



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. 2015. hal-01171361v2

HAL Id: hal-01171361

<https://hal.science/hal-01171361v2>

Preprint submitted on 5 Oct 2015 (v2), last revised 12 Mar 2017 (v4)

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‡

October 5, 2015

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 behaviour 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 with constant $R > 0$ (see [3] for instance) if for all probability measures ν in \mathbb{R}^n , absolutely continuous with respect to μ ,

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

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

$$H(\nu|\mu) = Ent_\mu(f) = \int f \log f d\mu \quad \text{and} \quad I(\nu|\mu) = \int \frac{|\nabla f|^2}{f} d\mu.$$

- A probability measure μ in \mathbb{R}^n satisfies a Talagrand transportation inequality [35] with constant $R > 0$ if for all ν absolutely continuous with respect to μ

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

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

†Institut Camille Jordan, Umr Cnrs 5208, Université Claude Bernard Lyon 1, Lyon, France. gentil@math.univ-lyon1.fr

‡Institut Universitaire de France and 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], [36]).

By the Otto-Villani Theorem [34], the logarithmic Sobolev inequality (1) implies the Talagrand inequality (2) with the same constant (see also [6], [36, 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 ,

$$\text{Var}_{\mu}(f) \leq \int \nabla f \cdot \text{Hess}(V)^{-1} \nabla f d\mu. \quad (3)$$

Here $\text{Var}_{\mu}(f) = \int f^2 d\mu - (\int f d\mu)^2$ is the variance of f under the measure μ , see [3, 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), (2) and (3) with $R = 1$. In fact, in the Gaussian case, the Brascamp-Lieb inequality (3) can be obtained from (1) by linearisation, namely by taking $\nu = f\mu$ with f close to 1. Let us note that in this case $\text{Hess}(V) = \text{Id}_n$, and the Brascamp-Lieb inequality becomes the Poincaré inequality. Moreover these inequalities are optimal for the Gaussian measure: by direct computation, equality holds in (1) and (2) 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 (3) for $f(x) = b \cdot x$, $b \in \mathbb{R}^n$ (see [3, Chap. 4 and 5]).

Inequalities (1), (2) and (3) 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 of 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

$$H(\nu|\gamma) \leq \frac{1}{2} \int |x|^2 d\nu - \frac{n}{2} + \frac{n}{2} \log \left(1 + \frac{1}{n} \left(I(\nu|\mu) + n - \int |x|^2 d\nu \right) \right) \quad (4)$$

of the logarithmic Sobolev inequality (1) has been obtained by D. Bakry and M. Ledoux [4] by self-improvement from the Euclidean logarithmic Sobolev inequality, or by semigroup arguments on the Euclidean heat semigroup (see also [3, Sect. 6.7.1] and the early work [13] by E. Carlen). The dimensional improvement

$$W_2^2(\nu, \gamma) \leq \int |x|^2 d\nu + n - 2n \exp \left(\int \frac{|x|^2}{2n} d\nu - \frac{1}{2} - \frac{1}{n} H(\nu|\mu) \right) \quad (5)$$

of the Talagrand inequality (2) has been derived in [2]; 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 [4] that linearising (4) leads to the dimensional improvement

$$\text{Var}_{\gamma}(f) \leq \int |\nabla f|^2 d\gamma - \frac{1}{2n} \left(\int (|x|^2 - n) f d\gamma \right)^2, \quad (6)$$

of the Brascamp-Lieb (or Poincaré) inequality (3) for the Gaussian measure (see also [3, Sect. 6.7.1]). On the other hand, by a spectral analysis of the Ornstein-Uhlenbeck semigroup, the bound

$$\text{Var}_{\gamma}(f) \leq \frac{1}{2} \int |\nabla f|^2 d\gamma + \frac{1}{2} \left| \int \nabla f d\gamma \right|^2 \quad (7)$$

has been established in [28, Sect. 6.2]. By the Cauchy-Schwarz inequality, it improves upon (6). Naturally, both inequalities (6) and (7) 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 (4), (5) and (7) 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), (2), (3) of the inequalities: namely, for the logarithmic Sobolev (*LSI* in short) and Talagrand (*Tal* in short) inequalities, lower bounds on the quantities

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

The problem of dimensional refinement 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 [23] (see also R. Eldan [18] or [21] for convex cones). Further recent results have been obtained on deficit in the logarithmic Sobolev inequality in the Gaussian case by Figalli-Maggi-Pratelli [24], Indrei-Marcon [30] and Bobkov&al [7]. In particular [7] rediscovers (4) and extends earlier results obtained in dimension one by Barthe-Kolesnikov [5] on the Talagrand deficit. Fathi-Indrei-Ledoux [20] also considers these deficits, particularly emphasizing the case where ν has additional properties, mainly a Poincaré inequality ensuring better constants in the logarithmic Sobolev inequality. Very recently D. Cordero-Erausquin [15] has studied refinements of the Talagrand and Brascamp-Lieb inequalities via optimal transport tools.

In this paper we prove dimensional improvements of logarithmic Sobolev, Talagrand and Brascamp-Lieb inequality via multiple tools. Let us quote for example C. Villani [36, 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.

In many cases we will compare our inequalities with other recent extensions. 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 (2) as soon as the potential V satisfies $\text{Hess}(V) \geq R \text{Id}_n$ with $R > 0$, as symmetric matrices. One of the goal of this paper is to extend the above dimensional inequalities under this condition with $R > 0$ or only $\text{Hess}(V) > 0$. Applications to concentration inequalities and short and long time behaviour for the laws of solutions to stochastic differential equations are also given.

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 [8, 9] by S. Bobkov and M. Ledoux. The method is based on a general convexity inequality given in Theorem 1.1.

In section 2 we propose a dimensional Talagrand inequality through optimal transportation in the spirit of Barthe-Kolesnikov [5] and D. Cordero-Erausquin [14] or the recent [15]. In Section 2.1 we apply our new Talagrand inequality to dimensional concentration inequalities.

Inspired by recent results on the equivalence between contraction and $CD(R, n)$ curvature dimension condition in abstract measure spaces (see [1, 19, 12]), in Section 3 we consider applications to refined dimensional contraction properties under $CD(R, \infty)$; we shall see how the dimension improves the asymptotic behaviour for the laws of solutions to stochastic differential equations (in the spirit of [10, 11]). Using the terminology of the Γ_2 -condition, the associated Markov generator $L = \Delta - \nabla V \cdot \nabla$ does not satisfy a $CD(R, n)$ 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.

In section 4 we prove two kinds of dimensional Brascamp-Lieb inequalities. Both will follow from a linearisation argument, in the optimal transport approach and via the Borell-Brascamp-Lieb inequality. They will be compared with other very recent dimensional refinements of the Brascamp-Lieb inequalities.

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 $\mu(f) = \int f d\mu$.

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 , and let $s, t \geq 0$ be fixed such that $t + s = 1$. Under the hypothesis

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

for any $x, y \in \mathbb{R}^n$, the Prékopa-Leindler inequality ensures that

$$\int H dx \geq \left(\int F dx \right)^t \left(\int G dx \right)^s$$

(see [36, Chap. 19] for instance). It appears as a functional form of the Brunn-Minkowski inequality: for any bounded measurable sets A and B in \mathbb{R}^n ,

$$\text{vol}(tA + sB) \geq \text{vol}(A)^t \text{vol}(B)^s.$$

The Borell-Brascamp-Lieb inequality is a stronger and dimensional form of the Prékopa-Leindler inequality. Let F, G and H be positive measurable functions on \mathbb{R}^n and let $s, t > 0$ be fixed such that $s + t = 1$. If $\int F dx = \int G dx = 1$ and

$$H(tx + sy) \geq \left(tF(x)^{-1/n} + sG(y)^{-1/n} \right)^{-n} \quad (8)$$

for any $x, y \in \mathbb{R}^n$, then the Borell-Brascamp-Lieb inequality asserts that $\int H dx \geq 1$ (see again [36]).

