Prékopa–Leindler type inequalities on Riemannian manifolds, Jacobi fields, and optimal transport

Dario Cordero-Erausquin[*],   Robert J. McCann[†], and Michael Schmuckenschläger[‡]

June 10, 2005

Abstract

We investigate Prékopa-Leindler type inequalities on a Riemannian manifold M equipped with a measure with density e^{-V} where the potential V and the Ricci curvature satisfy $\operatorname{Hess}_x V + \operatorname{Ric}_x \geq \lambda\, I$ for all $x \in M$, with some $\lambda \in \mathbb{R}$. As in our earlier work [14], the argument uses optimal mass transport on M, but here, with a special emphasis on its connection with Jacobi fields. A key role will be played by the differential equation satisfied by the determinant of a matrix of Jacobi fields. We also present applications of the method to logarithmic Sobolev inequalities (the Bakry-Emery criterion will be recovered) and to transport inequalities. A study of the displacement convexity of the entropy functional completes the exposition.

Résumé

Nous étudions l'extension d'inégalités de type Prékopa-Leindler au cas d'une variété riemannienne M équipée d'une mesure ayant une densité e^{-V} où le potentiel V et la courbure de Ricci vérifient $\operatorname{Hess}_x V + \operatorname{Ric}_x \geq \lambda\, I$ ($\forall x \in M$), pour un certain $\lambda \in \mathbb{R}$. Nous ferons appel, comme dans notre travail précédent [14], au transport optimal de mesure. Mais nous exploiterons plus encore son lien avec les champs de Jacobi, ce qui permettra de ramener la discussion à l'étude du déterminant d'une matrice de champs de Jacobi. Nous présentons également d'autres applications de la méthode, en particulier aux inégalités de Sobolev logarithmiques (critère de Bakry-Emery) et à l'étude de la convexité de déplacement de la fonctionnelle entropie.

[*]Laboratoire d'Analyse et de Mathématiques Appliquées, Université de Marne la Vallée, 77454 Marne la Vallée Cedex 2, France. `cordero@math.univ-mlv.fr`

[†]Department of Mathematics, University of Toronto, Toronto Ontario Canada M5S 3G3. `mccann@math.toronto.edu`

[‡]Institut für Analysis und Numerik, Universität Linz, A-4040 Linz, Österreich. `schmucki@bayou.uni-linz.ac.at`

1 Introduction

The Brunn-Minkowski inequality plays a central role in geometric convex analysis. It can be stated as follows: for $A, B \subset \mathbb{R}^n$,

$$\text{vol}(A + B)^{1/n} \geq \text{vol}(A)^{1/n} + \text{vol}(B)^{1/n},$$

where vol denotes the Lebesgue measure. By applying this inequality when B is an Euclidean ball of radius ε and by letting $\varepsilon \to 0$, one easily recovers the isoperimetric inequality in Euclidean space. Using homogeneity, the Brunn-Minkowski inequality can be restated in the following equivalent form: for $A, B \subset \mathbb{R}^n$ and $s \in [0, 1]$,

$$\text{vol}((1 - s)A + sB) \geq \text{vol}(A)^{1-s} \text{vol}(B)^s. \tag{1}$$

From the beginning of the nineteen-seventies there has been a significant and systematic effort to obtain functional versions of geometric inequalities (see for instance [8, 9, 36] and the surveys [16, 18] for different points of view...). These functional versions are not only more powerful (they can be applied in different settings) but they also shed new light on the geometric inequalities themselves. An inequality which has proved to be extremely useful is the so-called Prékopa-Leindler inequality [36, 24, 37].

Theorem 1 (Prékopa-Leindler inequality) *Let $s \in (0, 1)$ and $u, v, w : \mathbb{R}^n \longrightarrow \mathbb{R}_+$ be such that, for all $x, y \in \mathbb{R}^n$,*

$$w((1 - s)x + sy) \geq u^{1-s}(x)v^s(y). \tag{2}$$

Then, $\|w\|_{L^1} \geq \|u\|_{L^1}^{1-s}\|v\|_{L^1}^s$.

An equivalent and shorter way of stating this result is to say that for every $u, v : \mathbb{R}^n \longrightarrow \mathbb{R}_+$ one has

$$\int_{\mathbb{R}^n} \sup_{\{(x,y)\in\mathbb{R}^{2n}|z=(1-s)x+sy\}} \left\{u^{1-s}(x)v^s(y)\right\} \, dz \; \geq \; \left(\int_{\mathbb{R}^n} u\right)^{1-s}\left(\int_{\mathbb{R}^n} v\right)^s. \tag{3}$$

If we apply this inequality for $u = 1_A$ and $v = 1_B$, the indicator functions of sets $A, B \subset \mathbb{R}^n$, then we recover the Brunn-Minkowski inequality (1). For more information on Brunn-Minkowski and Prékopa-Leindler inequalities and their interplay with geometry and probability, we refer to [18, 28, 39, 23, 4].

We now want to take as reference measure a measure μ with density e^{-V}. The function $V : \mathbb{R}^n \to \mathbb{R}$ is called the potential and we are going to assume that

$$\text{Hess}_x V \geq \lambda \, \text{Id} \qquad \forall x \in \mathbb{R}^n, \tag{4}$$

for some $\lambda \in \mathbb{R}$, where Id denotes the identity matrix. It is clear after the work of Bakry and Emery, that this amounts to a "curvature" condition on the "space" $(\mathbb{R}^n, \mu)$. A typical case is the Gauss space obtained when μ is the standard Gaussian measure γ_n given by

$V(x) = |x|^2/2 + n\log(2\pi)/2$ and for which (4) holds with $\lambda = 1$. If $\lambda \geq 0$, the density e^{-V} is log-concave and the we immediately have that the Prékopa-Leindler inequality (3) extends to integration with respect to μ. However, if $\lambda > 0$, one can expect a stronger inequality to hold. This is indeed the case, as suggested by the remarkable paper by Maurey [27]. For $f, g, h : \mathbb{R}^n \to \mathbb{R}_+$, we apply the Prékopa-Leindler inequality to

$$u(x) = f(x)e^{-V(x)}, \ v(y) = g(y)e^{-V(x)}, w(z) = h(z)e^{-V(z)}.$$

We want to find under which conditions on f, g and h the assumption (2) is satisfied. The terms in V can be simplified. Indeed, we notice that using an integral Taylor expansion (on $[0, s]$ and on $[s, 1]$) we have for every smooth function $\alpha : [0, 1] :\longrightarrow \mathbb{R}$, $k \in \mathbb{R}$, $s \in [0, 1]$, that

$$\left(\forall t \in [0, 1], \ \alpha''(t) \geq k \right) \Longrightarrow (1 - s)\alpha(0) + s\alpha(1) - \alpha(s) \geq k\, s(1 - s)/2. \qquad (5)$$

Thus, by applying (5) for $\alpha(t) := V\big((1 - t)x + ty\big)$ we see that condition (4) implies that $(1-s)V(x)+sV(y)-V((1-s)x+sy) \geq \lambda\, s(1-s)|x-y|^2/2$ for every $x, y \in \mathbb{R}^n$. We shall use this simple and well known observation several times in this paper. As a consequence we obtain the following reformulation the Prékopa-Leindler inequality.

Theorem 2 (Weighted Prékopa-Leindler inequality) *Let μ be a measure of the form $d\mu = e^{-V} dx$ where V verifies (4). Let $s \in [0, 1]$ and $f, g, h : \mathbb{R}^n \longrightarrow \mathbb{R}_+$ be such that, $\forall x, y \in \mathbb{R}^n$*

$$h((1 - s)x + sy) \geq e^{-\lambda\, s(1-s)\, |x-y|^2/2}\, f^{1-s}(x)g^s(y). \qquad (6)$$

Then, $\displaystyle\int_{\mathbb{R}^n} h\, d\mu \geq \left(\int_{\mathbb{R}^n} f\, d\mu \right)^{1-s} \left(\int_{\mathbb{R}^n} g\, d\mu \right)^s.$

This result (one should rather say "observation") is of course not new. The fact that the Prékopa-Leindler inequality contains this relative form was used, after Maurey's work [27], by several authors; note that the Prékopa-Leindler inequality corresponds to the special case $V = 0$ and $\lambda = 0$. One can derive from this form concentration and log-Sobolev inequalities for the measure μ, as in [27, 38, 6].

