STRONG DISPLACEMENT CONVEXITY ON RIEMANNIAN MANIFOLDS

A. FIGALLI AND C. VILLANI

ABSTRACT. Ricci curvature bounds in Riemannian geometry are known to be equivalent to the weak convexity (convexity along at least one geodesic between any two points) of certain functionals in the space of probability measures. We prove that the weak convexity can be reinforced into strong (usual) convexity, thus solving a question left open in [4].

1. Introduction and main result

For the past few years, there has been ongoing research to study the links between Riemannian geometry and optimal transport of measures [9, 10]. In particular, it was recently found that lower bounds on the Ricci curvature tensor can be recast in terms of convexity properties of certain nonlinear functionals defined on spaces of probability measures [1, 4, 5, 6, 7, 8]. In this paper we solve a natural problem in this field by establishing the equivalence of several such formulations.

Before explaining our results in more detail, let us give some notation and background. Let (M, g) be a smooth complete connected n-dimensional Riemannian manifold, equipped with its geodesic distance d and its volume measure vol. Let $P(M)$ be the set of probability measures on M. For any real number $p \geq 1$, we denote by $P_p(M)$ the set of probability measures μ such that

$$\int_M d^p(x, x_0)\, d\mu(x) < \infty \qquad \text{for some } x_0 \in M.$$

The set $P_2(M)$ is equipped with the Wasserstein distance of order 2, denoted by W_2: This is the square root of the optimal transport cost functional, when the cost function $c(x, y)$ coincides with the squared distance $d^2(x, y)$; see for instance [10, Definition 6.1]. Then $P_2(M)$ is a metric space, and even a length space; that is, any two probability measures in $P_2(M)$ are joined by at least one geodesic curve $(\mu_t)_{0\leq t\leq 1}$. (Here and in the sequel, by convention geodesics are supposed to be globally minimizing and to have constant speed.)

A basic representation theorem (see [4, Proposition 2.10] or [10, Corollary 7.22]) states that any Wasserstein geodesic curve necessarily takes the form $\mu_t = (e_t)_*\Pi$, where Π is a probability measure on the set Γ of minimizing geodesics $[0, 1] \to M$, the symbol $*$ stands for push-forward, and $e_t : \Gamma \to M$ is the evaluation at time t: $e_t(\gamma) := \gamma(t)$. So the optimal transport problem between two probability measures μ_0 and μ_1 produces three related objects:

- an optimal coupling π of μ_0 and μ_1, which is a probability measure on $M \times M$ whose marginals are μ_0 and μ_1, achieving the lowest possible cost for the transport between these measures;

- a path $(\mu_t)_{0\leq t\leq 1}$ in the space of probability measures;

- a probability measure Π on the space of geodesics, such that $(e_t)_*\Pi = \mu_t$ and $(e_0, e_1)_*\Pi = \pi$. Such a Π is called a dynamical optimal transference plan [10, Definition 7.20].

The core of the studies in [1, 4, 5, 6, 7, 8] lies in the analysis of the convexity properties of certain nonlinear functionals along geodesics in $P_2(M)$, defined below:

Definition 1.1 (Nonlinear functionals of probability measures)**.** *Let ν be a reference measure on M, absolutely continuous with respect to the volume measure. Let $U : \mathbb{R}_+ \to \mathbb{R}$ be a continuous convex function with $U(0) = 0$; let $U'(\infty)$ be the limit of $U(r)/r$ as $r \to \infty$. Let μ be a probability measure on M and let $\mu = \rho\nu + \mu_s$ be its Lebesgue decomposition with respect to ν.*

(i) If $U(\rho)$ is bounded below by a ν-integrable function, then the quantity $U_\nu(\mu)$ is defined by the formula

$$U_\nu(\mu) = \int_M U(\rho(x))\,\nu(dx) + U'(\infty)\,\mu_s[M].$$

(ii) If π is a probability measure on $M \times M$, admitting μ as first marginal, β is a positive function on $M \times M$, and $\beta\, U(\rho/\beta)$ is bounded below (as a function of x, y) by a ν-integrable function of x, then the quantity $U^\beta_{\pi,\nu}(\mu)$ is defined by the formula

$$U^\beta_{\pi,\nu}(\mu) = \int_{M\times M} U\left(\frac{\rho(x)}{\beta(x,y)}\right)\beta(x,y)\pi(dy|x)\,\nu(dx) + U'(\infty)\,\mu_s[M],$$

where $\pi(dy|x)$ is the disintegration of $\pi(dx\,dy)$ with respect to the x variable.

