

This is a self-archived version of an original article. This version may differ from the original in pagination and typographic details.

Author(s): Schultz, Timo

Title: Existence of optimal transport maps in very strict $CD(K, \infty)$ -spaces

Year: 2018

Version: Accepted version (Final draft)

Copyright: © Springer-Verlag GmbH Germany, part of Springer Nature 2018.

Rights: In Copyright

Rights url: <http://rightsstatements.org/page/InC/1.0/?language=en>

Please cite the original version:

Schultz, T. (2018). Existence of optimal transport maps in very strict $CD(K, \infty)$ -spaces. *Calculus of Variations and Partial Differential Equations*, 57(5), Article 139. <https://doi.org/10.1007/s00526-018-1414-y>

EXISTENCE OF OPTIMAL TRANSPORT MAPS IN VERY STRICT $CD(K, \infty)$ -SPACES

TIMO SCHULTZ

ABSTRACT. We introduce a more restrictive version of the strict $CD(K, \infty)$ -condition, the so-called very strict $CD(K, \infty)$ -condition, and show the existence of optimal maps in very strict $CD(K, \infty)$ -spaces despite the possible lack of uniqueness of optimal plans.

1. INTRODUCTION

Consider a (complete and separable) metric space (X, d) . In the theory of optimal mass transportation probably the most studied problem is the one with the quadratic cost i.e. the minimisation problem

$$(1) \quad \inf \int_{X \times X} d^2(x, y) \, d\varpi(x, y),$$

where the infimum is taken over all transport plans ϖ between a given starting measure μ_0 and a final measure μ_1 , in other words over all Borel probability measures ϖ on $X \times X$ with marginals μ_0 and μ_1 . This problem is interpreted as an optimal mass transportation problem in the following way. The quantity $d^2(x, y)$ tells how much it costs to transport a unit mass from x to y , and for a given transport plan ϖ , mass from x is to be transported to y , if and only if (x, y) belongs to the support of ϖ . The total cost of sending all the mass μ_0 to μ_1 according to the plan ϖ is given by the integral

$$\int_{X \times X} d^2(x, y) \, d\varpi(x, y).$$

The above formulation is the so-called Kantorovich formulation of the optimal mass transportation problem, which is a relaxed version of the original Monge formulation of the problem, where instead of the infimum (1) one considers the infimum

$$(2) \quad \inf \int_X d^2(x, T(x)) \, d\mu_0(x),$$

where the infimum is taken over all Borel mappings $T: X \rightarrow X$ for which $T_{\#}\mu_0 = \mu_1$ i.e. all the Borel maps sending the mass μ_0 to μ_1 . One natural and interesting question is whether these two infima agree and when is the optimal plan in (1) given by an optimal

Date: May 17, 2018.

2000 Mathematics Subject Classification. Primary 53C23.

Key words and phrases. Ricci curvature, optimal mass transportation, metric measure spaces, existence of optimal maps, branching geodesics.

Author is partially supported by the Academy of Finland.

map in (2). More precisely, when is the optimal plan ϖ of the form $\varpi = (\text{id}, T)_{\#}\mu_0$ for some Borel mapping $T: X \rightarrow X$.

The existence of optimal map was proven in the Euclidean setting for absolutely continuous measures by Brenier [6]. Later this result was generalized to the Riemannian framework by McCann [18], and further to some cases of the sub-Riemannian setting by Ambrosio and Rigot [2], Agrachev and Lee [1], and by Figalli and Rifford [10]. In metric space setting Bertrand proved the existence of optimal map in the so-called Alexandrov spaces [5]. Under the assumption of non-branching of geodesics, the existence of optimal map was proven for metric measure spaces with Ricci curvature bounded from below i.e. for the spaces satisfying the so-called curvature dimension condition ($CD(K, N)$ -condition for short) by Gigli [11] (see Section 2.2 for the definition of $CD(K, \infty)$ -space), and with milder assumptions by Cavalletti and Huesmann [8].

The non-branching assumption plays a crucial role in both of those proofs leaving the existence of optimal maps in general $CD(K, \infty)$ -spaces open. On the other hand, if one considers spaces satisfying only the so-called measure contraction property, which is weaker type of Ricci curvature lower bound condition, the existence of optimal maps may fail as what follows from the example of Ketterer and Rajala in [15].