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 [8, 9] (see also [26] 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. Proofs are based on Taylor expansions when $s \rightarrow 0$ or $t \rightarrow 0$.

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.

Here we let W^* be the Legendre transform of a function W on \mathbb{R}^n , defined for $x \in \mathbb{R}^n$ by

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

Theorem 1.1 (Convexity inequality) *Let g, W be smooth positive functions on \mathbb{R}^n such that $g(x), W(x), |\nabla g(x)|, |\nabla W(x)| \rightarrow \infty$ when $|x| \rightarrow \infty$. If $\int g^{-n} dx = \int W^{-n} dx = 1$, then*

$$\int \frac{W^*(\nabla g)}{g^{n+1}} dx \geq 0, \quad (9)$$

with equality if $g = W$ and is convex.

Proof

◁ The proof is based on a Taylor expansion of the Borell-Brascamp-Lieb inequality (8) when $s = 1 - t$ goes to 0.

Let $F = g^{-n}$ and $G = W^{-n}$ in (8), hence satisfying $\int F dx = \int G dx = 1$. Then the map H_t defined by

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

for $z \in \mathbb{R}^n$ satisfies $\int H_t dx \geq 1$. Let us now compute the first-order Taylor expansion of H_t when $s \rightarrow 0$, or equivalently $t \rightarrow 1$. We obtain

$$\begin{aligned} H_t(z)^{-1/n} &= \inf_{h \in \mathbb{R}^n} \left\{ (1-s)(g(z) + sh \cdot \nabla g(z)) + sW(z - h) + o(s) \right\} \\ &= \inf_{h \in \mathbb{R}^n} \left\{ g(z) + s(W(z - h) + h \cdot \nabla g(z) - g(z)) + o(s) \right\}, \end{aligned}$$

where the $o(s)$ can be chosen uniformly in z . Changing h into $z - u$ in the infimum, for given z , we get

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

and therefore

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

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 (9). \triangleright

The function W^* is convex, but we do not assume that so is W . 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 $W = e^{V/n}$. Then, for any $a \in \mathbb{R}$ and $y \in \mathbb{R}^n$,

$$W^*(y) \leq \frac{1}{n} e^a V^*(ne^{-a}y) + (a-1)e^a \quad (10)$$

by convexity of the exponential function.

1.2 Euclidean logarithmic Sobolev inequalities

As a warm up, let us first see how to quickly recover the classical Euclidean logarithmic Sobolev inequality, by starting from (10). Let $C : \mathbb{R}^n \rightarrow \mathbb{R}^+$ be a strictly convex function, and let us apply (10) with $V = C + \beta$ and $W = e^{V/n}$; here $\beta = \log \int e^{-C} dx$ so that $\int e^{-V} dx = 1$. Let also f a function such that $\int e^f dx = 1$, $g = e^{-f/n}$ and $a = -f/n + u$ where u is a real constant. Then $V^* = C^* - \beta$ and (10) can be written as

$$W^*(\nabla g) \leq \frac{1}{n} e^{-f/n+u} [C^*(-\nabla f e^{-u}) - \beta - f + nu - n].$$

Theorem 1.1 applied with $g = e^{-f/n}$ and $W = e^{V/n}$, for which $\int W^{-n} dx = \int g^{-n} dx = 1$, gives

$$\int f e^f dx \leq n(u-1) - \beta + \int C^*(-e^{-u} \nabla f) e^f dx$$

for all smooth functions f such that $\int e^f dx = 1$, and all u in \mathbb{R} .

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 $C^*(-e^{-u} \nabla f) = e^{-pu} C^*(-\nabla f)$. Optimising over u leads to

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

for any smooth function f such that $\int e^f dx = 1$.

Hence, we recover the optimal L^p -Euclidean log Sobolev inequality proved in [17] and [25].

1.3 Dimensional logarithmic Sobolev inequalities

In this section we consider again a probability measure μ with density e^{-V} and the function $W = e^{V/n}$, and a function f such that $\int e^f d\mu = 1$. Inequality (10) applied with $g = \exp(\frac{V-f}{n})$ and $a = \frac{V-f}{n} + u$ with $u \in \mathbb{R}$ gives

$$W^*(\nabla g) \leq e^{\frac{V-f}{n}+u} \left[\frac{1}{n} V^*(e^{-u} \nabla(V-f)) + \frac{V-f}{n} + u - 1 \right].$$

Hence

$$\int (V^*(e^{-u} \nabla(V-f)) + V - f + n(u-1)) e^f e^{-V} dx \geq 0$$

by Theorem 1.1, and then

$$\text{Ent}_\mu(e^f) \leq \int [V^*(s\nabla(V-f)) + V]e^f d\mu - n(1 + \log s) \quad (11)$$

for any f such that $\int e^f d\mu = 1$, and any $s = e^{-u} > 0$.

For $s = 1$ this simplifies as

$$\text{Ent}_\mu(e^f) \leq \int [V^*(\nabla(V-f)) + V - n]e^f d\mu, \quad \int e^f d\mu = 1.$$

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,

$$\text{Ent}_\gamma(e^f) \leq \frac{1}{2} \int |\nabla f|^2 e^f d\gamma, \quad \int e^f d\gamma = 1.$$

More generally, if V is a strictly convex function on \mathbb{R}^n , then $V(x) = \nabla V(x) \cdot x - V^*(\nabla V(x))$, and by integration by parts we recover the modified logarithmic Sobolev inequality

$$\text{Ent}_\mu(e^f) \leq \int [V^*(\nabla(V-f)) + x \cdot \nabla f - V^*(\nabla V)]e^f d\mu, \quad \int e^f d\mu = 1$$

proved by the second author in [26].

We now optimise over the parameter $s > 0$ to obtain dimensional logarithmic Sobolev inequalities in two diverse settings.

We first consider a q -homogeneous function V , that is, a p -homogeneous function V^* , with $1/p + 1/q = 1$.

Corollary 1.2 (L^p -dimensional logarithmic Sobolev inequality) *Let V_0 be a q -homogeneous function with $q > 1$, and let $\beta = \log \int e^{-V_0} dx$. Let μ be the probability measure with density $e^{-V_0 - \beta}$. Then*

$$\text{Ent}_\mu(e^f) \leq \frac{n}{p} \log \left(\frac{p}{n} \int V_0^*(\nabla(V_0 - f))e^f d\mu \right) + n \frac{1-p}{p} + \int V_0 e^f d\mu \quad (12)$$

for any smooth function f such that $\int e^f d\mu = 1$; here $1/p + 1/q = 1$.

For $V_0(x) = |x|^2/2$, then μ is the standard Gaussian measure γ , and (12) with $p = 2$ is exactly the dimensional Gaussian logarithmic Sobolev inequality (4) written with e^f instead of f .