Remark 1.2. Sufficient conditions for U_ν and $U^\beta_{\pi,\nu}$ to be well-defined are discussed in [10, Theorems 17.8 and 17.28, Application 17.29] and will not be addressed here.

Remark 1.3. If $U'(\infty) = \infty$, then finiteness of $U_\nu(\mu)$ implies that μ is absolutely continuous with respect to ν. This is not true if $U'(\infty) < \infty$.

The various notions of convexity that are considered in [4, 6, 7, 8] belong to the following ones:

Definition 1.4 (Convexity properties)**.** *(i) Let U and ν be as in Definition 1.1, and let $\lambda \in \mathbb{R}$. We say that the functional U_ν is* λ-displacement convex *if for all Wasserstein geodesics $(\mu_t)_{0\le t\le 1}$ whose image lies in the domain of U_ν,*

$$U_\nu(\mu_t) \le (1-t)\,U_\nu(\mu_0) + t\,U_\nu(\mu_1) - \frac{1}{2}\,\lambda\,t(1-t)W_2^2(\mu_0,\mu_1), \qquad \forall t \in [0,1]. \tag{1}$$

We say that the functional U_ν is displacement convex with distortion β *if for all Wasserstein geodesics $(\mu_t)_{0\le t\le 1}$ whose image lies in the domain of U_ν, if $\pi(dx\,dy)$ stands for the associated optimal coupling between μ_0 and μ_1, and $\check{\pi}$ is obtained from π by exchanging the two variables x and y, then*

$$U_\nu(\mu_t) \le (1-t)\,U^\beta_{\pi,\nu}(\mu_0) + t\,U^\beta_{\check{\pi},\nu}(\mu_1), \qquad \forall t \in [0,1]. \tag{2}$$

(ii) We say that U_ν is weakly λ-displacement convex *(resp.* weakly displacement convex with distortion β*) if for all μ_0, μ_1 in the domain of U_ν, there is some Wasserstein geodesic from μ_0 to μ_1 along which* (1) *(resp.* (2)*) is satisfied.*

(iii) We say that U_ν is weakly λ-a.c.c.s. displacement convex *(resp.* weakly a.c.c.s. displacement convex with distortion β*) if condition* (1) *(resp.* (2)*) is satisfied along some Wasserstein geodesic when we further assume that μ_0, μ_1 are absolutely continuous and compactly supported.*

Remark 1.5. *The Wasserstein geodesic in (ii) and (iii) above is implicitly assumed to have its image entirely contained in the domain of the functional U_ν.*

Remark 1.6. *If U_ν is a λ-displacement convex functional, then the function $t \mapsto U_\nu(\mu_t)$ is λ-convex on $[0,1]$, i.e. for all $0 \le s \le s' \le 1$ and $t \in [0,1]$,*

$$U_\nu(\mu_{(1-t)s+ts'}) \le (1-t)U_\nu(\mu_s) + tU_\nu(\mu_{s'}) - \frac{1}{2}\lambda t(1-t)(s'-s)^2 W_2(\mu_0,\mu_1)^2. \tag{3}$$

This is not a priori *the case if we only assume that U_ν is weakly λ-displacement convex.*

In short, *weakly* means that we require a condition to hold only for *some* geodesic between two measures, as opposed to *all* geodesics, and *a.c.c.s.* means that we only require the condition to hold when the two measures are absolutely continuous and compactly supported.