The Prékopa-Leindler inequality was extended by the authors [14] to Riemannian manifolds. *In this paper, M will denote a smooth, complete n-dimensional Riemannian manifold with geodesic distance d.* We will denote by $d\mathrm{vol}$ the Riemannian element of volume and $(T_xM, \cdot, |\cdot|)$ will be the Euclidean structure on the tangent space T_xM at $x \in M$. For $x, y \in M$ and $t \in [0, 1]$, let $Z_t(x, y)$ be the barycenter between x and y given by

$$Z_t(x, y) = \{z \in M \ ; \ d(x, z) = td(x, y) \ \text{ and } \ d(z, y) = (1 - t)d(x, y)\}.$$

Except maybe when x belongs to the cut locus $cut(y)$ of y, the set $Z_t(x, y)$ reduces to a single point. In fact, when $x \notin cut(y)$, the curve $t \longrightarrow Z_t(x, y)$ describes exactly the minimal geodesic joining x to y. The Riemannian Prékopa-Leindler inequality obtained in [14] takes the following form: for $u, v : M \longrightarrow \mathbb{R}_+$ and $s \in [0, 1]$ one has

$$\int_M \sup_{z \in Z_s(x,y)} \left\{ D_s(x, y)\, u^{1-s}(x)v^s(y) \right\} d\mathrm{vol}(z) \ \geq \ \left(\int_M u \right)^{1-s} \left(\int_M v \right)^s,$$

3

where $D_s(x, y) \geq 0$ is a distortion factor due to curvature depending on x, y and s only. We have $D_s(x, y) \leq 1$ if the curvature is non-negative, while the opposite inequality holds in spaces of non-positive curvature. Of course, $D_s(x, y) \equiv 1$ in Euclidean space (the Prékopa-Leindler inequality (3) is then recovered). When the Ricci curvature is bounded from below, $\text{Ric}_x \geq k(n-1)\,\text{Id}$ for all $x \in M$ and some $k \in \mathbb{R}$, then it is possible to bound from above the distortion factor $D_s(x, y)$ by a factor depending only on the distance $d(x, y)$ between x and y. The Ricci curvature Ric_x at point $x \in M$ will be seen either as a self-adjoint operator or as a bilinear form on $T_x M$. Introduce for $k \in \mathbb{R}$,

$$S_k(d) := \frac{\sin(\sqrt{k}\, d)}{\sqrt{k}\, d} = \begin{cases} (\sin d)/d & \text{for } k = 1 & \text{(spherical case)} \\ 1 & \text{for } k = 0 & (\mathbb{R}^n) \\ (\sinh d)/d & \text{for } k = -1 & \text{(hyperbolic case)} \end{cases}. \tag{7}$$

If M satisfies $\text{Ric} \geq k(n-1)$ then, as proved in [14], one has for the volume distortion along any geodesic of length $d := d(x, y)$:

$$D_s(x, y) \leq \left(\frac{S_k(d)}{S_k^{1-s}((1-s)d)\, S_k^s(sd)} \right)^{n-1}, \tag{8}$$

with equality if M is of constant sectional curvature equal to k. It is a computational exercise to check that for $s \in [0, 1]$, $k \in \mathbb{R}$ and $d \geq 0$ ($d < \pi/k^{1/2}$ when $k > 0$),

$$\left(\frac{S_k(d)}{S_k^{1-s}((1-s)d)\, S_k^s(sd)} \right) \leq e^{-s(1-s)\,k\,d^2/2} \tag{9}$$

(setting $t = \sqrt{|k|}d$ reduces to the cases $k = 1$ and $k = -1$). After taking the $(n-1)$ power in (9), we obtain an upper bound for $D_s(x, y)$ which depends on $\lambda := k(n-1)$. As a consequence, we obtain the following result which was therefore implicitly already contained in [14].

Theorem 3 (Riemannian inequality) *Assume that, for some $\lambda \in \mathbb{R}$, the Ricci curvature on M verifies:*

$$\text{Ric}_x \geq \lambda\,\text{Id} \qquad \forall x \in M. \tag{10}$$

If $u, v, w : M \to \mathbb{R}_+$ and $s \in [0, 1]$ are such that, for every $x, y \in M$ and $z \in Z_s(x, y)$,

$$w(z) \geq e^{-\lambda\,s(1-s)\,d^2(x,y)/2}\, u^{1-s}(x)v^s(y), \tag{11}$$

then, $\int_M w \geq \left(\int_M u \right)^{1-s} \left(\int_M v \right)^s.$

By comparing (4) and (6) with (10) and (11), we see once again the analogy between the curvature of the potential V and the curvature of the manifold. However, we would like to push forward our investigation and consider, on our Riemannian manifold M, a measure $d\mu = e^{-V}\,d\text{vol}$ with a potential $V : M \longrightarrow \mathbb{R}$. What weighted form of the Prékopa-Leindler

inequality can we expect? Of course, following the Euclidean method, we may combine (4) with (10) to obtain an inequality, but that is not what we have in mind. We seek instead an inequality which allows the curvature of the manifold to be compensated by the curvature of the potential, and vice versa. Following the work of Bakry and Emery, we are willing to assume

$$\mathrm{Hess}_x\, V + \mathrm{Ric}_x \geq \lambda\,\mathrm{Id} \qquad \forall x \in M \tag{12}$$

for some $\lambda \in \mathbb{R}$. It is not clear whether such an assumption can be handled with our Riemannian versions [14] of the Prékopa-Leindler inequality. But in any case, one of the aims of this paper is to present an alternative approach to such inequalities. We will of course use, as before, optimal mass transport on M. But we will make a more direct use of its connection with Jacobi fields along geodesics, and of the related ODE's. The main result is, as expected, a weighted version of the Prékopa-Leindler inequality on manifolds.

Theorem 4 (Weighted version) *Let μ be a measure on M of the form $d\mu = e^{-V}\,d\mathrm{vol}$ where V and the Ricci curvature verify (12) for some $\lambda \in \mathbb{R}$. Let $s \in [0,1]$ and $f, g, h : M \longrightarrow \mathbb{R}_+$ be such that, $\forall x, y \in M$ and $z \in Z_s(x,y)$,*

$$h(z) \geq e^{-\lambda\, s(1-s)\, d^2(x,y)/2}\, f^{1-s}(x)\, g^s(y). \tag{13}$$

Then, $\displaystyle\int_M h\,d\mu \geq \left(\int_M f\,d\mu\right)^{1-s} \left(\int_M g\,d\mu\right)^s.$

Concentration of measure inequalities can be recovered following the ideas of Maurey's (τ)-property. Let us assume that μ is a *probability* measure on M of the form $d\mu = e^{-V}\,d\mathrm{vol}$ verifying (12). Then, by applying the previous result for $s = 1/2$ and $h \equiv 1$, we obtain that for every $\varphi : M \longrightarrow \mathbb{R}$,

$$\int_M e^{-\varphi}\,d\mu \int_M e^{Q_\lambda(-\varphi)}\,d\mu \leq 1, \qquad \text{where}\quad Q_\lambda\varphi(y) := \inf_{x \in M} \left(\lambda\, d^2(x,y)/4 - \varphi(x)\right). \tag{14}$$

Given a set $A \subset M$, we can apply this inequality with $\varphi(y) = 0$ on A and $+\infty$ outside A. Setting $d(\cdot, A) := \inf_{x \in A} d(\cdot, x)$, one gets

$$\int_M e^{\lambda\, d^2(\cdot, A)/4}\,d\mu \leq \mu(A)^{-1},$$

and therefore $\mu\big(\{x \in M\,;\ d(x, A) \geq \varepsilon\}\big) \leq \mu(A)^{-1}\, e^{-\lambda\, \varepsilon^2/4}$. This type of concentration of measure inequality was first proved on compact Riemannian manifolds with positive Ricci curvature ($V = 0$ and $d\mu = d\mathrm{vol}/\mathrm{vol}(M)$) by Gromov and Milman [19]. The reader can consult Ledoux [23] for references and background.

Theorem 4 is the main new result of the paper. It clearly extends the weighted Prékopa-Leindler inequality from Euclidean space stated in Theorem 2. Our other results consist mainly of proofs which give new insights into known results — as in the case of the Bakry-Emery criterion (Theorem 7) — or else lend rigour to heuristic arguments — as in the case

of Otto and Villani's [35] displacement semiconvexity of Riemannian entropy (Theorem 10).