Assume now that V_0 is only almost q -homogeneous, that is, there exists $c > 0$ such that

$$V_0(\lambda x) \geq c^{1-q} \lambda^q V_0(x)$$

for any $\lambda > 0$ and $x \in \mathbb{R}^n$. Equivalently V^* is almost p -homogeneous: for the same c

$$V_0^*(\alpha y) \leq c\alpha^p V_0^*(y)$$

for any $\alpha > 0$ and $y \in \mathbb{R}^n$. Then, by the same computation as in Corollary 1.2,

$$\text{Ent}_\mu(e^f) \leq \frac{n}{p} \log \left(\frac{p}{n} \int cV_0^*(\nabla(V_0 - f))e^f d\mu \right) + n \frac{1-p}{p} + \int V_0 e^f d\mu,$$

for any smooth function f such that $\int e^f d\mu = 1$.

We now assume uniform *convexity* on V . Suppose that V is \mathcal{C}^2 with $\text{Hess}(V) \geq R \text{Id}_n$ for $R > 0$. Then, for their inverse matrices, $\text{Hess}(V^*) \leq R^{-1} \text{Id}_n$ 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 $\nabla V^*(\nabla V(x)) = x$ and $V^*(\nabla V(x)) + V(x) = x \cdot \nabla V(x)$. For $z = s \nabla(V - f)$ at point x , and by (11), this leads to

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

By integration by parts we finally obtain:

Corollary 1.3 (Dimensional logarithmic Sobolev under Γ_2 -condition) *Let μ be a probability measure with density e^{-V} where V is \mathcal{C}^2 with $\text{Hess}(V) \geq R \text{Id}_n$ for $R > 0$. Then*

$$\text{Ent}_\mu(e^f) \leq n(s - 1 - \log s) + \frac{1}{2R} \int |(1 - s)\nabla V + s\nabla f|^2 e^f d\mu \quad (13)$$

for any $s > 0$ and any smooth function f such that $\int e^f d\mu = 1$.

When $s = 1$, we recover the classical logarithmic Sobolev inequality (1) under the Bakry-Émery condition. Moreover, as in (24) or (15) below for the Talagrand inequality, the bound (13) can be written as a deficit in the log Sobolev inequality.

Let us observe that the right-hand side in (13) 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 (4).

We will see in Section 3.1 that (13) 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 $\text{Hess}(V) \geq R \text{Id}_n$ with $R > 0$. Then for all $\nu \in P_2(\mathbb{R}^n)$*

$$\frac{R}{2} W_2^2(\mu, \nu) \leq \nu(V) - \mu(V) + n - n \exp \left[\frac{1}{n} \left(\nu(V) - \mu(V) - H(\nu|\mu) \right) \right]. \quad (14)$$

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

$$\delta_{\text{Tal}}(\nu|\mu) \geq \max \left\{ \delta_n \left(H(\nu|\mu) + \mu(V) - \nu(V) \right), \Lambda_n \left(\nu(V) - \mu(V) - \frac{R}{2} W_2^2(\nu, \mu) \right) \right\}. \quad (15)$$

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 (14) implies the classical Talagrand inequality (2) under the condition $\text{Hess}(V) \geq R \text{Id}_n$. When μ is the standard Gaussian measure γ on \mathbb{R}^n , then $R = 1$ and we recover the dimensional Talagrand inequality (5).

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 $\nu(V) \leq \mu(V)$,*

$$\delta_{\text{Tal}}(\nu|\mu) \geq \delta_n(H(\nu|\mu)) \geq \frac{1}{e} \min \left(H(\nu|\mu), \frac{H(\nu|\mu)^2}{n} \right). \quad (16)$$

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; it has been introduced and proved in [34] and [14] under the Bakry-Émery condition.

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 $\text{Hess}(V) \geq R \text{Id}_n$ 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*

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

Recall that the Fisher information I has been introduced in (1). For $g = 1$ and $\nu = f\mu$, this bound can be written as the dimensional *HWI* inequality

$$n \exp \left[\frac{1}{n} \left(H(\nu|\mu) + \mu(V) - \nu(V) \right) \right] - n \leq \mu(V) - \nu(V) + W_2(\mu, \nu) \sqrt{I(\nu|\mu)} - \frac{R}{2} W_2^2(\mu, \nu). \quad (17)$$

As in (15) for the Talagrand inequality, this can equivalently be written as a deficit in the *HWI* inequality. It is classical that a *HWI* inequality implies a logarithmic Sobolev inequality (see [34] for instance). Likewise, from (17), one can obtain a dimension dependent logarithmic Sobolev inequality. We refer to Section 2.4 for further details.

2.1 An application to concentration

Let us quickly revisit K. Marton's argument for concentration via Talagrand's inequality (as in [36, Chap. 22] for instance) and see how the refined inequality (14) in Theorem 2.1 gives sharpened information for large deviations.

Let $d\mu = e^{-V} dx$ satisfy inequality (14). 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

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

by (14), or its weaker form (2). 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 (14) again we get, for $r > c_A$,

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

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 $r > c_A$*

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

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

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

of the Talagrand inequality (2), see again [36, 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 $\text{Hess}(V) \geq \text{Id}_n$, so by Corollary 2.4 with $R = 1$ there exists a constant $C = C(p, n)$ such that for all $r > C$

$$\mu(|x| > r + 1) = \mu(A_r) \leq \exp \left[c_V - V_r + n \log(1 + V_r/n) \right].$$

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

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

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

Deficit in the Gaussian Talagrand inequality (for $\mu = \gamma$) and for centered measures ν has been investigated in one dimension in [5] and [7], in the form

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

Here the c 's are diverse numerical constants, the infimum runs over couplings π of γ and ν , and W_1 is the Wasserstein distance between one-dimensional measures, for the cost $|y - x|$, $x, y \in \mathbb{R}$.

This second lower bound has been extended in [20, Th. 5] to higher dimension, as

$$\delta_{\text{Tal}}(\nu|\gamma) \geq c \min \left(\frac{W_{1,1}(\nu, \gamma)^2}{n}, \frac{W_{1,1}(\nu, \gamma)}{\sqrt{n}} \right) \quad (18)$$

as soon as ν has mean 0 ; here c is a numerical constant independent of the dimension n and

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

Still under a centering condition, the bound (18) has been improved in [15, Prop. 3] by replacing the quantity $W_{1,1}/\sqrt{n}$ by the larger W_2 , and extended to reference measures μ with density e^{-V} where $\text{Hess}(V) \geq R \text{Id}_n$.

In comparison, our bound (15) 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 [36, Chap. 22]), and the relative entropy can be much larger than the weak distance W_2 .

As considered in [15] and [20], 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 $\delta_{\text{Tal}}(\nu^N|\mu^N) = N \delta_{\text{Tal}}(\nu|\mu)$ by tensorisation properties of both H and W_2^2 . However, the above bound (18) in [15] (so with W_2 instead of $W_{1,1}/\sqrt{n}$) leads to

$$\delta_{\text{Tal}}(\nu^N|\mu^N) \geq c \min \left(NW_2(\nu, \mu)^2, \sqrt{N} W_2(\nu, \mu) \right);$$

it has the good order in N only for small W_2 -perturbations ν of the reference measure μ .

On the contrary, 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

$$H(\nu^N|\mu^N) + \mu^N(V^N) - \nu^N(V^N) = N (H(\nu|\mu) + \mu(V) - \nu(V));$$

hence Theorem 2.1 leads to

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

which has the correct order in N .