In this paper we go towards understanding the question in general $CD(K, \infty)$ -spaces by considering a possibly more restrictive version of $CD(K, \infty)$ -property without the non-branching assumption, namely we require the entropy to be convex not only along one optimal geodesic plan but instead along all plans that we get by restricting and weighting a given particular plan (see 2.1 for the definition of very strict $CD(K, \infty)$ -space). We prove the existence of optimal maps in very strict $CD(K, \infty)$ -spaces between measures μ_0 and μ_1 that are absolutely continuous with respect to the reference measure \mathbf{m} . We actually prove the following stronger statement saying that there exists an optimal dynamical transport plan that is induced by a map.

Theorem 1.1. *Let (X, d, \mathbf{m}) be a very strict $CD(K, \infty)$ -space and let $\mu_0, \mu_1 \in \mathcal{P}_2(X)$ be absolutely continuous with respect to \mathbf{m} . Then there exists a measure $\pi \in \text{OptGeo}(\mu_0, \mu_1)$ that is induced by a map i.e. there exists a Borel mapping $T: X \rightarrow \text{Geo}(X)$ so that $\pi = T_{\#}\mu_0$.*

In [19] Rajala and Sturm (see also [13]) were able to remove the a priori non-branching assumption by considering a more restrictive version of Ricci curvature lower bounds, namely the strong $CD(K, \infty)$ -property. They introduced the definition of essential non-branching spaces and proved that strong $CD(K, \infty)$ -spaces are essentially non-branching, from which they deduced by using the idea of the proof of Gigli [11] that every optimal plan is given by a map. The result of Rajala and Sturm applies also in the measured Gromov Hausdorff -stable setting of metric measure spaces with Riemannian Ricci curvature bounded from below (the so-called $RCD(K, \infty)$ -spaces, see [4, 3]). While the $CD(K, \infty)$ -condition is stable [20, 12], and the strong $CD(K, \infty)$ is not, it remains open whether very strict $CD(K, \infty)$ -, or strict $CD(K, \infty)$ -condition is stable.

Question 1. *Let $(X_i, d_i, \mathbf{m}_i)_{i \in \mathbb{N}}$ be a sequence of (very) strict $CD(K, \infty)$ -spaces that converge to a metric measure space (X, d, \mathbf{m}) in suitable sense (for example in pointed measured Gromov sense). Is (X, d, \mathbf{m}) a (very) strict $CD(K, \infty)$ -space?*

Another related open question is the relation of $CD(K, \infty)$ -, strict $CD(K, \infty)$ -, and very strict $CD(K, \infty)$ -conditions:

Question 2. *Does the strict $CD(K, \infty)$ -condition imply very strict $CD(K, \infty)$ -condition? Does $CD(K, \infty)$ imply (very) strict $CD(K, \infty)$?*

With the new notion of essential non-branching introduced by Rajala and Sturm, Cavalletti and Mondino further proved the existence of optimal transport maps for essentially non-branching spaces with the measure contraction property [9]. Continuing from the work of Cavalletti and Huesmann [8] and the work of Cavalletti and Mondino [9], Kell proved that in a metric space endowed with qualitatively non-degenerate measure, and therefore especially in spaces satisfying the measure contraction property, the condition of being essentially non-branching is equivalent with having the existence and uniqueness of the optimal transport maps [14].

In the previous results the existence of optimal map is shown by proving that every optimal plan is given by a map and hence also the uniqueness of the plan is guaranteed. In very strict $CD(K, \infty)$ -spaces optimal plans may fail to be unique – which can be observed by looking for example at the space \mathbb{R}^n equipped with supremum norm – and therefore this strategy cannot work in our setting. Instead, we should consider one special plan that is given by the definition of very strict $CD(K, \infty)$ -space. Notice that even though very strict $CD(K, \infty)$ -spaces may fail to be non-branching (even in the sense of essential non-branchingness), still this specific optimal plan does not see any branching geodesics.

Our proof follows the ideas of Rajala and Sturm in [19] and of Gigli in [11]. Instead of proving the existence via the non-branchingness of the optimal plan, we do the proof directly, since it is not clear how to implement the idea of the mixing procedure of [19] in the very strict $CD(K, \infty)$ -setting.

2. PRELIMINARIES

Throughout this paper (X, d, \mathbf{m}) is assumed to be a complete and separable metric space endowed with a locally finite Borel measure \mathbf{m} . Since we are considering only transportations between absolutely continuous measures, all the Wasserstein geodesics of our concern live in the set of absolutely continuous measures due to the K -convexity of the entropy. Thus we may restrict to the case where $X = \text{spt } \mathbf{m}$.