The organization of the paper is as follows. The next section §2 recalls some relevant facts about mass transport and about its link to Jacobi fields. Then, in section §3 we recall a useful Lemma on the determinant of a matrix of Jacobi fields and use it to prove Theorem 4. Section §4 is devoted to logarithmic Sobolev and transport inequalities for a measure μ with density e^{-V} under the assumption (12). We shall give a mass transportation proof of a criterion of Bakry and Emery. A unified point of view on log-Sobolev and transport inequalities is given in section §5 with study of the displacement convexity of the entropy functional

$$E(u) := \int_M u \log u + \int_M uV \tag{15}$$

2   Mass transport and Jacobi fields

We will recall several results taken from our earlier works [32] and [14] concerning optimal transport on Riemannian manifolds. All the results needed for the proof of Theorem 4 are collected in Proposition 5. The reader is free to take this proposition for granted and to jump to the next section. However, without the explanations we collect below, it appears as a rather mysterious black box.

Let us start from the classical Euclidean theory of mass transport. The interplay between mass transport and geometric functional analysis (and Brunn-Minkowski theory) has a long history and we would like to mention a few steps relevant for the present work. In 1957, Knothe [20] gave a proof of the Brunn-Minkowski inequality using mass transport. The map constructed by Knothe, sometimes referred to as the Knothe map, was later used by Gromov [33] to give a direct proof of the Euclidean isoperimetric inequality in its functional form. In the mid nineteen-eighties, a completely different mathematical landscape led Brenier to discover a new mass transportation map by solving a variational problem [10]. Connections between the resulting *optimal* map and the isoperimetric [41], Brunn-Minkowski [31], and Prékopa-Leindler [29] inequalities were discovered independently by Trudinger [41] and McCann [29, 31] (whose *displacement convexity* inequalities are based on this map). Shortly afterward, Barthe [5] used the same map to prove deep convolution inequalities. Since then, the Brenier map has found many interesting geometric applications. For instance, it has been used to derive Aleksandrov-Fenchel inequalities by Alesker, Dar, and Milman [1], and to derive sharp Sobolev and related inequalities on $\mathbb{R}^n$ [15] and on bounded domains [26] by Cordero-Nazaret-Villani [15] and Maggi-Villani [26]. One can consult Villani's book [43] for more details and background on optimal mass transportation theory.

The result of Brenier [10], as improved by McCann [30], is as follows. Let μ and ν be two probability measures on $\mathbb{R}^n$ and assume that μ is absolutely continuous with respect to Lebesgue measure. Then, there exists a convex function φ such that ν is the image measure of μ under the map $T = \nabla \varphi$ (we say that $\nabla \varphi$ *pushes-forward* or *transports* μ onto ν). Let us recall that given two (Borel) measures μ and ν on some measurable (topological)

space, one says that ν is the *image measure* (or *push-forward*) of μ under the (Borel) map T if,

$$\int b \circ T \, d\mu = \int b \, d\nu, \tag{16}$$

for every nonnegative Borel function b. Equivalently, $\nu(B) = \mu(T^{-1}(B))$ for every Borel set B. In this case we may write either $\nu = T(\mu)$ or $\nu = T_{\#}\mu$.

The map $T = \nabla\varphi$ between μ and ν is uniquely determined and is sometimes called the Brenier map or Brenier-McCann map. It is known to solve Monge's optimization problem: among all maps transporting μ onto ν, find T as close to the identity map as possible in the least-square sense, i.e. in $L^2((\mathbb{R}^n, \mu); \mathbb{R}^n)$. An important feature of this map is that its Jacobian matrix $\operatorname{Hess}\varphi \geq 0$ is symmetric and non-negative definite, which makes various matrix inequalities available.

It is *a priori* not obvious how the map $T = \nabla\varphi$ should generalize to Riemannian manifolds. In order to focus on the displacement, it is convenient to introduce $\theta(x) := \varphi(x) - |x|^2/2$ so that T takes the form $T(x) := x + \nabla\theta(x)$. Of course, the inequality $\operatorname{Hess}\theta \geq -I$ satisfied by θ (distributionally) is less familiar to interpret than convexity of φ: the directional derivatives $\partial^2_{nn}\theta(x)$ can take either sign, depending on whether $T(x)$ expands or contracts locally, but they are bounded below since the contraction is not so severe as to become orientation reversing.

Brenier's result was extended to Riemannian manifolds by McCann [32]. Let μ and ν be two probability measures on M and suppose μ is absolutely continuous with respect to $d$vol. For technical reasons, it is simplest to assume that the probability measures μ and ν are compactly supported. Then, there exists a function $\theta : M \longrightarrow \mathbb{R}$ such that $-\theta$ is $d^2/2$-*concave* and such that ν is the image measure of μ under the map $F(x) = \exp_x(\nabla\theta(x))$. This map is uniquely defined and will be called the optimal map pushing μ forward to ν, because it again minimizes the mean-square Riemannian distance to the identity map, the mean being computed with respect to μ. Of course, it remains to clarify what $d^2/2$-concavity of φ means. One precise definition is that $\varphi = Q_2^2(\varphi)$ with Q_2 from (14); see also [32, 14]. We only summarize the properties which are needed for our work. Again, the explanations below are given solely as a motivation for Proposition 5.

First, θ is locally Lipschitz and therefore the map F is well defined almost everywhere. In fact, it was proved in [14] that θ has almost everywhere a second order derivative (in a suitable sense) and that almost everywhere $F(x) \notin cut(x)$. To be more specific, θ is locally semi-convex in the sense that for every x_0, there exists $C \geq 0$ such that the function $x \longrightarrow \theta(x) + Cd^2(x, x_0)$ is locally geodesically convex around x_0. As a consequence, its distributional second derivative is a measure given by the sum of an absolutely continuous part and a nonnegative singular part. In [14] we used the notion of Hessian in the sense of Aleksandrov, which means we say $\operatorname{Hess}_x \theta$ exists at $x \in M$ if $\nabla\theta(x)$ exists and if there exists a symmetric matrix H_x such that, for every direction $u \in T_x M$,

$$\theta(\exp_x(tu)) = \theta(x) + t\nabla\theta(x) \cdot u + \frac{t^2}{2}H_x(u) \cdot u + o(t^2).$$

We then write $\operatorname{Hess}_x \theta := H_x$. For a locally semi-convex function, the Hessian in the

sense of Aleksandrov exists almost everywhere and coincides almost everywhere with the absolutely continuous part of the distributional second derivative.

The $d^2/2$-concavity of $-\theta$ implies

$$\left[\mathrm{Hess}_x\, d^2_{F(x_0)}/2\, +\, \mathrm{Hess}_x\, \theta\right]_{x=x_0} \geq 0 \tag{17}$$

for almost every x_0 in the support of μ, where $d_y(\cdot) := d(\cdot, y)$ for $y \in M$. Note that the function $d^2_{F(x_0)}$ is smooth around x when $x \notin cut(F(x_0))$. In the Euclidean case, $F(x) = \exp_x \nabla\theta(x)$ with (17) reduces to $T(x) = x + \nabla\theta(x)$ with $\mathrm{Hess}\,\theta \geq -I$, and we therefore see that the map F encodes part of the geodesic structure of M.

Let us assume that μ and ν have densities u_0 and u_1, respectively. It is natural to expect from (16) that $u_0(x) = u_1(F(x))|\det dF_x|$. This is true in some weak sense. Let us *define* dF almost everywhere by $dF_{x_0} := Y(H + \mathrm{Hess}_{x_0}\, \theta)$, where the matrices Y and H are defined, when $F(x_0) \notin cut(x_0)$, by

$$Y := d[\exp_{x_0}]_{\nabla\theta(x_0)} \quad \text{and} \quad H = [\mathrm{Hess}_x\, d^2_{F(x_0)}/2]_{x=x_0}.$$

Then $u_0(x) = u_1(F(x))\det dF_x$ holds $d\mu$-almost everywhere. We now interpolate between the identity and the optimal map: Given F and $t \in [0, 1]$, we introduce

$$F_t(x) := \exp_x(t\nabla\theta(x)).$$

Observe that if $F(x) \notin cut(x)$, then $t \longrightarrow F_t(x) = Z_t(x, F(x))$ is the minimal geodesic joining $x = F_0(x)$ to $F(x) = F_1(x)$. Note also that for every y (close to x) the curve $t \longrightarrow F_t(y)$ is a again a geodesic. Therefore, it is natural to expect that $t \longrightarrow d(F_t)_x$ defines a *matrix of Jacobi fields* (in an orthonormal moving frame) along the geodesic $t \longrightarrow F_t(x)$.