There are obvious implications (with or without distorsion)

$$\begin{array}{c}\lambda\text{-displacement convex}\\ \Downarrow\\ \text{weakly } \lambda\text{-displacement convex}\\ \Downarrow\\ \text{weakly } \lambda\text{-a.c.c.s. displacement convex.}\end{array}$$

Although the natural convexity condition is arguably the one appearing in (i), that is, holding true along all Wasserstein geodesics, this condition is quite more delicate to study than the weaker conditions appearing in (ii) and (iii), in particular for stability issues: See [4, 6, 7]. In the same references the equivalence between (ii) and (iii) was established, at least for compact spaces [4, Proposition 3.21]. But the implication (ii) $\Rightarrow$ (i) remained open (and was listed as an open problem in a preliminary version of [10]). In the present paper we shall fill this gap (at least for the functionals defined above), thus answering a natural question about the notion of displacement convexity. Here is our main result:

Theorem 1.7. *Let U, ν and β be as in Definition 1.1. Assume that U is Lipschitz. For each $a > 0$, define $U_a(r) = U(ar)/a$. Then*

(i) If $(U_a)_\nu$ is weakly λ-a.c.c.s. displacement convex for any $a \in (0,1]$, then U_ν is λ-displacement convex;

(ii) If $(U_a)_\nu$ is weakly a.c.c.s. displacement convex with distortion β for any $a \in (0,1]$, then U_ν is displacement convex with distortion β.

Among the consequences of Theorem 1.7 is the following corollary:

Corollary 1.8. *Let M be a smooth complete Riemannian manifold with nonnegative Ricci curvature and dimension n. Let $U(r) = -r^{1-1/n}$, and let ν be the volume measure on M. Then U_ν is displacement convex on $P_p(M)$, where $p = 2$ if $n \ge 3$, and p is any real number greater than 2 if $n = 2$.*

More generally, Theorem 1.7 makes it possible to drop the "weakly" in all displacement convexity characterizations of Ricci curvature bounds.

Before turning to the proof of Theorem 1.7, let us explain a bit more about the difficulties and the strategy of proof. Obviously, there are two problems to tackle: first, the possibility that μ_0 and/or μ_1 do not have compact support; and secondly, the possibility that μ_0 and/or μ_1 are singular with respect to the volume measure.

It was shown in [4, 6, 7] that inequalities such as (1) or (2) are *stable* under (weak) convergence. Then it is natural to approximate μ_0, μ_1 by compactly supported, absolutely continuous measures, and pass to the limit. This scheme of proof is enough to show the implication (iii) $\Rightarrow$ (ii) in Definition 1.4, but does not guarantee that we can attain *all* Wasserstein geodesics in this way — unless of course we know that there is a unique Wasserstein geodesic between μ_0 and μ_1.

To treat the difficulty arising from the possible non-compactness, we use recent results by Fathi and the first author [2], showing that the Wasserstein geodesic between any two absolutely continuous probability measures on a Riemannian manifold M is unique, even if they are not compactly supported. (This exact statement does not appear in [2], but it is a simple consequence of the results there, and the reasoning in [3, Proposition 3.1]. See also [10, Corollary 7.23].)

The difficulty arising from the possible singularity of μ_0, μ_1 is less simple. If μ_0 and μ_1 are both singular, then there are in general several Wasserstein geodesics joining them. A most simple example is constructed by taking $\mu_0 = \delta_{x_0}$ and $\mu_1 = \delta_{x_1}$, where δ_x stands for the Dirac mass at x, and x_0, x_1 are joined by multiple geodesics. So it is part of the problem to regularize μ_0, μ_1 into absolutely continuous measures $\mu_{0,k}$, $\mu_{1,k}$ so that, as $k \to \infty$, the optimal transport between $\mu_{0,k}$ and $\mu_{1,k}$ converges to a *given* optimal transport between μ_0 and μ_1.

We handle this by a rather nonstandard regularization procedure, which roughly goes as follows. We start from a given dynamical optimal transference plan Π between μ_0 and μ_1, leave intact that part $\Pi^{(a)}$ of Π which corresponds to the absolutely continuous part of μ_0. Then we let displacement occur for a very short time at the level of that part $\Pi^{(s)}$ of Π corresponding to the singular part of μ_0. Next we regularize the resulting contribution of $\Pi^{(s)}$.

