\Big)\,d\mu \tag{71}$$

by integration. At this point one recognizes terms in the proof of [31, Th. 1] : one observes that $f = -L\theta_1$ so $\nabla f = M^2\theta_1 - X$ by differentiation, where $X\in\mathbb{R}^n$ is the vector with coordinates $L(\partial_i\theta_1)$; hence

$$|M\nabla\theta_1 - M^{-1}\nabla f|^2 = |M^{-1}X|^2 \geq \frac{1}{S}|X|^2$$

if moreover $\mathrm{Hess}(V)\leq S$. In particular

$$\int |M\nabla\theta_1 - M^{-1}\nabla f|^2\,d\mu \geq \frac{1}{S}\sum_i\int\Big(L(\partial_i\theta_1)\Big)^2\,d\mu \geq \frac{R}{S}\sum_{i,j}\int\big(\partial^2_{ji}\theta_1\big)^2\,d\mu$$

by (55), if $\mathrm{Hess}(V)\geq R\,\mathrm{Id}_n$. Hence

$$\int\Big(|M\nabla\theta_1 - M^{-1}\nabla h|^2 + \mathrm{tr}\big[(\mathrm{Hess}(\theta_1))^2\big]\Big)d\mu \geq \Big(1+\frac{R}{S}\Big)\sum_{i,j}\int\big(\partial^2_{ji}\theta_1\big)^2 d\mu \geq \frac{1}{n}\Big(1+\frac{R}{S}\Big)\left(\int\Delta\theta_1\,d\mu\right)^2 \tag{72}$$

since moreover by the Cauchy-Schwarz inequality

$$\left(\int\Delta\theta_1\,d\mu\right)^2 = \left(\sum_i\int\partial_{ii}\theta_1\,d\mu\right)^2 \leq n\sum_i\left(\int\partial_{ii}\theta_1\,d\mu\right)^2 \leq n\sum_{i,j}\left(\int\partial_{ij}\theta_1\,d\mu\right)^2.$$

By (71) and (72) we finally recover (53) since by integration by parts and (55)

$$\int\Delta\theta_1\,d\mu = \int\nabla\theta_1\cdot\nabla V\,e^{-V}\,dx = -\int L\theta_1\,V\,e^{-V}\,dx = \int f\,V\,d\mu.$$

Acknowledgements. The authors are grateful to both referees for a careful reading of the manuscript and helpful comments and questions which improved the paper. In particular they raised most relevant issues. This work was partly written while the authors were visiting Institut Mittag-Leffler in Stockholm; it is a pleasure for them to thank this institution for its kind hospitality and participants for discussions on this and related works. This research was supported by the French ANR-12-BS01-0019 STAB project.

References

[1] L. Ambrosio, N. Gigli and G. Savaré. Gradient flows in metric spaces and in the space of probability measures. Lectures in Math. ETH Zürich. Birkhäuser, Basel, 2008.

[2] M. Arnaudon, M. Bonnefont and A. Joulin. Intertwinings and generalized Brascamp-Lieb Inequalities. To appear on Rev. Mat. Iberoam.

[3] D. Bakry, F. Bolley, and I. Gentil. Dimension dependent hypercontractivity for Gaussian kernels. Prob. Theor. Rel. Fields, 154(3): 845–874, 2012.

[4] D. Bakry, I. Gentil and M. Ledoux. Analysis and geometry of Markov diffusion operators, volume 348 of Grund. Math. Wiss. Springer, Berlin, 2014.

[5] D. Bakry and M. Ledoux. A logarithmic Sobolev form of the Li-Yau parabolic inequality. Rev. Mat. Iberoam., 22(2):683–702, 2006.

[6] F. Barthe and A. Kolesnikov. Mass transport and variants of the logarithmic Sobolev inequality. J. Geom. Anal., 18(4):921–979, 2008.

[7] S. G. Bobkov, I. Gentil and M. Ledoux. Hypercontractivity of Hamilton-Jacobi equations. J. Math. Pures Appl. (9), 80(7):669–696, 2001.

[8] S. Bobkov, N. Gozlan, C. Roberto and P.-M. Samson. Bounds on the deficit in the logarithmic Sobolev inequality. J. Funct. Anal., 267:4110–4138, 2014.

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

[10] S. Bobkov and M. Ledoux. From Brunn-Minkowski to sharp Sobolev inequalities. Ann. Mat. Pura Appl. (4), 187(3):369–384, 2008.

[11] S. Bobkov and M. Ledoux. Weighted Poincaré-type inequalities for Cauchy and other convex measures. Ann. Probab., 37(2):403–427, 2009.