A Jacobi field is obtained by considering any perturbation of a geodesic by geodesics (see for instance [17] for a more detailed presentation). Let $\gamma : [0, 1] \to M$ be a geodesic. Consider a perturbation γ_s of γ by a geodesic $[0, 1] \ni t \longrightarrow \gamma_s(t)$ for each $s \in (-\varepsilon, \varepsilon)$, with $\gamma_0 = \gamma$. Introduce, for a fixed $t \in [0, 1]$,

$$J(t) := \frac{d}{ds}\Big|_{s=0}\gamma_s(t).$$

Since $\gamma_0 = \gamma$, we have $J(t) \in T_{\gamma(t)}M$. The vector field $t \longrightarrow J(t)$ along γ is called a *Jacobi field*. It satisfies the following linear differential equation

$$\nabla^2_{\dot\gamma(t)}J(t) + R_{\gamma(t)}(\dot\gamma(t), J(t))\dot\gamma(t) = 0.$$

where $\dot\gamma(t) := \dfrac{d\gamma}{dt}(t)$, $\nabla^2_{\dot\gamma(t)}$ denotes the second covariant derivative along γ, and $R_x : T_xM \times T_xM \times T_xM \longrightarrow T_xM$ is the Riemann curvature tensor at $x \in M$. Since Jacobi's equation is second order, a Jacobi field is uniquely determined by $J(0)$ and $\nabla_{\dot\gamma(t)}J(0)$. It is very convenient to introduce a moving frame and to work with the coordinates of Jacobi

8

fields. Let $e_1(0), \ldots, e_n(0)$ be an orthonormal basis of $T_{\gamma(0)}M$ and consider the parallel transport of this basis along γ. We obtain an orthonormal basis of $T_{\gamma(t)}M$ denoted by $e_1(t), \ldots, e_n(t)$. As usual, we will impose that $e_1(t) = \dot{\gamma}(t)/|\dot{\gamma}(t)|$ for all $t \in [0, 1]$. Let $Y(t) \in \mathbb{R}^n$ be the coordinates of a Jacobi field $J(t)$ in the basis $e_1(t), \ldots, e_n(t)$. Then $t \longrightarrow Y(t)$ satisfies a second order linear ODE:

$$Y''(t) + R(t)Y(t) = 0 \qquad (18)$$

for some symmetric matrix $R(t)$. The matrix $R(t)$ is the matrix of the operator

$$
\begin{aligned}
T_{\gamma(t)}M &\longrightarrow T_{\gamma(t)}M \\
v &\longrightarrow R(\dot{\gamma}(t), v)\dot{\gamma}(t)
\end{aligned}
\qquad (19)
$$

When $M = \mathbb{R}^n$, we have $R \equiv 0$. More generally, R is of the form

$$R = R^* = \begin{pmatrix} 0 & 0 \\ 0 & D \end{pmatrix},$$

and the trace of $R(t)$ gives the Ricci curvature in the direction $\dot{\gamma}(t)$:

$$\operatorname{tr} R(t) = \operatorname{Ric}_{\gamma(t)}\big(\dot{\gamma}(t), \dot{\gamma}(t)\big).$$

For example, $D = I_{n-1}$ and $\operatorname{tr} R = n-1$ in the case of the sphere $M = S^n$.

Let us come back to our optimal transport map $F_t(x) = \exp_x(t\nabla\theta(x))$. As before, it is possible to compute in some weak sense the differential of F_t:

$$d(F_t)_{x_0} := Y(t)(H(t) + t\operatorname{Hess}_{x_0}\theta)$$

where the matrices $Y(t)$ and $H(t)$ are defined, when $F(x_0) \notin cut(x_0)$ (and therefore $F_t(x_0) \notin cut(x_0)$), by

$$Y(t) := d[\exp_{x_0}]_{t\nabla\theta(x_0)} \quad \text{and} \quad H(t) = [\operatorname{Hess}_x d^2_{F_t(x_0)}/2]_{x=x_0}.$$

It is easy to see that $A(t) := d(F_t)_x$ defines a matrix of Jacobi fields along the geodesic $t \longrightarrow F_t(x)$ with the properties $A(0) = I$ and $A'(t) = \operatorname{Hess}_x\theta$. One can prove that the image of $u_0(x)\,dx$ under the map F_t defines a probability measure which has a density. If we denote by u_t this density, then we have again, $u_0(x)\,dx$-almost everywhere,

$$u_0(x) = u_t(F_t(x))\det d(F_t)_x. \qquad (20)$$

One can also check that for every Borel set $K \subset M$, $\int_{F_s(K)} u_s = \int_K u_0$. We can now summarize all these properties in the following proposition. In fact, for the proof of Theorem 4, we shall not need any other property of the optimal map F besides those stated below: the reader can as well take this as a "definition" of the map F....

9

Proposition 5 ([14]) *Let u_0 and u_1 be two compactly supported probability densities on the manifold M and let $F(x) := \exp_x(\nabla\theta(x))$ be the optimal map pushing $u_0(x)d\mathrm{vol}(x)$ forward to $u_1(y)d\mathrm{vol}(y)$. Let us fix $s \in [0,1]$ and set $F_t(x) := \exp_x(t\nabla\theta(x))$ for every $t \in (0,1)$. Then there exists a σ-compact set $K \subset M$ and a probability density u_s such that*

$$\int_{F_t(K)} u_t \, d\mathrm{vol} = 1 \quad for \quad t \in \{0, s, 1\},$$

and $F(x) \notin cut(x)$ and $\mathrm{Hess}_x\,\theta$ exist for every $x \in K$.

Furthermore, if we introduce, for a fixed $x \in K$, the minimal geodesic $\gamma_x(t) := F_t(x)$, (joining $x = \gamma(0)$ to $F(x) = \gamma(1)$) we have

$$u_0(x) = u_t(\gamma_x(t)) \det A_x(t) > 0 \quad for \quad t \in \{0, s, 1\}, \tag{21}$$

where $t \longrightarrow A_x(t)$ is the unique matrix of Jacobi fields along the geodesic γ_x verifying $A_x(0) = I$ and $A_x'(0) = \mathrm{Hess}_x\,\theta$.

The need to specify $s \in [0,1]$ in the previous Proposition is technical, not essential. We were not able to ensure the existence of a common K suitable for all $t \in [0,1]$. Of course, one can find a set of full μ-measure K suitable for any countable family of t's.

3   Proof of Theorem 4

It is clear from the preceding discussion that for geometrical applications of optimal transportation, it is essential to understand the behavior of the determinant of an invertible ($\det A_x(t) > 0$) matrix of Jacobi fields. We emphasize the next Lemma which is a well known comparison result in Riemannian geometry (see e.g. Chavel[12]). A (standard) proof is included for completeness.

Lemma 6 *Let $\gamma : [0,1] \longrightarrow M$ be a geodesic and $t \longrightarrow A(t)$ an invertible matrix of Jacobi fields along γ such that $A(0) = I$ and $A'(0)$ is symmetric. If $\varphi(t) := -\log\det A(t)$ and $r(t) = \mathrm{Ric}_{\gamma(t)}\big(\dot\gamma(t), \dot\gamma(t)\big)$, then for each $t \in [0,1]$,*

$$\varphi''(t) - \frac{1}{n}\varphi'(t)^2 - r(t) \geq 0.$$

Proof. Once the problem is expressed by using a moving frame along γ, it becomes rather elementary linear algebra. Recall that $r(t)$ is exactly the trace of $R(t)$ the matrix of the operator defined in (19). By definition of Jacobi fields (18), we have, for every $t \in [0,1]$,

$$A''(t) + R(t)A(t) = 0.$$

Besides the symmetry of $R(t)$ and the invertibility of $A(t)$, we shall use that $A(0) = I$ and that $A'(0)$ is symmetric. Introduce $B(t) := A'(t)A(t)^{-1}$. We have $\varphi'(t) = -\operatorname{tr} B$ and $\varphi''(t) = -\operatorname{tr} B'$. The important point is to check that B is symmetric. Note that

$$B^* - B = A^{*-1}(A'^*A - A^*A')A^{-1}.$$

10

The term $A'^*A - A^*A'$ is constant since, using $R = R^*$, we see that its derivative vanishes. Using that $A'^*A - A^*A'$ is 0 at $t = 0$ and therefore at every t, we conclude that $B = B^*$. Since $A'A^{-1} + A(A^{-1})' = 0$, we find $B' = A''A^{-1} - B^2 = -R - B^2$ and so, after taking traces,

$$\varphi'' = -\operatorname{tr} B' = \operatorname{tr}(B^2) + \operatorname{tr} R.$$

To conclude, use Hölder's inequality $\operatorname{tr}(B^2) = \operatorname{tr}(BB^*) \geq \frac{1}{n}(\operatorname{tr} B)^2$. Since $\operatorname{tr} B = -\varphi'$, we indeed find that $\varphi'' \geq \frac{1}{n}\varphi'^2 + \operatorname{tr} R$. $\qquad\square$

We can now prove the weighted version of the Riemannian inequality, Theorem 4. Let f, g, h be as in the Theorem. Without loss of generality we can assume that these functions are compactly supported and that $\int_M f \, d\mu = \int_M g \, d\mu = 1$. Our goal is to prove that $\int_M h \, d\mu \geq 1$.