2.3 Proof of Theorem 2.3

Theorem 2.3 is a consequence of the relation

$$H(h\mu|\mu) - \mu(hV) = Ent_{dx}(he^{-V}) \quad (19)$$

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 two smooth functions such that $f\mu$ and $g\mu$ belong to $P_2(\mathbb{R}^n)$. Let φ be a convex map on \mathbb{R}^n such that (the Brenier map) $\nabla\varphi$ transports $f\mu$ onto $g\mu$. Then*

$$\begin{aligned} \int V g d\mu - \int V f d\mu - \int (\nabla\varphi - x) \cdot \nabla f d\mu &\geq n \exp \left[\frac{1}{n} \left(Ent_{dx}(fe^{-V}) - Ent_{dx}(ge^{-V}) \right) \right] - n \\ &+ \int_0^1 \int (\nabla\varphi(x) - x) \cdot \text{Hess}(V)(x + t(\nabla\varphi(x) - x)) (\nabla\varphi(x) - x)(1-t) dt f(x) d\mu(x). \end{aligned}$$

Indeed, if $\text{Hess}(V) \geq R \text{Id}_n$, 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, the last term in the left-hand side is bounded by $W_2\sqrt{I}$ by the Cauchy-Schwarz inequality. This implies Theorem 2.3.

Proof

◁ By the Taylor formula,

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

for almost every x in \mathbb{R}^n . We now integrate with respect to $f\mu$ and use the relation

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

Here $\Delta\varphi$ is the trace of the Hessian $\text{Hess}(\varphi)$ of φ in the sense of Alexandrov, see [14] and [32]. This leads to

$$\begin{aligned} \int V g d\mu - \int V f d\mu - \int (\nabla\varphi - x) \cdot \nabla f d\mu &\geq \int \Delta\varphi f d\mu - n \\ &+ \int_0^1 \int (\nabla\varphi(x) - x) \cdot \text{Hess}(V)(x + t(\nabla\varphi(x) - x)) (\nabla\varphi(x) - x)(1-t) dt f(x) d\mu(x). \end{aligned} \quad (20)$$

Then Lemma 2.5 is a consequence of the following Lemma. ▷

Lemma 2.6 *Let $\mu_1, \mu_2 \in P_2(\mathbb{R}^n)$ absolutely continuous with respect to the Lebesgue measure, with respective densities also denoted μ_1 and μ_2 . Let φ be the convex map on \mathbb{R}^n such that $\nabla\varphi$ transports μ_1 onto μ_2 . Then*

$$\int \Delta\varphi d\mu_1 \geq n \exp \left[\frac{Ent_{dx}(\mu_1) - Ent_{dx}(\mu_2)}{n} \right]. \quad (21)$$

Here $\Delta\varphi$ is the trace of the Hessian of φ in the sense of Alexandrov.

Proof

◁ By [32] the Monge-Ampère equation

$$\mu_1(x) = \mu_2(\nabla\varphi(x)) \det(\text{Hess}(\varphi)(x)) \quad (22)$$

holds a.e. in \mathbb{R}^n , in the sense of Alexandrov. Taking logarithms and integrating with respect to μ_1 lead to

$$Ent_{dx}(\mu_1) = Ent_{dx}(\mu_2) + \int \log \det(\text{Hess}(\varphi)) d\mu_1. \quad (23)$$

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

$$\int \log \det(\text{Hess}(\varphi)) d\mu_1 = n \frac{1}{n} \sum_i \int \log(\varphi_i) d\mu_1 \leq n \log \left(\int \frac{1}{n} \sum_i \varphi_i d\mu_1 \right) = n \log \left(\frac{1}{n} \int \Delta \varphi d\mu_1 \right).$$

This concludes the proof. \triangleright

Remark 2.7 *In the Gaussian case, we have already observed that translations of the Gaussian measure are extremals. As observed in [14], 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 $\text{Hess}(V) \geq R \text{Id}_n$, then equality in the Talagrand inequality implies that the potential is necessarily Gaussian and that extremals are translations of the Gaussian measure.

2.4 Logarithmic Sobolev inequalities by transport

As observed in [34], a HWI inequality classically implies a logarithmic Sobolev inequality by bounding from above the second order polynomial in W_2 in HWI by its maximum. Likewise, the dimensional HWI inequality (17) 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 $\text{Hess}(V) \geq R \text{Id}_n$ with $R > 0$. Then

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

for all ν . Equivalently, in terms of deficit,

$$\delta_{LSI}(\nu|\mu) \geq R \max \left\{ \delta_n \left(\nu(V) - \mu(V) - H(\nu|\mu) \right), \Lambda_n \left(\frac{I(\nu|\mu)}{2R} - \nu(V) + \mu(V) \right) \right\}. \quad (24)$$

In the Gaussian case, then $R = 1$ and we obtain a bound which is slightly worse than (4), 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 [7, Th. 1.1]. As observed by the authors, the HWI inequality in [7] does not seem either to imply (4). We could not compare the HWI in [7] to our bound (17) in full generality. However, if $\nu(|x|^2) = n = \gamma(|x|^2)$ then they can respectively be written as

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

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 [7, 20] it is also possible to get refined logarithmic Sobolev inequalities by combining the HWI and Talagrand inequalities. Here, if $\text{Hess}(V) \geq R \text{Id}_n$ with $R > 0$, then (17) can be written as

$$H + \delta_n(-h) \leq W_2 \sqrt{I} - \frac{R}{2} W_2^2 \quad (25)$$

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

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

Then, by (25) 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 (24). 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 [7] and [30], measures ν satisfying a (weaker) Poincaré inequality as in [20]. 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 [30], the deficit is controlled by the Wasserstein for the class of centered function with upper and lower bounded curvature. The authors in [7] also give new bounds in terms of conditionally centered vectors. Further improvements are given in [20] 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 $\text{Hess}(V) \geq R \text{Id}_n$, 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

$$\frac{\partial u_t}{\partial t} = \Delta u_t + \nabla \cdot (u_t \nabla V), \quad t > 0, x \in \mathbb{R}^n \quad (26)$$

as in [1, Chap. 11.2] and [16, Th. 4.20 and 4.21]. 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 .

Under the assumption $\text{Hess}(V) \geq R \text{Id}_n$, the map $H(\cdot|\mu)$ is geodesically R -convex on $P_2(\mathbb{R}^n)$. Moreover, the interpretation of (26) 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 [36, Chap. 24]). Let u and v be solutions to (26). Then

$$H(u_t|\mu) \leq \frac{W_2^2(u_0, \mu)}{2t} e^{2 \max\{-R, 0\} t}, \quad t > 0 \quad (27)$$

and

$$W_2(u_t, v_t) \leq e^{-Rt} W_2(u_0, v_0), \quad t \geq 0. \quad (28)$$

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

$$W_2(u_t, \mu) \leq e^{-Rt} W_2(u_0, \mu), \quad t \geq 0. \quad (29)$$

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 (26) is given by the Mehler formula (see [3, 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 $\sigma_t^2 = 1 - e^{-2t}$:

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

Its relative entropy can be computed as

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

Of course this is coherent with (27), with $R = 0$, since

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

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 (26), 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

$$\frac{d}{dt} \int |x|^2 du_t = 2n - 2 \int |x|^2 du_t. \quad (30)$$

In particular, in the notation $h(t) = H(u_t|\gamma)/n$ and $i(t) = I(u_t|\gamma)$, the dimensional Gaussian logarithmic Sobolev inequality (4) simplifies as $2h \leq \log(1 + i)$. Hence

$$h'(t) = -i(t) \leq 1 - e^{2h(t)}, \quad t > 0.$$

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

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

which gives the same short-time behaviour.