Let us illustrate this in the most basic case when $\mu_0 = \delta_{x_0}$ and $\mu_1 = \delta_{x_1}$. Let $\gamma = (\gamma_t)_{0 \le t \le 1}$ be a given geodesic between x_0 and x_1; we wish to approximate the Wasserstein geodesic $(\delta_{\gamma_t})_{0 \le t \le 1}$. Instead of directly regularizing μ_0 and μ_1, we shall first replace μ_0 by $\mu_\tau = \delta_{\gamma_\tau}$, where τ is positive but very small; and then regularize δ_{γ_τ} and δ_{x_1} into probability measures $\mu_{\tau,\varepsilon}$ and $\mu_{1,\varepsilon}$. What we have gained is that the geodesic joining γ_τ to $x_1 = \gamma_1$ is unique, so we may let $\tau \to 0$ and $\varepsilon \to 0$ in such a way that the Wasserstein geodesic joining $\mu_{\tau,\varepsilon}$ to $\mu_{1,\varepsilon}$ does converge to $(\delta_{\gamma_t})_{0 \le t \le 1}$.

In a more general context, the procedure will be more tricky, and what will make it work is the following important property [10, Theorem 7.29]: *Geodesics in dynamical optimal transport plans do not cross at intermediate times.* In fact, if Π is a given dynamical optimal transport plan, then for each $t \in (0,1)$ one can define a measurable map $F_t : M \to \Gamma$ by the requirement that $F_t \circ e_t = \mathrm{Id}$, Π-almost surely. In understandable words, if γ is a geodesic along which there is optimal transport, then the position of γ at time t determines the whole geodesic γ. This property will ensure that $\Pi^{(a)}$ and $\Pi^{(s)}$ "do not overlap at intermediate times".

Finally, we note that the results in this paper can be extended to more general situations outside the category of Riemannian manifolds: It is sufficient that the optimal transport between any two absolutely continuous probability measures be unique. In fact, there is a more general framework where these results still hold true, namely the case of *nonbranching* locally compact, complete length spaces. This extension will be established, by a slightly different approach, in [10, Chapter 30].

Acknowledgement: John Lott played a double role in the genesis of this paper, first because some of the methods and results used here are taken from his joint work with the second author [4], but also because it was him who explicitly asked whether displacement convexity

inequalities would hold true along all displacement interpolations between two singular measures on a Riemannian manifold.

2. Proofs

In the sequel, we shall use the notation $U_{a,\nu}$ for $(U_a)_\nu$. An important ingredient in the proof of Theorem 1.7 will be the following lemma, which has interest on its own (and will be used for different purposes in [10, Chapter 30]).

Lemma 2.1. *Let U be a Lipschitz convex function with $U(0) = 0$. Let μ_1, μ_2 be any two probability measures on M, and let Z_1, Z_2 be two positive numbers with $Z_1 + Z_2 = 1$. Then*

(i) $U_\nu(Z_1\mu_1 + Z_2\mu_2) \geq Z_1\, U_{Z_1,\nu}(\mu_1) + Z_2\, U_{Z_2,\nu}(\mu_2)$, with equality if μ_1 and μ_2 are singular to each other;

(ii) Let π_1, π_2 be two probability measures on $M \times M$, and let β be a positive measurable function on $M \times M$. Then

$$U^\beta_{Z_1\pi_1+Z_2\pi_2,\nu}(Z_1\mu_1 + Z_2\mu_2) \geq Z_1\, U^\beta_{Z_1,\pi_1,\nu}(\mu_1) + Z_2\, U^\beta_{Z_2,\pi_2,\nu}(\mu_2),$$

with equality if μ_1 and μ_2 are singular to each other.

Proof of Lemma 2.1. We start by the following remark: If x, y are nonnegative numbers, then

$$U(x + y) \geq U(x) + U(y). \tag{4}$$

Inequality (4) follows at once from the fact that $U(t)/t$ is a nondecreasing function of t, and thus

$$\frac{U(x)}{x} \leq \frac{U(x+y)}{x+y},\quad \frac{U(y)}{y} \leq \frac{U(x+y)}{x+y} \Longrightarrow xU(x+y) + yU(x+y) \geq (x+y)(U(x) + U(y)).$$

Next, with obvious notation,

$$U_\nu(Z_1\mu_1 + Z_2\mu_2) = \int U(Z_1\rho_1 + Z_2\rho_2)\, d\nu + U'(\infty)\big(Z_1\,\mu_{1,s}[M] + Z_2\,\mu_{2,s}[M]\big);$$

$$U_{Z_1,\nu}(\mu_1) = \frac{1}{Z_1}\int U(Z_1\rho_1)\, d\nu + U'(\infty)\mu_{1,s}[M];$$

$$U_{Z_2,\nu}(\mu_2) = \frac{1}{Z_2}\int U(Z_2\rho_2)\, d\nu + U'(\infty)\mu_{2,s}[M];$$

so part (i) of the lemma follows immediately from (4). The claim about equality is obvious since it amounts to say that $U(x + y) = U(x) + U(y)$ as soon as either x or y is zero.