The idea of the proof is to introduce the optimal map $F(x) = \exp_x(\nabla\theta(x))$ pushing $f \, d\mu$ forward to $g \, d\mu$ and to compute the integral $\int h(z) \, d\mu(z)$ using the change of variables $z = F_s(x) := \exp_x(s\nabla\theta(x))$. Then, the relation between the three functions f, g, h and the properties of $\varphi(t) := -\log \det d(F_t)_x$ combine nicely to give the result. However, the lack of regularity of F prevents us from doing this directly. As a substitute we will use Proposition 5 and the density u_s of $F_s \#(f \, d\mu)$ which satisfies the "change of variable" formula (21). We will then prove that $h \, e^{-V} \geq u_s$ which implies $\int_M h \, d\mu \geq 1$. Although we will no longer speak of optimal transport or of changing variables, for intuition it is important to bear in mind where the following computations came from.

For $u_0(x) := f(x)e^{-V(x)}$ and $u_1(y) := g(y)e^{-V(y)}$, let u_s and K be as in the Proposition 5. Using the notation of that proposition, we see that it is enough to prove for every $x \in K$ that

$$h(\gamma_x(s))e^{-V(\gamma_x(s))} \geq u_s(\gamma_x(s)). \tag{22}$$

Indeed, since $\displaystyle\int_{\{\gamma_x(s)\,;\ x \in K\}} u_s(z) \, dz = 1$, the previous inequality will give $\int h \, d\mu \geq 1$.

Fix $x \in K$ and denote by $\gamma := \gamma_x : [0,1] \longrightarrow M$ the geodesic joining $\gamma(0) = x$ to $\gamma(1) = F(x)$. Let $A(t) := A_x(t)$ be the invertible matrix of Jacobi fields given by Proposition 5. The geodesic γ has constant speed equal to

$$d := d(\gamma(0), \gamma(1)) = |\dot{\gamma}(t)| \qquad \forall t \in [0,1].$$

If we set $\varphi(t) := -\log \det A(t)$ for every $t \in [0,1]$, we have in particular:

$$f(\gamma(0)) \, e^{-V(\gamma(0))} = g(\gamma(1)) \, e^{-V(\gamma(1))-\varphi(1)} \quad \text{and} \quad f(\gamma(0)) \, e^{-V(\gamma(0))} = u_s(\gamma(s)) \, e^{-\varphi(s)}.$$

We have, using the assumption (13) of Theorem 4 with $x = \gamma(0)$ and $y = \gamma(1)$, and the previous equations:

$$\frac{h(\gamma(s))e^{-V(\gamma(s))}}{u_s(\gamma(s))} \geq \frac{f^{1-s}(\gamma(0)) \, g^s(\gamma(1)) \, e^{-V(\gamma(s))-\lambda\,s(1-s)\,d^2/2}}{u_s(\gamma(s))}$$

$$= \exp\Big\{(1-s)V(\gamma(0)) + sV(\gamma(1)) - V(\gamma(s)) + s\varphi(1) - \varphi(s)$$
$$-\lambda\,s(1-s)\,d^2/2\Big\}.$$

Thus, in order to get (22), it is enough to prove that

$$(1-s)V(\gamma(0)) + sV(\gamma(1)) - V(\gamma(s)) \ + \ s\varphi(1) - \varphi(s) \geq \lambda \, s(1-s) \, d^2/2. \qquad (23)$$

Noticing that $\varphi(0) = 0$, we see that (23) is equivalent to

$$(1-s)\alpha(0) + s\alpha(1) - \alpha(s) \geq \lambda \, s(1-s) \, d^2/2. \qquad (24)$$

where,

$$\alpha(t) := V(\gamma(t)) + \varphi(t), \qquad \forall t \in [0,1].$$

We note that, for every $t \in [0,1]$, we have, using Lemma 6 and the assumption (12)

$$
\begin{aligned}
\alpha''(t) &= \mathrm{Hess}_{\gamma(t)} \, V(\dot\gamma(t), \dot\gamma(t)) + \varphi''(t) \\
&\geq \mathrm{Hess}_{\gamma(t)} \, V(\dot\gamma(t), \dot\gamma(t)) + \frac{1}{n}\varphi'(t)^2 + \mathrm{Ric}_{\gamma(t)}(\dot\gamma(t), \dot\gamma(t)) \\
&\geq \left(\mathrm{Hess}_{\gamma(t)} \, V + \mathrm{Ric}_{\gamma(t)}\right)(\dot\gamma(t), \dot\gamma(t)) \\
&\geq \lambda \, |\dot\gamma(t)|^2
\end{aligned}
$$

Thus we have $\alpha'' \geq \lambda \, d^2$ on $[0,1]$. In view of (5), we get (24) and therefore (22). This ends the proof of Theorem 4. $\qquad\square$

4 On logarithmic Sobolev and transport inequalities

For background and applications of logarithmic Sobolev inequalities we refer to the excellent surveys and book by Ledoux [21, 22, 23].

Given a probability measure μ on M, the entropy w.r.t. μ of a nonnegative function $f : M \longrightarrow \mathbb{R}_+$ is defined by

$$
\begin{aligned}
\mathrm{Ent}_\mu(f) &:= \int_M f \log f \, d\mu - \left(\int_M f \, d\mu\right) \log \int_M f \, d\mu \\
&= \int_M f \log f \, d\mu \qquad \text{if} \quad \int_M f \, d\mu = 1. \qquad (25)
\end{aligned}
$$

The logarithmic Sobolev inequality we want to reproduce was proved by Bakry and Emery using semigroup tools.

Theorem 7 (Bakry and Emery [3]) *Let μ be a probability measure on M of the form $d\mu = e^{-V} \, d\mathrm{vol}$, where V and the Ricci curvature verify $\mathrm{Hess}_x \, V + \mathrm{Ric}_x \geq \lambda \, \mathrm{Id}$ for all $x \in M$ and some positive $\lambda > 0$. Then for every smooth function $f : M \to \mathbb{R}_+$ we have:*

$$\mathrm{Ent}_\mu(f) \leq \frac{1}{2\lambda} \int_M \frac{|\nabla f|^2}{f} \, d\mu.$$

Following the argument of Bobkov and Ledoux [6] inspired by [27], it is possible to derive by elementary but clever computations the previous inequality from Theorem 4. It is however instructive to give a direct proof using mass transport. A source of inspiration for us was the work of Otto and Villani [35] (after Otto's pioneering work [34]) on the relation between optimal transportation and log-Sobolev inequalities. In their paper, Otto and Villani gave a formal proof of Theorem 7 using optimal transport. However the arguments of their paper are quite different from those displayed in the present paper since they rely on the displacement interpolation and more precisely on the coupled Hamilton-Jacobi / transport system satisfied by the interpolated densities. The technicalities involved in this approach prevented them from giving a rigorous proof in the Riemannian case. We would like on the contrary to give a simple and direct proof along the lines of [13] and using the same techniques as above. In fact, as in [13], we will prove the following result from which log-Sobolev and transport inequalities can be recovered. As explained in the section to follow, (26) can be viewed as what Villani [43] has called the *above-tangent* formulation of displacement convexity for the entropy.