More generally :

Proposition 3.1 *Let u be a solution to (26) with $\text{Hess}(V) \geq R \text{Id}_n$, $R > 0$, 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*

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

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

◁ We shall let c denote diverse positive constants depending only on n, M and R . By Corollary 1.3 applied to the measure $e^f dx = u_t$, 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 all $t \geq 0$ and $s > 0$. Recall that I has been introduced in (1). 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 $H := H(u_t|\mu) \geq 1$, then $I := I(u_t|\mu) \geq 2R$ so that $s = \sqrt{2R/I}$ is smaller than 1. For this s we obtain

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

Hence

$$\frac{d}{dt} H = -I \leq -e^{2H/n-c}.$$

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. \triangleright

3.2 Refined contraction properties

Let us now see how to make (28) finer. Writing (26) as the continuity equation

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

with $\xi[u_t] = -\nabla V - \nabla \log u_t$, there holds

$$\begin{aligned} -\frac{1}{2} \frac{d}{dt} W_2^2(u_t, v_t) &= \int (\xi[v_t](\nabla \varphi_t(x)) - \xi[u_t](x)) \cdot (\nabla \varphi_t(x) - x) u_t(x) dx \\ &\geq \int [\Delta \varphi_t(x) + \Delta \varphi_t^*(\nabla \varphi_t(x)) - 2n + (\nabla V(\nabla \varphi_t(x)) - \nabla V(x)) \cdot (\nabla \varphi_t(x) - x)] u_t(x) dx \end{aligned}$$

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 . The equality follows from [36, Th. 23.9] (see also [1, Th. 8.4.7]); its assumptions are satisfied since (and likewise for v)

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

which is finite for any $t_2 > t_1 > 0$, as observed above. The inequality follows from a weak integration by parts, as in [31, Th. 1.5]; there $\Delta \varphi_t$ is the trace of the Alexandrov Hessian of φ_t .

Now, for given t and x , if $\text{Hess}(\varphi_t)(x)$ has the n positive eigenvalues $e^{2\lambda_i(x)}$, for $i = 1, \dots, n$, then its inverse matrix $\text{Hess}(\varphi_t^*)(\nabla \varphi_t(x))$ has eigenvalues $e^{-2\lambda_i(x)}$; hence at point x

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

Hence, by convexity of \sinh^2 and the Jensen inequality, and (23),

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

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

$$-\frac{1}{2} \frac{d}{dt} W_2^2(u_t, v_t) \geq 4n \sinh^2 \left(\frac{\text{Ent}_{dx}(v_t) - \text{Ent}_{dx}(u_t)}{2n} \right) + R W_2^2(u_t, v_t). \quad (31)$$

By time integration this ensures the following dimensional contraction property :

Proposition 3.3 *In the above notation, if $\text{Hess}(V) \geq R \text{Id}_n$ for $R \in \mathbb{R}$, then for any solutions to (26)*

$$W_2^2(u_t, v_t) \leq e^{-2Rt} W_2^2(u_0, v_0) - 8n \int_0^t e^{-2R(t-s)} \sinh^2 \left(\frac{\text{Ent}_{dx}(v_s) - \text{Ent}_{dx}(u_s)}{2n} \right) ds, \quad t \geq 0. \quad (32)$$

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 (32) has been derived in [11] and [12], and is also a consequence of [19]. 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 (32) 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 $\text{Hess}(V) \geq R \text{Id}_n$ should be replaced by the monotonicity condition $(A(y) - A(x)) \cdot (y - x) \geq R|y - x|^2$ for all x, y (see [10] for this non-gradient case).*

3.3 A gradient flow argument to Proposition 3.3

For all t let again $\nabla \varphi_t$ 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 : for each s the measure μ_t^s is the image measure of μ by the map $\nabla \varphi_t^s(x) = x + s(\nabla \varphi_t(x) - x)$. Then, formally and following [36, Chap. 23],

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

where for given $t > 0$ we let

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

We now use the following classical property: the map $\psi : s \mapsto \text{Ent}_{dx}(\mu_t^s)$ satisfies $\psi'' \geq \psi'^2/n$ on $[0, 1]$. Hence, by (34) in Lemma 3.5 below

$$E_t'(1) - E_t'(0) \geq 4n \sinh^2 \left(\frac{\text{Ent}_{dx}(v_t) - \text{Ent}_{dx}(u_t)}{2n} \right) + \int (\nabla V(\nabla \varphi_t(x)) - \nabla V(x)) \cdot (\nabla \varphi_t(x) - x) du_t$$

This leads to (31) and then (32) as soon as $\text{Hess}(V) \geq R \text{Id}_n$.

The following elementary lemma gives additional information to [19, Lem. 2.2].

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

- $\psi'' \geq \psi'^2/n$;
- for all r, s in $[0, 1]$,

$$n - \psi'(r)(s - r) \geq n e^{\frac{\psi(r) - \psi(s)}{n}}; \quad (33)$$

- for all r, s in $[0, 1]$,

$$(\psi'(s) - \psi'(r))(s - r) \geq 4n \sinh^2 \left(\frac{\psi(s) - \psi(r)}{2n} \right). \quad (34)$$

Remark 3.6 *Observe that (33) in Lemma 3.5 for $\psi(s) = \text{Ent}_{dx}(\mu^s)$, $r = 0$ and $s = 1$ leads to (21) in Lemma 2.6. Indeed*

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

Proof of Lemma 3.5. Let indeed $U = e^{-\psi/n}$, so that

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

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 , which is (33) when multiplying both sides by $e^{\psi(r)/n}$.

Adding (33) with the corresponding bound obtained with r, s instead of s, r leads to (34). Conversely, dividing (34) by $(s-r)^2$ and letting s go to r gives $\psi'' \geq \psi'^2/n$ at point r . \triangleright

Let us now recall why for given t the map $\psi : s \mapsto \text{Ent}_{dx}(\mu_t^s)$ satisfies $\psi'' \geq \psi'^2/n$ on $[0,1]$. For this, let t be fixed and let us drop the dependence on t . Let also $\theta_i := \theta_i(x)$ be the eigenvalues of $\text{Hess}(\varphi)(x) - I$ and let us write (23) with the measures $\mu_1 = \mu$ and $\mu_2 = \mu^s$. We obtain

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

Hence

$$\psi'(s) = - \sum_i \int \frac{\theta_i}{1 + s\theta_i} d\mu \tag{35}$$

and then by the Cauchy-Schwarz inequality

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

Remark 3.7 Identity (35) can also be checked by using the continuity equation

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

solved by $(\mu^s)_{s \in [0,1]}$. Here the vector field v^s satisfies $v^s(\nabla \varphi^s(x)) = \nabla \varphi(x) - x$. Indeed

$$\psi'(s) = \frac{d}{ds} \int \mu^s \log \mu^s dx = - \int \nabla \cdot (v^s \mu^s) \log \mu^s dx = - \int \nabla \cdot v^s d\mu^s = - \int (\nabla \cdot v^s)(\nabla \varphi^s(x)) d\mu(x)$$

by integration by parts. Identity (35) follows since by chain rule

$$(\nabla \cdot v^s)(\nabla \varphi^s(x)) = \text{tr} \left[(H\varphi(x) - I)(I + s(H\varphi(x) - I))^{-1} \right] = \sum_i \frac{\theta_i}{1 + s\theta_i}.$$