The proof of part (ii) is based on a similar type of reasoning. First note that (with the conventions $U(0)/0 = U'(0)$, $U(\infty)/\infty = U'(\infty)$ and μ_s-almost surely, $d\mu/d\nu = +\infty$)

$$\begin{aligned} &U^\beta_{Z_1\pi_1+Z_2\pi_2,\nu}(Z_1\mu_1 + Z_2\mu_2) \\ &\qquad = \int_{M\times M} U\left(\frac{Z_1\rho_1(x) + Z_2\rho_2(x)}{\beta(x,y)}\right) \frac{\beta(x,y)}{Z_1\rho_1(x) + Z_2\rho_2(x)}\,(Z_1\pi_1 + Z_2\pi_2)(dx\,dy); \end{aligned}$$

$$U^\beta_{Z_1,\pi_1,\nu}(\mu_1) = \int U\left(\frac{Z_1\rho_1(x)}{\beta(x,y)}\right) \frac{\beta(x,y)}{Z_1\rho_1(x)}\, Z_1\pi_1(dx\,dy);$$

$$U^\beta_{Z_2,\pi_2,\nu}(\mu_2) = \int U\left(\frac{Z_2\rho_2(x)}{\beta(x,y)}\right) \frac{\beta(x,y)}{Z_2\rho_2(x)}\, \pi_2(dx\,dy).$$

So the proof of the lemma will be complete if we can show that

$$U\left(\frac{Z_1\rho_1+Z_2\rho_2}{\beta}\right)\frac{\beta}{Z_1\rho_1+Z_2\rho_2}\,(Z_1\pi_1+Z_2\pi_2) \geq U\left(\frac{Z_1\rho_1}{\beta}\right)\frac{\beta}{Z_1\rho_1}\,(Z_1\pi_1)+U\left(\frac{Z_2\rho_2}{\beta}\right)\frac{\beta}{Z_2\rho_2}\,(Z_2\pi_2). \tag{5}$$

Since $U(r)/r$ is a nondecreasing function of r, if X_1, X_2, p_1, p_2 are any four nonnegative numbers then

$$\frac{U(X_1+X_2)}{X_1+X_2}(p_1+p_2)\geq \frac{U(X_1)}{X_1}\,p_1+\frac{U(X_2)}{X_2}\,p_2.$$

To recover (5), it suffices to apply the latter inequality with

$$X_1=\frac{Z_1\rho_1(x)}{\beta(x,y)},\quad X_2=\frac{Z_2\rho_2(x)}{\beta(x,y)},$$

$$p_1=\frac{d(Z_1\pi_1)}{d(Z_1\pi_1+Z_2\pi_2)}(x,y),\quad p_2=\frac{d(Z_2\pi_2)}{d(Z_1\pi_1+Z_2\pi_2)}(x,y)$$

and to integrate against $(Z_1\pi_1+Z_2\pi_2)(dx\,dy)$. □

Proof of Theorem 1.7. First we observe that U_ν is well-defined on $P_2(M)$ since, if $\mu=\rho\nu+\mu_s$ is the Lebesgue decomposition of a probability measure $\mu\in P(M)$, then

$$U(\rho)\geq -\|U\|_{\mathrm{Lip}}\,\rho\in L^1(M,\nu).$$

In fact, there is also an upper bound, so U_ν is well-defined on the whole of $P_2(M)$ with values in $\mathbb{R}$. Moreover, by an approximation argument, we may replace the assumptions of weak a.c.c.s. displacement convexity by weak displacement convexity on the whole of $P_2(M)$. (The proof is the same as in [4, Proposition 3.21] (in the compact case) or [10, Theorem 30.5].)