Theorem 8 *Let μ be a probability measure on M of the form $d\mu = e^{-V} d\mathrm{vol}$ where V and the Ricci curvature verify $\mathrm{Hess}_x V + \mathrm{Ric}_x \geq \lambda \mathrm{Id}$ for all $x \in M$ and some $\lambda \in \mathbb{R}$. Let f and g be two smooth compactly supported nonnegative functions such that $\int_M f \, d\mu = \int_M g \, d\mu = 1$ and let $F(x) := \exp_x(\nabla \theta(x))$ be the optimal map pushing $f \, d\mu$ forward to $g \, d\mu$. Then one has*

$$\mathrm{Ent}_\mu(g) - \mathrm{Ent}_\mu(f) \geq \int \nabla f \cdot \nabla \theta \, d\mu + \frac{\lambda}{2} \int |\nabla \theta|^2 f \, d\mu. \tag{26}$$

Let us explain how the logarithmic Sobolev inequality of Theorem 7 follows from Theorem 8. Without loss of generality we can assume that f is compactly supported and that $\int f \, d\mu = 1$. We now apply Theorem 8 with an arbitrary g. We use $a^2/2 + b^2/2 \geq -ab$ in (26) in order to get

$$\mathrm{Ent}_\mu(g) - \mathrm{Ent}_\mu(f) \geq -\frac{1}{2\lambda} \int_M \frac{|\nabla f|^2}{f} \, d\mu.$$

Letting $g \longrightarrow 1$ (or equivalently, taking the infimum over all g's) we find the desired inequality

$$\frac{1}{2\lambda} \int_M \frac{|\nabla f|^2}{f} \, d\mu \geq \mathrm{Ent}_\mu(f).$$

Proof of Theorem 8. We will use the result and the notation of Proposition 5 relating the probability densities $u_0(x) := f(x)e^{-V(x)}$ and $u_1(y) := g(y)e^{-V(y)}$ with the optimal map $F(x) = \exp_x(\nabla \theta(x))$ pushing forward $u_0(x) \, dx$ to $u_1(y) \, dy$. By definition of the image measure we have

$$\mathrm{Ent}_\mu(g) - \mathrm{Ent}_\mu(f) = \int (\log g \circ F - \log f) \, f \, d\mu.$$

We fix $x \in K$, $\gamma = \gamma_x$ and $A = A_x$ as in Proposition 5 (we shall not use the information about $s \in [0,1]$). We have, using $F(x) = \gamma(1)$ and $f(x)e^{-V(x)} = g(\gamma(1))e^{-V(\gamma(1))} \det A(1)$,

$$
\begin{aligned}
(\log g \circ F - \log f)(x) &= \log g(\gamma(1)) - \log f(\gamma(0)) \\
&= V(\gamma(1)) - V(\gamma(0)) - \log \big(\det A(1) \big) \\
&= \alpha(1) - \alpha(0)
\end{aligned}
$$

where $\alpha(t) := V(\gamma(t)) - \log \det A(t)$ for every $t \in [0,1]$. Setting $\varphi(t) = -\log \det A(t)$ and combining (as in the proof of Theorem 4) Lemma 6 with the assumption on the curvature, we see that

$$
\begin{aligned}
\alpha''(t) &= \operatorname{Hess}_{\gamma(t)} V(\dot\gamma(t), \dot\gamma(t)) + \varphi''(t) \\
&\geq \operatorname{Hess}_{\gamma(t)} V(\dot\gamma(t), \dot\gamma(t)) + \frac{1}{n}\varphi'(t)^2 + \operatorname{Ric}_{\gamma(t)}(\dot\gamma(t), \dot\gamma(t)) \\
&\geq \big(\operatorname{Hess}_{\gamma(t)} V + \operatorname{Ric}_{\gamma(t)} \big)(\dot\gamma(t), \dot\gamma(t)) \\
&\geq \lambda |\dot\gamma(t)|^2 \\
&= \lambda\, d^2(x, F(x)).
\end{aligned}
$$

Using (5), this implies that $\alpha(1) - \alpha(0) \geq \alpha'(0) + \lambda d^2(x, F(x))/2$. Back to the original notations and remembering that $\gamma(t) = F_t(x) := \exp_x(t\nabla\theta(x))$ we have

$$
(\log g \circ F - \log f)(x) \geq \nabla V(x) \cdot \nabla\theta - \Delta\theta + \lambda|\nabla\theta(x)|^2/2. \tag{27}
$$

Here we used that the derivative of the determinant at the identity is the trace in order to write $\varphi'(0) = -\Delta\theta$. Because $\operatorname{Hess}\theta$ was defined in a weak form, the function $\Delta\theta$ is only the absolutely continuous part of the distributional Laplacian. But since θ is locally semi-convex, it is easy to see that one has $\Delta\theta \leq \Delta_{\mathcal{D}'}\theta$ in the sense of distributions, where $\Delta_{\mathcal{D}'}\theta$ is the distributional Laplacian. Thus we can integrate by parts, using that θ is locally Lipschitz and the definition of $\Delta_{\mathcal{D}'}$ to get

$$
\int_M (\nabla V(x) \cdot \nabla\theta - \Delta\theta)\, f(x)\, e^{-V(x)}\, dx \geq \int_M \nabla f \cdot \nabla\theta\, e^{-V}.
$$

Integration of (27) with respect to $f\, d\mu$ therefore gives

$$
\int_M (\log g \circ F - \log f)\, f\, d\mu \geq \int_M \nabla f \cdot \nabla\theta\, d\mu + \frac{\lambda}{2}\int_M |\nabla\theta|^2\, f\, d\mu.
$$

$\square$

Let us also comment here and in the next section on transport inequalities. As was already mentioned (but not explicitly used), the map $F(x) := \exp_x(\nabla\theta(x))$ exhibited by McCann [32] is optimal in some sense. Let $\mathcal{P}_{ac}(M)$ be the space of probability densities, compactly supported for convenience:

$$
\mathcal{P}_{ac}(M) := \left\{ u \in L^1(M)\,;\ u \geq 0\ \text{compactly supported with}\ \int_M u = 1 \right\}. \tag{28}
$$

We shall identify without further comment a density $u \in \mathcal{P}_{ac}(M)$ and the probability measure u with that density. One can define on $\mathcal{P}_{ac}(M)$ the so-called *Wasserstein distance* $d_2(u_0, u_1)$ between two elements of $\mathcal{P}_{ac}(M)$ as follows:

$$d_2^2(u_0, u_1) := \inf_{\pi \in \Gamma(u_0, u_1)} \int_{M \times M} d^2(x, y) \, d\pi(x, y). \tag{29}$$

Here the unsubscripted $d(x, y)$ denotes the Riemannian geodesic distance, while $\Gamma(u_0, u_1)$ denotes the set of Borel probability measures on $M \times M$ with marginals u_0 and u_1, respectively. In particular, if a Borel map $G : M \longrightarrow M$ pushes u_0 forward to u_1, we see by taking $\pi := (I \times G)(u_0)$ that

$$d_2^2(u_0, u_1) \leq \int_M d^2(x, G(x)) \, u_0(x) \, dx.$$

The map $F(x) = \exp_x(\nabla\theta(x))$ given in [32] with a $d^2/2$-concave $-\theta$ is precisely characterized among all Borel maps pushing u_0 forward to u_1 by the property that

$$d_2^2(u_0, u_1) = \int_M d^2(x, F(x)) \, u_0(x) \, dx = \int_M |\nabla\theta|^2 \, u_0. \tag{30}$$

Now, let us fix as before a reference measure μ of the form $d\mu = e^{-V} d\mathrm{vol}$. For simplicity we will assume that μ is compactly supported, meaning $V = +\infty$ a.e. outside a compact set. Let $g \geq 0$ be such that $\int_M g \, d\mu = 1$ (in other words, $g \, e^{-V} \in \mathcal{P}_{ac}(M)$). We apply Theorem 8 with $f \equiv 1$. Then (26) becomes

$$\mathrm{Ent}_\mu(g) \geq \frac{\lambda}{2} \int |\nabla\theta|^2 \, d\mu,$$

where $F(x) = \exp_x(\nabla\theta(x))$ pushes μ forward to $g \, d\mu$. Thus we have proved the following result.

Theorem 9 (Riemannian transportation inequality) *Let μ be a (compactly supported) probability measure on M of the form $d\mu = e^{-V} d\mathrm{vol}$ where V and the Ricci curvature verify* $\mathrm{Hess}_x V + \mathrm{Ric}_x \geq \lambda \, \mathrm{Id}$ *for all $x \in M$ and some positive $\lambda > 0$. Then, for every $g \geq 0$ with* $\int_M g \, d\mu = 1$ *one has:*

$$d_2^2(g \, d\mu, \mu) \leq \frac{2}{\lambda} \mathrm{Ent}_\mu(g).$$

This result easily extends to non-compactly supported probabilities μ. This kind of inequality is called a *transportation inequality*. We refer to [22, 23] for background and references. We mention however that Bobkov, Gentil and Ledoux [7] showed that there exist many connections between transportation inequalities, Prékopa-Leindler inequalities and hypercontractivity of Hamilton-Jacobi equations. Also, it is easy to deduce Otto and Villani's "HWI"-type inequalities [35] from Theorem 8 as in [13].