[12] F. Bolley, I. Gentil and A. Guillin. Convergence to equilibrium in Wasserstein distance for Fokker-Planck equations. J. Funct. Anal., 263(8):2430–2457, 2012.

[13] F. Bolley, I. Gentil and A. Guillin. Dimensional contraction via Markov transportation distance. J. London Math. Soc., 90(1):309–332, 2014.

[14] F. Bolley, I. Gentil, A. Guillin and K. Kuwada. Equivalence between dimensional contractions in Wasserstein distance and curvature-dimension condition. To appear on Annali della Scuola Norm. Sup. di Pisa.

[15] E. Carlen. Superadditivity of Fisher's information and logarithmic Sobolev inequalities. J. Funct. Anal., 101:194–211, 1991.

[16] D. Cordero-Erausquin. Some applications of mass transport to Gaussian type inequalities. Arch. Rat. Mech. Anal., 161:257–269, 2002.

[17] D. Cordero-Erausquin. Transport inequalities for log-concave measures, quantitative forms and applications. To appear on Canadian J. Math.

[18] S. Daneri and G. Savaré. Lecture notes on gradient flows and optimal transport. Optimal transportation, London Math. Soc. Lecture Note Ser., 413:100–144, 2014.

[19] M. Del Pino and J. Dolbeault. The optimal Euclidean L^p-Sobolev logarithmic inequality. J. Funct. Anal., 197(1):151–161, 2003.

[20] R. Eldan. A two-sided estimate for the Gaussian noise stability. Invent. Math. 201(2):561–624, 2015.

[21] M. Erbar, K. Kuwada and K.-T. Sturm. On the equivalence of the entropic curvature-dimension condition and Bochner's inequality on metric measure spaces. Invent. Math. 201(3):993–1071, 2015.

[22] M. Fathi, E. Indrei and M. Ledoux. Quantitative logarithmic Sobolev inequalities and stability estimates. Disc. Cont. Dynamical Syst., 36:6835–6853, 2016.

[23] A. Figalli and E. Indrei. A sharp stability result for the relative isoperimetric inequality inside convex cones. J. Geom. Anal., 23:938–969, 2013.

[24] A. Figalli and D. Jerison. Quantitative stability for the Brunn-Minkowski inequality. To appear on Adv. Math.

[25] A. Figalli, F. Maggi and A. Pratelli. A mass transportation approach to quantitative isoperimetric inequalities. Invent. Math., 182:167–211, 2010.

[26] A. Figalli, F. Maggi and A. Pratelli. Sharp stability theorems for the anisotropic Sobolev and log-Sobolev inequalities on functions of bounded variation. Adv. Math., 242:80–101, 2013.

[27] I. Gentil. The general optimal L^p-Euclidean logarithmic Sobolev inequality by Hamilton-Jacobi equations. J. Funct. Anal. 202(2):591–599, 2003.

[28] I. Gentil. From the Prékopa-Leindler inequality to modified logarithmic Sobolev inequality. Ann. Fac. Sci. Toulouse, 6, 17(2):291–308, 2008.

[29] I. Gentil. Dimensional contraction in Wasserstein distance for diffusion semigroups on a Riemannian manifold. Potential Anal., 42(4):861–873, 2015.

[30] L. Goldstein, I. Nourdin and G. Peccati. Gaussian phase transitions and conic intrinsic volumes: Steining the Steiner formula. Preprint, 2014.

[31] G. Hargé. Reinforcement of an inequality due to Brascamp and Lieb. J. Funct. Anal., 254(2):267–300, 2008.

[32] E. Indrei and D. Marcon. A quantitative log-Sobolev inequality for a two parameter family of functions. Int. Math. Res. Not., 20:5563–5580, 2014.

[33] S. Lisini. Nonlinear diffusion equations with variable coefficients as gradient flows in Wasserstein spaces. ESAIM Contr. Opt. Calc. Var., 15:712–740, 2009.

[34] R. McCann. A convexity principle for interacting gases. Adv. Math., 128(1):153–179, 1997.

[35] V. H. Nguyen. Dimensional variance estimates of Brascamp-Lieb type and a local approach to dimensional Prékopa theorem. J. Funct. Anal., 266(2):931–955, 2014.

[36] F. Otto and C. Villani. Generalization of an inequality by Talagrand and links with the logarithmic Sobolev inequality. J. Funct. Anal., 173(2):361–400, 2000.

[37] M. Talagrand. Transportation cost for Gaussian and other product measures. Geom. Funct. Anal., 6(3):587–600, 1996.

[38] C. Villani. Topics in Optimal transportation, volume 58 of Grad. studies in math. Amer. Math. Soc, Providence, 2003.

[39] C. Villani. Optimal transport, Old and new, volume 338 of Grund. Math. Wiss. Springer, Berlin, 2009.