Let μ_0, μ_1 be any two measures in $P_2(M)$, and let Π be an optimal dynamical transference plan between μ_0 and μ_1. Let further

$$\mu_0=\rho_0\,\nu+\mu_{0,s}$$

be the Lebesgue decomposition of μ_0 with respect to ν. Let $E^{(a)}$ and $E^{(s)}$ be two disjoint Borel subsets of M such that $\rho_0\,\nu$ is concentrated on $E^{(a)}$ and $\mu_{0,s}$ is concentrated on $E^{(s)}$. We decompose Π as

$$\Pi=\Pi^{(a)}+\Pi^{(s)}, \tag{6}$$

where

$$\Pi^{(a)}:=\Pi\llcorner\Big\{\gamma\in\Gamma\mid\gamma(0)\in E^{(a)}\Big\},\qquad \Pi^{(s)}:=\Pi\llcorner\Big\{\gamma\in\Gamma\mid\gamma(0)\in E^{(s)}\Big\}.$$

Taking the marginals at time t in (6) we get

$$\mu_t=\mu_t^{(a)}+\mu_t^{(s)}.$$

In the end, we renormalize $\mu_t^{(a)}$ and $\mu_t^{(s)}$ into probability measures: we define

$$Z^{(a)}=\Pi^{(a)}[\Gamma]=\mu_0^{(a)}[M]=\mu_t^{(a)}[M];\qquad Z^{(s)}=\Pi^{(s)}[\Gamma],$$

and

$$\hat{\Pi}^{(a)}:=\frac{\Pi^{(a)}}{Z^{(a)}},\qquad \hat{\mu}_t^{(a)}:=\frac{\mu_t^{(a)}}{Z^{(a)}};\qquad \hat{\Pi}^{(s)}:=\frac{\Pi^{(s)}}{Z^{(s)}},\qquad \hat{\mu}_t^{(s)}:=\frac{\mu_t^{(s)}}{Z^{(s)}}.$$

So

$$\mu_t = Z^{(a)} \hat{\mu}_t^{(a)} + Z^{(s)} \hat{\mu}_t^{(s)}. \tag{7}$$

We remark that by the results in [2] $\mu_t^{(a)}$ is absolutely continuous for any $t \in [0,1)$, but $\mu_t^{(s)}$ is not necessary completely singular.

It follows from [10, Theorem 7.29 (v)] that for any $t \in (0,1)$ there is a Borel map F_t such that $F_t(\gamma_t) = \gamma_0$, $\Pi(d\gamma)$-almost surely. Then $\mu_t^{(s)}$ is concentrated on $F_t^{-1}(E^{(s)})$, while $\mu_t^{(a)}$ is concentrated on $F_t^{-1}(E^{(a)})$; so these measures are singular to each other. Then by Lemma 2.1 and (7), for any $t \in (0,1)$,

$$U_\nu(\mu_t) = Z^{(a)} U_{Z^{(a)},\nu}(\hat{\mu}_t^{(a)}) + Z^{(s)} U_{Z^{(s)},\nu}(\hat{\mu}_t^{(s)}). \tag{8}$$

In the sequel, we focus on part (i) of Theorem 1.7, since the reasoning is quite the same for part (ii). By construction and the restriction property of optimal transport [10, Theorem 7.29], $\hat{\Pi}^{(a)}$ is an optimal dynamical transference plan between $\hat{\mu}_0^{(a)}$ and $\hat{\mu}_1^{(a)}$, and the associated Wasserstein geodesic is $(\hat{\mu}_t^{(a)})_{0\le t\le 1}$. Since by construction $\hat{\mu}_0^{(a)}$ is absolutely continuous, by the results in [2] (or by [3, Proposition 3.1] or by [10, Corollary 7.23]) $(\hat{\mu}_t^{(a)})$ is the *unique* Wasserstein geodesic joining $\hat{\mu}_0^{(a)}$ to $\hat{\mu}_1^{(a)}$. Then we can apply the displacement convexity inequality of the functional $U_{Z^{(a)},\nu}$ along that geodesic:

$$U_{Z^{(a)},\nu}(\hat{\mu}_t^{(a)}) \le (1-t)\, U_{Z^{(a)},\nu}(\hat{\mu}_0^{(a)}) + t\, U_{Z^{(a)},\nu}(\hat{\mu}_1^{(a)}) - \frac{\lambda}{2}\, t\,(1-t)\, W_2^2(\hat{\mu}_0^{(a)}, \hat{\mu}_1^{(a)}). \tag{9}$$