3.4 Improved convergence rates

In this section we consider a solution u to (26) in the Gaussian case where $\mu = \gamma$, and for which we can take $R = 1$ above. Let us see how the contraction property (32) can make the convergence estimate (28) 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 (30). Hence (19) and the Talagrand inequality (14) ensure that $0 \leq W_2^2(u_t, \gamma) < 2n$ and

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

In particular the right-hand side is non negative. Moreover, for the stationary solution $v_t = v_0 = \gamma$, the contraction property (32) with $R = 1$, in the form (31), implies

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

where $x(t) = W_2^2(u_t, \gamma)/(2n) \in [0, 1]$. 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,

$$x(t) \leq 1 - \left(1 - (2x(0) - x(0)^2)e^{-2t}\right)^{\frac{1}{2}}.$$

By the lower bound

$$1 - (2x(0) - x(0)^2)e^{-2t} \geq (1 - x(0)e^{-2t})^2 \quad (36)$$

it implies the classical bound (28). It also improves it: this can be seen for instance by writing it as

$$W_2^2(u_t, \gamma) \leq W_2^2(u_0, \gamma)e^{-2t} \frac{2 - x(0)}{1 + \left(1 - (2x(0) - x(0)^2)e^{-2t}\right)^{\frac{1}{2}}} \leq W_2^2(u_0, \gamma)e^{-2t} \frac{1 - W_2^2(u_0, \gamma)/(4n)}{1 - W_2^2(u_0, \gamma)e^{-2t}/(4n)}$$

by (36), where the last quotient is smaller than 1.

Remark 3.8 *The Gaussian assumption is used here only to ensure uniform convexity of the potential (hence the Talagrand inequality), and that $\int V du_t \leq \int V e^{-V}$ as soon as this holds at $t = 0$.*

4 Brascamp-Lieb inequalities

It is classical that linearising 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 linearisation in the optimal transport Lemma 2.5, and a second one by linearisation in the Borell-Brascamp-Lieb inequality (8).

4.1 Brascamp-Lieb inequality by optimal transport 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 C^2 function satisfying $\text{Hess}(V) > 0$. Then*

$$\text{Var}_\mu(f) \leq \int \nabla f \cdot \text{Hess}(V)^{-1} \nabla f d\mu - \frac{\left(\int V f d\mu - \int V d\mu \int f d\mu\right)^2}{n - \text{Var}_\mu(V)} \quad (37)$$

for all smooth functions f .

We will observe in the proof that $\text{Var}_\mu(V) < n$ as soon as $\text{Hess}(V) > 0$. In fact, it follows from the bound (37) for $f = V$ that $\text{Var}_\mu(V) \leq \frac{nI}{n+I} < n$ where $I = \int \nabla V \cdot \text{Hess}(V)^{-1} \nabla V d\mu$. In particular, if $R\text{Id}_n \leq \text{Hess}(V) \leq S\text{Id}_n$, then $I \leq R^{-1} \int |\nabla V|^2 d\mu = R^{-1} \int \Delta V d\mu \leq nS/R$ and $\text{Var}_\mu(V) \leq \frac{nS}{R+S}$. Equality holds (to $n/2$) for the Gaussian measure with any variance, for which $R = S$.

If $\mu = \gamma$ is the standard Gaussian measure then (37) is exactly the dimensional (Poincaré) inequality (6) (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 [29, Th. 1] : if V is a C^2 function satisfying $R\text{Id}_n \leq \text{Hess}(V) \leq S\text{Id}_n$ for constants $0 \leq R \leq S$, then

$$\text{Var}_\mu(f) \leq \int \nabla f \cdot \text{Hess}(V)^{-1} \nabla f d\mu - \frac{1 + R/S}{n} \left(\int V f d\mu - \int V d\mu \int f d\mu\right)^2 \quad (38)$$

for all f .

We do not know in full generality which of the coefficients $(n - \text{Var}_\mu(V))^{-1}$ and $n^{-1}(1 + R/S)$ in the corrective terms of (37) and (38) 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 (37) is always strictly larger than n^{-1} whereas the coefficient in (38) is n^{-1} when $R = 0$ (no uniform convexity) or $S = +\infty$ (no upper bound on $\text{Hess}(V)$): hence at least in these cases our bound is stronger.

The bound (38) has been obtained in [29] by a L^2 argument. We shall see in the appendix that it can be recovered by linearisation in the Monge-Ampère equation.

Proof of Proposition 4.1. Start from Lemma 2.5 with $f = 1$ and $\nu = g\mu$, or equivalently from

$$\begin{aligned} & \int_0^1 \int (\nabla\varphi(x) - x) \cdot \text{Hess}(V)(x + t(\nabla\varphi(x) - x))(\nabla\varphi(x) - x)(1-t) dt d\mu(x). \\ & \leq \lambda H(\nu|\mu) + n(1 - \lambda + \lambda \log \lambda) + (1 - \lambda) \left(\int V d\nu - \int V d\mu \right) \quad (39) \end{aligned}$$

for all $\lambda > 0$, by linearizing the exponential term.

Let a be a constant and h a smooth function to be chosen later, such that $\int h d\mu = 0$. For $\varepsilon > 0$ small, let us take $\nu = (1 + \varepsilon h)\mu$, and $\lambda = 1 + a\varepsilon$. Then the right-hand side in (39) is

$$\frac{\varepsilon^2}{2} \left[\int h^2 d\mu + na^2 - 2a \int Vh d\mu \right] + o(\varepsilon^2).$$

Moreover the transport map $\nabla\varphi$ is given by $\nabla\varphi(x) = x + \varepsilon \nabla\omega(x) + o(\varepsilon)$ where $L\omega = -g$ using an expansion of the Monge-Ampère equation; hence the left-hand side in (39) is

$$\frac{1}{2} \int (\nabla\varphi(x) - x) \cdot \text{Hess}(V)(x)(\nabla\varphi(x) - x) d\mu(x) + o(\varepsilon^2) \geq \int Q_1^V \phi d\nu - \int \phi d\mu + o(\varepsilon^2),$$

where for any $t > 0$ and y

$$Q_t^V \phi(y) = \inf_x \left\{ \phi(x) + \frac{1}{2t} (y - x) \cdot \text{Hess}(V)(x) (y - x) \right\}.$$

To sum up,

$$\int Q_1^V \phi d\nu - \int \phi d\mu \leq \frac{\varepsilon^2}{2} \left[\int h^2 d\mu + na^2 - 2a \int Vh d\mu \right] + o(\varepsilon^2)$$

for all ϕ, a and g .