By a geodesic we mean a constant speed curve $\gamma: [0, 1] \rightarrow X$ that is length minimizing i.e. $l(\gamma) = d(\gamma_0, \gamma_1)$, where l denotes the length of γ . We denote by $\text{Geo}(X)$ the set of all geodesics of the space X endowed with the supremum metric.

2.1. Optimal mass transportation and Wasserstein geodesics. We denote by $\mathcal{P}(X)$ the space of all Borel probability measures on X , and by $\mathcal{P}_2(X)$ the set of all $\mu \in \mathcal{P}(X)$

with finite second moment i.e. those μ for which we have

$$\int d^2(x, x_0) d\mu < \infty$$

for some – and thus for all – $x_0 \in X$.

We define the Wasserstein 2-distance W_2 in $\mathcal{P}_2(X)$ as

$$W_2(\mu, \nu) := \left(\inf_{\sigma \in \mathcal{A}(\mu, \nu)} \int_{X \times X} d^2(x, y) d\sigma \right)^{\frac{1}{2}},$$

where $\mathcal{A}(\mu, \nu) := \{\sigma \in \mathcal{P}(X \times X) : P_{\#}^1 \sigma = \mu, P_{\#}^2 \sigma = \nu\}$ is the set of all admissible plans. The square of the Wasserstein distance is nothing else but the total cost in the mass transportation problem with quadratic cost between the masses μ and ν . We denote the set of admissible plans realising the above infimum as $\text{Opt}(\mu, \nu)$.

Even though we do not assume the space (X, d) to be geodesic, at the end of Section 2 we point out that from the definition of very strict $CD(K, \infty)$ -space we actually get that the space X is a length space – keeping in mind that $X = \text{spt } \mathbf{m}$. Since X is a complete and separable metric space with length structure, we also have that the Wasserstein space $(\mathcal{P}_2(X), W_2)$ is a complete and separable length space (see [21] and [16]). Furthermore, a curve $t \mapsto \mu_t$ in $\mathcal{P}_2(X)$ is a geodesic if and only if there exists $\pi \in \mathcal{P}(\text{Geo}(X))$ such that $(e_t)_{\#} \pi = \mu_t$ and $(e_0, e_1)_{\#} \pi \in \text{Opt}(\mu_0, \mu_1)$. Here $e_t: \text{Geo}(X) \rightarrow X$ is the evaluation map $\gamma \mapsto \gamma_t := \gamma(t)$. The set of all $\pi \in \mathcal{P}(\text{Geo}(X))$ for which $(e_0, e_1)_{\#} \pi \in \text{Opt}(\mu_0, \mu_1)$ is denoted by $\text{OptGeo}(\mu_0, \mu_1)$.

2.2. Ricci curvature bounded from below. The notion of synthetic Ricci curvature lower bounds for metric measure spaces were first introduced by Sturm [20] and independently by Lott and Villani [17]. The definitions are based on the connection of Ricci curvature and optimal mass transportation; namely, the convexity properties of suitable entropy functionals along Wasserstein geodesic. For the definition of Ricci curvature lower bounds let us first introduce the entropy functional $\text{Ent}_{\infty}: \mathcal{P}(X) \rightarrow [-\infty, \infty]$ that is defined as

$$\text{Ent}_{\infty}(\mu) = \begin{cases} \int \rho \log \rho d\mathbf{m} & , \text{ if } \mu \ll \mathbf{m} \text{ and } (\rho \log \rho)_+ \in L^1(\mathbf{m}), \\ \infty & \text{ otherwise,} \end{cases}$$

where ρ is the density of μ with respect to \mathbf{m} i.e. $\mu = \rho \mathbf{m}$, and $(\rho \log \rho)_+ = \max\{\rho \log \rho, 0\}$.

A metric measure space (X, d, \mathbf{m}) is said to have Ricci curvature bounded from below by $K \in \mathbb{R}$, if for every $\mu_0, \mu_1 \in \mathcal{P}_2(X)$ absolutely continuous with respect to \mathbf{m} , there exists $\pi \in \text{OptGeo}(\mu_0, \mu_1)$ such that the entropy Ent_{∞} is K -convex along π , that is the inequality

$$(3) \quad \text{Ent}_{\infty}(\mu_t) \leq (1-t)\text{Ent}_{\infty}(\mu_0) + t\text{Ent}_{\infty}(\mu_1) - \frac{K}{2}t(1-t)W_2^2(\mu_0, \mu_1)$$

holds for all $t \in [0, 1]$, where $\mu_t := (e_t)_{\#} \pi$. Such a space is called a $CD(K, \infty)$ -space. If the K -convexity holds along every $f\pi$, where $f: \text{Geo}(X) \rightarrow \mathbb{R}$ is any non-negative Borel function for which $\int f d\pi = 1$, then the space is called a strict $CD(K, \infty)$ -space (see [4]). In this paper a more restrictive version of strict $CD(K, \infty)$ -condition is used, namely the

convexity of the entropy is not only required between points 0, t and 1, but also between any points $t_1 < t_2 < t_3$. To emphasise the difference, we call these spaces *very strict* $CD(K, \infty)$ -spaces.