Next, let $\varepsilon_k \to 0$ be a sequence of positive numbers. From the nonbranching property of $P_2(M)$ [10, Corollary 7.31], there is only one Wasserstein geodesic joining $\hat{\mu}_{\varepsilon_k}^{(s)}$ to $\hat{\mu}_1^{(s)}$ and it is obtained by reparameterizing $(\hat{\mu}_t^{(s)})_{\varepsilon_k \le t\le 1}$ on $[0,1]$ (with an affine reparameterization in t). So we can also apply the displacement convexity inequality of the functional $U_{Z^{(s)},\nu}$ along that geodesic, and get

$$\begin{aligned} U_{Z^{(s)},\nu}(\hat{\mu}_t^{(s)}) \le \left(\frac{1-t}{1-\varepsilon_k}\right) U_{Z^{(s)},\nu}(\hat{\mu}_{\varepsilon_k}^{(s)}) + \left(\frac{t-\varepsilon_k}{1-\varepsilon_k}\right) U_{Z^{(s)},\nu}(\hat{\mu}_1^{(s)}) \\ - \frac{\lambda}{2}\,(t-\varepsilon_k)\,(1-t)\, W_2^2(\hat{\mu}_0^{(s)}, \hat{\mu}_1^{(s)}). \end{aligned} \tag{10}$$

(For the latter term we have used the fact that if $(\mu_t)_{0\le t\le 1}$ is any Wasserstein geodesic, then $W_2(\mu_s, \mu_t) = |t-s|\, W_2(\mu_0, \mu_1)$.)

The first term in the right-hand side of (10) can be trivially bounded by $U'(\infty)$, which coincides with $U_{Z^{(s)},\nu}(\hat{\mu}_0^{(s)})$ since $\hat{\mu}_0^{(s)}$ is totally singular. Indeed, since $\frac{U(r)}{r} \leq U'(\infty)$, we have

$$\begin{aligned}
U_{Z^{(s)},\nu}(\hat{\mu}_{\varepsilon_k}^{(s)}) &= \frac{1}{Z^{(s)}} \int_M U\left(Z^{(s)}\hat{\rho}_{\varepsilon_k}^{(s)}\right) d\nu + U'(\infty)\,\hat{\mu}_{\varepsilon_k,s}^{(s)}(M) \\
&= \frac{1}{Z^{(s)}} \int_{\{\hat{\rho}_{\varepsilon_k}^{(s)}>0\}} U\left(Z^{(s)}\hat{\rho}_{\varepsilon_k}^{(s)}\right) d\nu + U'(\infty)\,\hat{\mu}_{\varepsilon_k,s}^{(s)}(M) \\
&= \int_{\{\hat{\rho}_{\varepsilon_k}^{(s)}>0\}} \frac{U\big(Z^{(s)}\hat{\rho}_{\varepsilon_k}^{(s)}\big)}{Z^{(s)}\hat{\rho}_{\varepsilon_k}^{s}} \hat{\rho}_{\varepsilon_k}^{(s)}\, d\nu + U'(\infty)\,\hat{\mu}_{\varepsilon_k,s}^{(s)}(M) \\
&\leq \int_{\{\hat{\rho}_{\varepsilon_k}^{(s)}>0\}} U'(\infty)\hat{\rho}_{\varepsilon_k}^{(s)}\, d\nu + U'(\infty)\,\hat{\mu}_{\varepsilon_k,s}^{(s)}(M) \\
&= U'(\infty)\,\hat{\mu}_{\varepsilon_k}^{(s)}(M) = U'(\infty).
\end{aligned}$$

Then by passing to the lim inf as $k \to \infty$ in (10), we recover

$$U_{Z^{(s)},\nu}(\hat{\mu}_t^{(s)}) \leq (1-t)\,U_{Z^{(s)},\nu}(\hat{\mu}_0^{(s)}) + t\,U_{Z^{(s)},\nu}(\hat{\mu}_1^{(s)}) - \frac{\lambda}{2}\,t(1-t)\,W_2^2(\hat{\mu}_0^{(s)},\hat{\mu}_1^{(s)}). \tag{11}$$