Let now f be a fixed smooth function such that $\int f d\mu = 0$, and let $\phi = \varepsilon f$. Observing that

$$Q_1(\varepsilon f) = \varepsilon f - \frac{\varepsilon^2}{2} \nabla f \cdot \text{Hess}(V)^{-1} \nabla f + o(\varepsilon^2),$$

we obtain

$$2 \int fh d\mu - \int \nabla f \cdot \text{Hess}(V)^{-1} \nabla f d\mu \leq \int h^2 d\mu + na^2 - 2a \int Vh d\mu.$$

We minimise over h under the condition $\int h d\mu = 0$, by choosing $h = f + a(V - \int V d\mu)$. Hence

$$\int f^2 d\mu \leq \int \nabla f \cdot \text{Hess}(V)^{-1} \nabla f d\mu + I a^2 - 2a \int V f d\mu$$

for all a , where $I = n - \text{Var}_\mu(V)$. Necessarily I is positive. 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 $a = I^{-1} \int V f d\mu$. This concludes the proof of Proposition 4.1. \triangleright

Remark 4.2 The fact that $I > 0$ (needed in the proof above) as soon as $\text{Hess}(V) > 0$, can be simply recovered in the following way: let $f = V - \int V d\mu$, $L = \Delta - \nabla V \cdot \nabla$ and g solve $Lg = f$; then

$$\begin{aligned} 0 &= \int (f - Lg)^2 d\mu = \int [f^2 - 2fLg + (Lg)^2] d\mu = \int [f^2 + 2\nabla g \cdot \nabla f + \sum_{i,j} (\partial_{ij}g)^2 + \nabla g \cdot \text{Hess}(V) \nabla g] d\mu \\ &> \int (f^2 + 2\Delta g + \sum_{i,j} (\partial_{ij}g)^2) d\mu = \int (f^2 - n + \sum_i (g_i + 1)^2) d\mu \geq \int f^2 d\mu - n \end{aligned}$$

unless $\nabla g = 0$, that is $V - \int V d\mu = 0$, which is impossible. Here the g_i are the n eigenvalues of the symmetric matrix $\text{Hess}(g)$. Hence

$$I = - \int f^2 d\mu + n > 0.$$

Here we have used the integration by parts relations (see [3])

$$\int \nabla f \cdot \nabla g d\mu = - \int Lf g d\mu, \quad \int (Lg)^2 d\mu = \int \left[\sum_{i,j} (\partial_{ij} g)^2 + \nabla g \cdot \text{Hess}(V) \nabla g \right] d\mu. \quad (40)$$

The fact that $I > 0$ is also a consequence of the bound (15) in [29] : indeed

$$\int f^2 d\mu \geq \frac{1}{n} \int \left(\Delta g - \int \Delta g d\mu \right)^2 d\mu + \frac{1}{n} \left(\int f^2 d\mu \right)^2 + \int \nabla g \cdot \text{Hess}(V) \nabla g d\mu > \frac{1}{n} \left(\int f^2 d\mu \right)^2$$

again unless $\nabla g = 0$, which is impossible. This means again that $I > 0$. Let us finally mention that V. H. Nguyen [33] has observed that $I \geq 0$ as soon as V is convex.

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

The following result gives an improved version of the Brascamp-Lieb inequality (3) 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 C^2 function satisfying $\text{Hess}(V) > 0$. Then for any smooth function f such that $\int f d\mu = 0$,*

$$\text{Var}_\mu(f) \leq \int \nabla f \cdot \text{Hess}(V)^{-1} \nabla f d\mu - \int \frac{(f - \nabla f \cdot \text{Hess}(V)^{-1} \nabla V)^2}{n + \nabla V \cdot \text{Hess}(V)^{-1} \nabla V} d\mu. \quad (41)$$

For the standard Gaussian measure γ we obtain a new dimensional Poincaré inequality

$$\text{Var}_\gamma(f) \leq \int |\nabla f|^2 d\gamma - \int \frac{(f - \nabla f \cdot x)^2}{n + |x|^2} d\gamma \quad (42)$$

for any smooth function f such that $\int f d\gamma = 0$. By the Cauchy-Schwarz inequality and integration by part,

$$\int \frac{(f - \nabla f \cdot x)^2}{n + |x|^2} d\gamma \geq \frac{(\int \nabla f \cdot x d\gamma)^2}{2n} = \frac{(\int \Delta f d\gamma)^2}{2n} = \frac{(\int f |x|^2 / 2 d\gamma)^2}{n - \text{Var}_\gamma(|x|^2 / 2)}.$$

Therefore, for the Gaussian measure, inequality (42) is stronger than (6) mentioned in the introduction (and naturally equality still holds for $f = |x|^2 / 2$).

Proof

\triangleleft We adapt the argument of [8]. Let f be a smooth compactly supported function satisfying $\int f d\mu = 0$. We apply (8) 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

$$\phi_\delta(z) = -n \log \inf_{h \in \mathbb{R}^n} \left\{ Z_\delta^{1/n} \exp \left(-\frac{2\delta}{n} f(z+h) + \frac{V(z+h)}{n} \right) + \exp \left(\frac{V(z-h)}{n} \right) \right\} + n \log(2) + V(z). \quad (43)$$

Then the Borell-Brascamp-Lieb inequality insures 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.

We first have to estimate h_δ at which the infimum in (43) is attained. For any $\delta > 0$, h_δ satisfies

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

From this we are deducing a Taylor expansion of h_δ . First, since $\int f d\mu = 0$,

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

Moreover f is smooth with compact support and V is strictly convex, so $h_\delta \rightarrow 0$ as $\delta \rightarrow 0$. A first-order Taylor expansion of the above equality implies

$$-\delta \nabla f(z) + \text{Hess}(V)(z)h_\delta - \frac{\delta}{n} f(z) \nabla V(z) + \frac{h_\delta \cdot \nabla V(z)}{n} \nabla V(z) + o(\delta) = 0, \quad (44)$$

where $o(\delta)$ can be chosen uniformly in z . By taking the scalar product of (44) with $\text{Hess}(V)^{-1} \nabla V$ one gets

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

for the point z , and then

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

again by (44); here $X = \nabla f \cdot \text{Hess}(V)^{-1} \nabla V$ and $Y = \nabla V \cdot \text{Hess}(V)^{-1} \nabla V$.

Hence, again for the point z ,

$$\phi_\delta = -n \log \left[1 - \frac{\delta}{n} f - \frac{\delta h_\delta \cdot \nabla f}{n} + \frac{h_\delta \cdot \text{Hess}(V) h_\delta}{2n} + \frac{\delta^2}{n^2} f^2 + \frac{(h_\delta \cdot \nabla V)^2}{2n^2} - \frac{\delta f}{n^2} h_\delta \cdot \nabla V + \frac{\delta^2}{n} \int f^2 d\mu \right] + o(\delta^2).$$

From the identities $\log(1+x) = x - x^2/2 + o(x^2)$ and (44), we get

$$\phi_\delta = \delta f + \frac{h_\delta \cdot \text{Hess}(V) h_\delta}{2} - \frac{\delta^2}{2n} f^2 + \frac{(h_\delta \cdot \nabla V)^2}{2n} - \frac{\delta^2}{4} \int f^2 d\mu + o(\delta^2).$$

The above expressions of h_δ and $h_\delta \cdot \nabla V$ finally give

$$\phi_\delta = \delta f + \frac{\delta^2}{2} \nabla f \cdot \text{Hess}(V)^{-1} \nabla f - \frac{\delta^2}{2} \frac{(f - X)^2}{n + Y} - \frac{\delta^2}{4} \int f^2 d\mu + o(\delta^2).$$

In conclusion, the second-order Taylor expansion of the Borell-Brascamp-Lieb inequality $\int e^{\phi_\delta} d\mu \geq 1$ implies

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

for all smooth f compactly supported such that $\int f d\mu = 0$. This concludes the argument by definition of X and Y . The inequality can be extended to smooth functions such that all terms are well defined. \triangleright

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 (37) with G. Hargé's bound, as the same covariance term appears in our optimal transport method. Let us now compare (41) 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 (37) and (41). In the Gaussian case we have already observed that (41) is stronger than (37). 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 (37) vanishes for functions f for which the one in (41) does not.