Finally, we would like to comment on the sharpness of the results. In general, the result of Theorem 7 is optimal (think of the Gaussian measure on $\mathbb{R}^n$). However, when one works on a compact n-dimensional manifold with $\mu := \sigma = 1/\mathrm{vol}(M)$, the (normalized) uniform Riemannian probability measure on M (which means that V is constant), the result can be significantly improved (see [3, 21]). Indeed, if $\mathrm{Ric}_x \geq \lambda \, \mathrm{Id}$ for all $x \in M$ and some $\lambda > 0$, then for every smooth $f : M \longrightarrow \mathbb{R}_+$ one has

$$\mathrm{Ent}_\sigma(f) \leq \frac{n-1}{2\,n\,\lambda} \int_M \frac{|\nabla f|^2}{f} \, d\sigma.$$

On the sphere S^n of constant Ricci curvature equal to $(n-1)I$, the inequality takes the form

$$\mathrm{Ent}_\sigma(f) \leq \frac{1}{2n} \int_{S^n} \frac{|\nabla f|^2}{f} \, d\sigma.$$

This is a sharp inequality since after a Taylor expansion around the constant function, we recover that the spherical Laplacian has a spectral gap of size n. A direct application of Theorem 7 would instead give gap $n-1$. We do not presently see how to reach such sharp constants using a mass transportation approach. Perhaps a more clever variant or application of Lemma 6 is needed.

5 The *displacement convexity* point of view

This section illuminates some of the connections between the inequalities we have discussed and the notion of displacement convexity [31]. For some recent developments concerning displacement convexity, see [11] and the references there.

The space $\mathcal{P}_{ac}(M)$ of compactly supported probability measures on M forms a metric space when endowed with the Wasserstein distance (29) of the preceding section. Moreover, to each pair of measures $u_0, u_1 \in \mathcal{P}_{ac}(M)$ corresponds a unique path $s \in [0,1] \longrightarrow u_s \in \mathcal{P}_{ac}(M)$ which interpolates between them and forms a geodesic in the sense that

$$d_2(u_0, u_s) = s d_2(u_0, u_1) \quad \text{and} \quad d_2(u_s, u_1) = (1-s) d_2(u_0, u_1) \qquad \text{when } s \in [0,1]. \quad (31)$$

In fact, $u_s = F_s(u_0)$ is nothing but the image measure of u_0 under the map F_s of Proposition 5 [14]. Thus the path is constructed by gradually displacing the mass of u_0 at x along the geodesic on M which leads to $F(x)$, where it eventually forms $u_1 = F(u_0)$. The metric space $\mathcal{P}_{ac}(M)$ forms a *length space*, and the path u_s is called either the Wasserstein *geodesic* or *displacement interpolation* between u_0 and u_1.

A functional $E : \mathcal{P}_{ac}(M) \longrightarrow \mathbb{R} \cup \{+\infty\}$ is said to be *displacement convex* if $E(u_s)$ is a convex function of $s \in [0,1]$ along each Wasserstein geodesic; in other words, if (31) implies $E(u_s) \leq (1-s)E(u_0) + sE(u_1)$. The functional $E(u)$ is said to be displacement *semiconvex* with constant $\lambda \in \mathbb{R}$ if for every u_0, u_1 and $s \in [0,1]$, (31) implies

$$(1-s)E(u_0) + sE(u_1) - E(u_s) \geq \lambda \, s(1-s) d_2^2(u_0, u_1)/2, \qquad (32)$$

or equivalently, if $d^2 E(u_s)/ds^2 \geq \lambda d_2^2(u_0, u_1)$ in the distributional sense. Note that with this terminology, displacement *semi*convexity implies displacement *convexity* when $\lambda \geq 0$; moreover, this convexity is uniform when $\lambda > 0$.

We now rigorously prove the displacement semiconvexity of the Boltzmann type entropy in a Riemannian setting — another result derived formally by Otto and Villani [35]. The case $V = 0$ has been treated by Sturm and von Renesse [42] also, exploiting [14]. We recently learned that the next result was independently proved by Sturm [40] and by Lott and Villani [25] along their investigation of curvature in abstract length spaces. Interestingly enough, they also proved a converse statement, namely that a lower bound on the displacement convexity constant of the entropy in turn gives a lower bound for the curvature tensor.

Theorem 10 (Entropy is displacement semiconvex in a Riemannian setting)
Assume there exists $\lambda \in \mathbb{R}$ such that $V : M \longrightarrow \mathbb{R}$ satisfies $\operatorname{Hess}_x V + \operatorname{Ric}_x \geq \lambda I$ for all $x \in M$. Then the functional $E(u)$ defined on $\mathcal{P}_{ac}(M)$ by (15) is displacement semiconvex with constant λ.

Proof. The proof is similar to that of Theorem 8. Fix $u_0, u_1 \in \mathcal{P}_{ac}(M)$ and $s \in [0, 1]$. Let $F(x) = \exp_x(\nabla\theta(x))$ be the optimal map pushing u_0 forward to u_1 and set $F_s(x) := \exp_x(s\nabla\theta(x))$. We will use the notation and results of Proposition 5. We have, using the definition of the image measure and taking the log of the equations $u_0(x) = u_t(F_t(x)) \det A_x(t)$ for $t \in \{0, s, 1\}$,

$$
\begin{aligned}
(1-s)E(u_0) + sE(u_1) - E(u_s) &= \int_M \Big[(1-s)\log u_0(x) + s\log u_1(F(x)) - \log u_s(F_s(x)) \\
&\qquad\qquad + (1-s)V(x) + sV(F(x)) - V(F_s(x)) \Big] u_0(x)\, dx \\
&= \int_K \Big[(1-s)\alpha_x(0) + s\alpha_x(1) - \alpha_x(s) \Big] u_0(x)\, dx
\end{aligned}
$$

where $\alpha_x(t) := -\log\det(A_x(t)) + V(\gamma_x(t))$ for $x \in K$ and $t \in [0, 1]$. As in the previous proofs, we can use Lemma 6 to see that $\alpha_x'' \geq \lambda d^2(x, F(x))$ on $[0, 1]$. Indeed, for a fixed $x \in K$, we have, suppressing the dependence on x and setting $\varphi(t) := -\log\det A(t)$,

$$
\begin{aligned}
\alpha''(t) &= \operatorname{Hess}_{\gamma(t)} V(\dot\gamma(t), \dot\gamma(t)) + \varphi''(t) \\
&\geq \operatorname{Hess}_{\gamma(t)} V(\dot\gamma(t), \dot\gamma(t)) + \frac{1}{n}\varphi'(t)^2 + \operatorname{Ric}_{\gamma(t)}(\dot\gamma(t), \dot\gamma(t)) \\
&\geq \big(\operatorname{Hess}_{\gamma(t)} V + \operatorname{Ric}_{\gamma(t)} \big) (\dot\gamma(t), \dot\gamma(t)) \\
&\geq \lambda\, |\dot\gamma(t)|^2 \\
&= \lambda\, d^2(x, F(x)).
\end{aligned}
$$

This gives, by integral Taylor expansion,

$$
(1-s)\alpha_x(0) + s\alpha_x(1) - \alpha_x(s) \geq \lambda\, s(1-s)\, d^2(x, F(x))/2.
$$

17

Integrating this inequality with respect to u_0 on K gives the desired result (32). $\qquad\square$

To elucidate the connection of displacement semiconvexity with logarithmic Sobolev and transportation inequalities, let us rewrite the absolute entropy (15) in the form

$$
\begin{aligned}
E(u) &= \int_M u[\log u - \log \mu]\, dx \\
&= \int_M \frac{u}{\mu}\left[\log \frac{u}{\mu}\right]\mu(x)\, dx
\end{aligned}
$$

where $\mu := e^{-V(x)}$. If μ is a probability measure, then $E(u) = \mathrm{Ent}_\mu(f)$ is nothing but the relative entropy (25), with $f = u/\mu$ the Radon-Nikodym derivative of u with respect to the reference measure μ. If μ is not a probability measure but has finite total mass, it can always be normalized by adding a constant to $V(x)$ without affecting the displacement (semi-)convexity of $E(u)$.