By combining together (8), (9) and (11), we obtain

$$\begin{aligned}
U_\nu(\mu_t) \leq (1-t)\left[Z^{(a)}U_{Z^{(a)},\nu}(\hat{\mu}_0^{(a)}) + Z^{(s)}U_{Z^{(s)},\nu}(\hat{\mu}_0^{(s)})\right] + t\left[Z^{(a)}U_{Z^{(a)},\nu}(\hat{\mu}_1^{(a)}) + Z^{(s)}U_{Z^{(s)},\nu}(\hat{\mu}_1^{(s)})\right] \\
- \frac{\lambda}{2}\,t(1-t)\left[Z^{(a)}W_2^2(\hat{\mu}_0^{(a)},\hat{\mu}_1^{(a)}) + Z^{(s)}W_2^2(\hat{\mu}_0^{(s)},\hat{\mu}_1^{(s)})\right].
\end{aligned} \tag{12}$$

The last term inside square brackets can be rewritten as

$$\int d^2(\gamma_0,\gamma_1)\Pi^{(a)}(d\gamma) + \int d^2(\gamma_0,\gamma_1)\Pi^{(s)}(d\gamma) = \int d^2(\gamma_0,\gamma_1)\Pi(d\gamma) = W_2^2(\mu_0,\mu_1).$$

Plugging this back into (12) and using Lemma 2.1, we conclude that

$$U_\nu(\mu_t) \leq (1-t)\,U_\nu(\mu_0) + t\,U_\nu(\mu_1) - \frac{\lambda}{2}\,t(1-t)\,W_2^2(\mu_0,\mu_1).$$

This finishes the proof of Theorem 1.7. □

Proof of Corollary 1.8. Let $U := r \to -r^{1-1/N}$. By the estimates derived in [4, Proposition E.17], U_ν is well-defined on $P_p(M)$. (This is made more explicit in [10, Theorem 17.8 and Example 17.9].)

Let $\mathcal{DC}_n$ be the displacement convex class of order n, that is the class of functions $U \in C^2(0,\infty) \cap C([0,+\infty))$ such that $U(0) = 0$ and $\delta^n U(\delta^{-n})$ is a convex function of δ. (See [10, Definition 17.1]). Obviously, $U \in \mathcal{DC}_n$. By [10, Proposition 17.7], there is a sequence $(U^{(\ell)})_{\ell\in\mathbb{N}}$ of Lipschitz functions, all belonging to $\mathcal{DC}_n$, such that $U^{(\ell)}$ converges monotonically to U as $\ell \to \infty$.

Since $U^{(\ell)}$ lies in $\mathcal{DC}_n$, it is by now classical (see [10, Theorem 17.15], which summarizes the works of many authors) that $U_\nu^{(\ell)}$ it is a.c.c.s-displacement convex. By Theorem 1.7, this functional is also displacement convex. Then it follows by an easy limiting argument that U_ν itself is displacement convex. □

References

[1] Cordero-Erausquin, D., McCann, R. J., and Schmuckenschläger, M. A Riemannian interpolation inequality à la Borell, Brascamp and Lieb. *Invent. Math. 146*, 2 (2001), 219–257.

[2] Fathi, A., and Figalli, A. Optimal transportation on non-compact manifolds. Preprint (2006).

[3] Figalli, A. Existence, uniqueness and regularity of optimal transport maps. To appear in *SIAM, Journal Math. Anal.* Available online via `http://cvgmt.sns.it/people/figalli/`.

[4] Lott, J., and Villani, C. Ricci curvature for metric-measure spaces via optimal transport. To appear in *Ann. of Math.* Available online via `http://www.umpa.ens-lyon.fr/~cvillani/`.

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

[6] Sturm, K.-T. On the geometry of metric measure spaces. I. *Acta Math. 196*, 1 (2006), 65–131.

[7] Sturm, K.-T. On the geometry of metric measure spaces. II. *Acta Math. 196*, 1 (2006), 133–177.

[8] Sturm, K.-T., and von Renesse, M.-K. Transport inequalities, gradient estimates, entropy and Ricci curvature. *Comm. Pure Appl. Math. 58*, 7 (2005), 923–940.

[9] Villani, C. *Topics in optimal transportation*, vol. 58 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2003.

[10] Villani, C. Optimal transport, old and new. Notes from the Saint-Flour 2005 Summer School. Available online at `www.umpa.ens-lyon.fr/~cvillani/`.