In fact, the additional term in (41) vanishes if and only if there exist $a \in \mathbb{R}$ and $b \in \mathbb{R}^n$ such that $f = a + \nabla V \cdot b$ (and then $a = \int f e^{-V}$, $b = \int (x - \int x e^{-V}) f e^{-V}$). But it is classical that these

functions f are exactly those for which equality holds in the Brascamp-Lieb inequality (3). Hence the additional term in (41) can be seen as a (weighted) way of measuring the distance of a function to the optimisers in the Brascamp-Lieb inequality (3).

Very recently, and under the same hypothesis as in Theorem 4.3, D. Cordero-Erausquin in [15, Prop. 5] proved that

$$\text{Var}_\mu(f) \leq \int \nabla f \cdot \text{Hess}(V)^{-1} \nabla f d\mu - C_\mu \int (f - \nabla V \cdot b)^2 d\mu \quad (45)$$

for all f satisfying $\int f d\mu = 0$; here $b = \int y f d\mu$ and C_μ depends on the Poincaré constant of the measure μ and numerical constants. The additional term appears here as a non-weighted distance to the optimisers. A more quantitative comparison between (41) and (45) can not easily be performed as C_μ depends on numerical constants.

- We now turn to the Gaussian case when $\mu = \gamma$. We have already observed that (41) is stronger than (37), which is exactly (6). On the other hand, (7) is a purely spectral inequality. We have numerically checked that (41) implies (7) for the Hermite polynomial functions H_k , $k \in \{1, \dots, 7\}$. 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

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

has been proved in [9, Sect. 2]. There 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.

Appendix

G. Hargé's bound (38) can be recovered by linearisation in the Monge-Ampère equation (22). 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 (22) with such μ_1 and μ_2 and observing that

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

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

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

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

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

so that

$$\int f^2 d\mu = \int \nabla f \cdot \text{Hess}(V)^{-1} \nabla f d\mu - \int \left(|M \nabla\theta_1 - M^{-1} \nabla f|^2 + \text{tr}[(\text{Hess}(\theta_1))^2] \right) d\mu \quad (46)$$

by integration. At this point one recognizes terms in the proof of [29, 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 $\text{Hess}(V) \leq S$. In particular

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

by (40), if $\text{Hess}(V) \geq R \text{Id}_n$. Hence

$$\int \left(|M \nabla \theta_1 - M^{-1} \nabla h|^2 + \text{tr}[(\text{Hess}(\theta_1))^2] \right) d\mu \geq \left(1 + \frac{R}{S} \right) \sum_{i,j} \int (\partial_{ji}^2 \theta_1)^2 d\mu \geq \frac{1}{n} \left(1 + \frac{R}{S} \right) \left(\int \Delta \theta_1 d\mu \right)^2 \quad (47)$$

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 (46) and (47) we finally recover (38) since by integration by parts and (40)

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

Acknowledgements. 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] D. Bakry, F. Bolley, and I. Gentil. Dimension dependent hypercontractivity for Gaussian kernels. Prob. Theor. Rel. Fields, 154(3): 845–874, 2012.
- [3] D. Bakry, I. Gentil and M. Ledoux. Analysis and geometry of Markov diffusion operators, volume 348 of Grund. Math. Wiss. Springer, Berlin, 2014.
- [4] D. Bakry and M. Ledoux. A logarithmic Sobolev form of the Li-Yau parabolic inequality. Rev. Mat. Iberoam., 22(2):683–702, 2006.
- [5] F. Barthe and A. Kolesnikov. Mass transport and variants of the logarithmic Sobolev inequality. J. Geom. Anal., 18(4):921–979, 2008.
- [6] S. G. Bobkov, I. Gentil and M. Ledoux. Hypercontractivity of Hamilton-Jacobi equations. J. Math. Pures Appl. (9), 80(7):669–696, 2001.
- [7] 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.
- [8] S. Bobkov and M. Ledoux. From Brunn-Minkowski to Brascamp-Lieb and to logarithmic Sobolev inequalities Geom. Funct. Anal., 10:1028–1052, 2000.
- [9] S. Bobkov and M. Ledoux. Weighted Poincaré-type inequalities for Cauchy and other convex measures. Ann. Probab., 37(2):403–427, 2009.
- [10] 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.
- [11] F. Bolley, I. Gentil and A. Guillin. Dimensional contraction via Markov transportation distance. J. London Math. Soc., 90(1):309–332, 2014.
- [12] F. Bolley, I. Gentil, A. Guillin and K. Kuwada. Equivalence between dimensional contractions in Wasserstein distance and curvature-dimension condition. In preparation.
- [13] E. Carlen. Superadditivity of Fisher’s information and logarithmic Sobolev inequalities. J. Funct. Anal., 101:194–211, 1991.
- [14] D. Cordero-Erausquin. Some applications of mass transport to Gaussian type inequalities. Arch. Rat. Mech. Anal., 161:257–269, 2002.
- [15] D. Cordero-Erausquin. Transport inequalities for log-concave measures, quantitative forms and applications. Preprint, 2015.
- [16] 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.
- [17] M. Del Pino and J. Dolbeault. The optimal Euclidean L^p -Sobolev logarithmic inequality. J. Funct. Anal., 197(1):151–161, 2003.

- [18] R. Eldan. A two-sided estimate for the Gaussian noise stability. To appear in Invent. Math.
- [19] 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.
- [20] M. Fathi, E. Indrei and M. Ledoux. Quantitative logarithmic Sobolev inequalities and stability estimates. Preprint, 2014.
- [21] A. Figalli and E. Indrei. A sharp stability result for the relative isoperimetric inequality inside convex cones. J. Geom. Anal., 23:938–969, 2013.
- [22] A. Figalli and D. Jerison. Quantitative stability for the Brunn-Minkowski inequality. Preprint, 2014.
- [23] A. Figalli, F. Maggi and A. Pratelli. A mass transportation approach to quantitative isoperimetric inequalities. Invent. Math., 182:167–211, 2010.
- [24] 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.
- [25] I. Gentil. The general optimal L^p -Euclidean logarithmic Sobolev inequality by Hamilton-Jacobi equations. J. Funct. Anal. 202(2):591–599, 2003.
- [26] I. Gentil. From the Prékopa-Leindler inequality to modified logarithmic Sobolev inequality. Ann. Fac. Sci. Toulouse, 6, 17(2):291–308, 2008.
- [27] I. Gentil. Dimensional contraction in Wasserstein distance for diffusion semigroups on a Riemannian manifold. Potential Anal., 42(4):861–873, 2015.
- [28] L. Goldstein, I. Nourdin and G. Peccati. Gaussian phase transitions and conic intrinsic volumes: Steining the Steiner formula. Preprint, 2014.
- [29] G. Hargé. Reinforcement of an inequality due to Brascamp and Lieb. J. Funct. Anal., 254(2):267–300, 2008.
- [30] 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.
- [31] S. Lisini. Nonlinear diffusion equations with variable coefficients as gradient flows in Wasserstein spaces. ESAIM Contr. Opt. Calc. Var., 15:712–740, 2009.
- [32] R. McCann. A convexity principle for interacting gases. Adv. Math., 128(1):153–179, 1997.
- [33] 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.
- [34] 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.
- [35] M. Talagrand. Transportation cost for Gaussian and other product measures. Geom. Funct. Anal., 6(3):587–600, 1996.
- [36] C. Villani. Optimal transport, Old and new, volume 338 of Grund. Math. Wiss. Springer, Berlin, 2009.