By Taylor's theorem, any semiconvex function $h : \mathbb{R} \longrightarrow \mathbb{R}$ with constant $\lambda \in \mathbb{R}$ satisfies for all $t \in \mathbb{R}$

$$
h(t_1) \geq h(t_0) + h'(t_0)(t_1 - t_0) + \lambda(t_1 - t_0)^2/2.
$$

The second order correction bounds how low $h(t_1)$ can be relative to the tangent estimate of $h(t)$ from t_0. The same inequality extends to semiconvex functions which are not C^2-smooth by an approximation argument, with $h'(t_0)$ intrepreted as a directional derivative from the left or the right. Applied to the semiconvex function $s \mapsto E(u_s)$ of Theorem 10, this yields

$$
E(u_1) \geq E(u_0) + \left.\frac{dE(u_s)}{ds}\right|_{s=0^+} + \lambda d_2^2(u_0, u_1)/2
$$

for every Wasserstein geodesic $\{u_s\}_{s\in[0,1]}$ in $\mathcal{P}_{ac}(M)$. Since the endpoints of $f_s = u_s/\mu$ can be arbitrary probability densities, the identities $E(u_s) = \mathrm{Ent}_\mu(f_s)$ and (30) yield

$$
\mathrm{Ent}_\mu(f_1) - \mathrm{Ent}_\mu(f_0) \geq \left.\frac{d\mathrm{Ent}_\mu(f_s)}{ds}\right|_{s=0^+} + \frac{\lambda}{2}\int_M |\nabla\theta|^2 f_0 d\mu.
$$

The form of this inequality already suggests the fundamental result (26) of [13] and the preceding section. To complete the identification, it remains only to find the directional derivative of $E(u_s)$, or at least show

$$
\left.\frac{dE(u_s)}{ds}\right|_{s=0^+} \geq \int \nabla f_0 \cdot \nabla\theta\, d\mu. \tag{33}
$$

This can be accomplished using the transport equation

$$
\left.\frac{\partial u_s}{\partial s}\right|_{s=0} + \nabla \cdot [u_0\nabla\theta] = 0 \tag{34}
$$

which arises as the linearization of (20), and is satisfied distributionally. Indeed,

$$
\begin{aligned}
\left.\frac{dE(u_s)}{ds}\right|_{s=0^+} &= \int \left.\frac{\partial u_s}{\partial s}\right|_{s=0^+} \frac{\delta E}{\delta u} \\
&= \int u_0 \nabla\theta \cdot \nabla(\log u_0 + V + 1) \\
&= \int f_0 \mu \nabla\theta \cdot \nabla(\log(f_0\mu) - \log\mu) \\
&= \int \mu \nabla\theta \cdot \nabla f_0
\end{aligned}
$$

formally. However, a rigorous justification of these identities to obtain (33) requires appropriate hypotheses and care [34, 2, 11].

Acknowledgements

References

[1] S. ALESKER, S. DAR AND V. MILMAN, A remarkable measure preserving diffeomorphism between two convex bodies in $\mathbb{R}^n$, *Geom. Dedicata* 74 (1999) 201–212.

[2] L.A. AMBROSIO, N. GIGLI AND G. SAVARÉ, Gradient flows with metric and differentiable structures, and applications to the Wasserstein space. To appear in the Academy of Lincei proceedings on "Nonlinear evolution equations", Rome.

[3] D. BAKRY AND M. EMERY, Diffusions hypercontractives, in *Séminaire de Probabilités XIX*, Lecture Notes in Math., 1123, Springer (1985), 177–206.

[4] K.M. BALL, An elementary introduction to modern convex geometry, in *Flavors of geometry*, Math. Sci. Res. Inst. Publ. 31, Cambridge Univ. Press (1997), 1–58.

[5] F. BARTHE, On a reverse form of the Brascamp-Lieb inequality, *Invent. Math.* 134, 2 (1998), 335–361.

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

[7] S. BOBKOV, I. GENTIL AND M. LEDOUX, Hypercontractivity of Hamilton-Jacobi equations, *J. Math. Pures Appl.* (9) 80 (2001), no. 7, 669–696.

[8] C. Borell, Convex set functions in d-space, *Period. Math. Hungar.* 6 (1975), 111–136.

[9] H.J. Brascamp and E.H. Lieb, On extensions of the Brunn-Minkowski and Prékopa-Leindler theorems, including inequalities for log concave functions, and with an application to the diffusion equation *J. Funct. Anal.* 22 (1976), 366–389.

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

[11] J.A. Carrillo, R.J. McCann and C. Villani, Contractions in the 2-Wasserstein length space and thermalization of granular media, *Arch. Rational Mech. Anal.,* to appear.

[12] I. Chavel, *Riemannian Geometry—a Modern Introduction*, Cambridge Tracts in Math. 108, Cambridge University Press, Cambridge, 1993.

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

[14] D. Cordero-Erausquin, R.J. McCann and M. Schmuckenschläger, A Riemannian interpolation inequality à la Borell, Brascamp and Lieb, *Invent. Math.* 146 (2001), 219–257.

[15] D. Cordero-Erausqauin, B. Nazaret and C. Villani, A mass-transportation approach to sharp Sobolev and Gagliardo-Nirenberg inequalities, *Adv. Math.* 182 (2004), no. 2, 307–332.

[16] S. Das Gupta, Brunn-Minkowski inequality and its aftermath, *J. Multivariate Anal.* 10 (1980), 296–318.

[17] S. Gallot, D. Hulin and J. Lafontaine, *Riemannian Geometry*, Springer-Verlag (Universitext), Berlin, 1990.

[18] R.J. Gardner, The Brunn-Minkowski inequality, *Bull. Amer. Math. Soc. (N.S.)* 39 (2002), no. 3, 355–405.

[19] M. Gromov and V. Milman, A topological application of the isoperimetric inequality, *Amer. J. Math* 105 (1983), 843–854.

[20] H. Knothe, Contributions to the theory of convex bodies, *Michigan Math. J.* 4 (1957), 39–52.

[21] M. Ledoux, The geometry of Markov diffusion generators, *Ann. Fac. Sci. Toulouse Math.* (6) 9 (2000), no. 2, 305–366.

[22] M. LEDOUX, *Measure concentration, transportation cost, and functional inequalities*, Summer School on Singular Phenomena and Scaling in Mathematical Models, Bonn, 10-13 June 2003 (http://www.lsp.ups-tlse.fr/Ledoux).

[23] M. LEDOUX, *The concentration of measure phenomenon*, American Mathematical Society, Providence, RI, 2001.

[24] L. LEINDLER, On a certain converse of Hölder's inequality, *Acta Sci. Math.* 33 (1972), 217–233.

[25] J. LOTT AND C. VILLANI, Ricci curvature for metric-measure spaces via optimal transport, *Preprint.*

[26] F. MAGGI AND C. VILLANI, Balls have the worst best Sobolev inequality, *Preprint.*

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

[28] B. MAUREY, Inégalité de Brunn-Minkowski-Lusternik, et autres inégalités géométriques et fonctionnelles, *Séminaire Bourbaki*, Novembre 2003.

[29] R.J. MCCANN, *A Convexity Principle for Interacting Gases and Equilibrium Crystals,* PhD Thesis, Princeton University (1994).

[30] R.J. MCCANN, Existence and uniqueness of monotone measure-preserving maps, *Duke. Math. J.* 80 (1995), 309–323.

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

[32] R.J. MCCANN, Polar factorization of maps on Riemannian manifolds, *Geom. Funct. Anal.* 11 (2001), no. 3, 589–608.

[33] V.D. MILMAN AND G. SCHECHTMAN, *Asymptotic theory of finite-dimensional normed spaces*, Springer-Verlag, Berlin, 1986. With an appendix by M. Gromov.

[34] F. OTTO, The geometry of dissipative evolution equations: the porous medium equation, *Comm. Partial Differential Equations* 26 (2001), no. 1-2, 101–174.

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

[36] A. PRÉKOPA, Logarithmic concave measures with application to stochastic programming, *Acta Sci. Math.* 32 (1971), 301–315.

[37] A. PRÉKOPA, On logarithmic concave measures and functions, *Acta Sci. Math. (Szeged)* 34 (1973), 335–343.

[38] M. SCHMUCKENSCHLÄGER, A concentration of measure phenomenon on uniformly convex bodies, in *GAFA Seminar (1992-1994)*, Oper. Theory Adv. Appl. 77, 275–287, Birkaäuser, 1995.

[39] R. SCHNEIDER, *Convex Bodies: the Brunn-Minkowski Theory,* Cambridge University Press, Cambridge, 1993.

[40] K.T. STURM, Convex functionals of probability measures and nonlinear diffusions on manifolds,. *J. Math. Pures Appl.* 84 (2005).

[41] N.S. TRUDINGER, Isoperimetric inequalities for quermassintegrals, *Ann. Inst. H. Poincaré Anal. Non Linéaire* 11 (1994), 411–425.

[42] K.-T. STURM AND M.-K. VON RENESSE, Transport inequalities, gradient estimates, entropy and Ricci curvature, *Comm. Pure Appl. Math.* 58 (2005), 923–940.

[43] C. VILLANI, *Topics in Optimal Transportation,* Graduate Studies in Math. 58, American Mathematical Society, Providence, RI, 2003.