Definition 2.1. *A metric measure space (X, d, \mathbf{m}) is called a very strict $CD(K, \infty)$ -space, if for every $\mu_0, \mu_1 \in \mathcal{P}_2(X)$ absolutely continuous with respect to the reference measure \mathbf{m} there exists $\pi \in \text{OptGeo}(\mu_0, \mu_1)$ so that the entropy Ent_∞ is K -convex along $(\text{restr}_{t_0}^{t_1})_\#(f\pi)$ for all $t_0, t_1 \in [0, 1]$, $t_0 < t_1$, and for all non-negative Borel functions $f: \text{Geo}(X) \rightarrow \mathbb{R}$ with $\int f \, d\pi = 1$.*

Here $\text{restr}_{t_0}^{t_1}: \text{Geo}(X) \rightarrow \text{Geo}(X)$ is the restriction map defined as

$$\text{restr}_{t_0}^{t_1}(\gamma)(t) := \gamma(t(t_1 - t_0) + t_0).$$

It is worth of noticing that due to the Radon-Nikodym theorem, in the above definition one could equivalently require the convexity to hold along $(\text{restr}_{t_0}^{t_1})_\#\tilde{\pi}$ for all $\tilde{\pi} \in \mathcal{P}(X)$ that are absolutely continuous with respect to π .

As mentioned before, the very strict $CD(K, \infty)$ -condition implies that the space X is a length space: let $x, y \in X$ and define for $\varepsilon > 0$ the measures $\mu_0 := \mathbf{m}|_{B(x, \varepsilon)}$ and $\mathbf{m}|_{B(y, \varepsilon)}$. Let $\pi \in \text{OptGeo}(\mu_0, \mu_1)$ which exists by the definition. Now any $\gamma \in \text{spt}\pi$ is a geodesic from $B(x, \varepsilon)$ to $B(y, \varepsilon)$ and thus the point $\gamma_{\frac{1}{2}}$ is an ε -midpoint of x and y . Thus by completeness we have that X is a length space [7].

3. THE MAIN THEOREM

In the paper [19] Rajala and Sturm prove the existence of the optimal map by first proving that every optimal plan $\pi \in \text{OptGeo}(X)$ is essentially non-branching and then as a corollary of this they prove that every such π is actually given by a map. While the proof for the essential non-branching of the $\pi \in \text{OptGeo}(X)$ given by the definition of very strict $CD(K, \infty)$ -space can be carried through with relatively small changes, the proof of the corollary is more problematic.

In their proof of the existence of optimal map they first divide the original measure π into two measures π^1 and π^2 that intersect at time t , and then construct a new measure π^{mix} by mixing these measures π^1 and π^2 essentially in the way that at time t you may change from a geodesic in the support of one of the measures to a geodesic in the support of the other. By doing this they end up with a new plan that is still optimal, but due to this mixing the plan is not essentially non-branching anymore. The problem in applying this strategy to our case is that the mixing procedure should be done in such a way that in the end the constructed measure π^{mix} is absolutely continuous with respect to the original measure π , which is something that one should not expect from this kind of mixing. To overcome this obstacle, we prove directly the existence of a map by still using the idea from the proof of the essential non-branching in [19].

In the proof of the existence of optimal map we will use the following two lemmas.

Lemma 3.1 (cf. [11, Lemma 3.2]). *Let $\pi \in \text{OptGeo}(\mu_0, \mu_1)$ be the plan given by the definition of very strict $CD(K, \infty)$. Then for all $\tilde{\pi} \ll \pi$ with $\mathbf{m}(\{\tilde{\rho}_0 > 0\}) < \infty$ and*

$\text{Ent}_\infty(\tilde{\mu}_0), \text{Ent}_\infty(\tilde{\mu}_1) \in \mathbb{R}$ it holds that

$$\mathbf{m}(\{\tilde{\rho}_0 > 0\}) \leq \liminf_{t \rightarrow 0} \mathbf{m}(\{\tilde{\rho}_t > 0\}),$$

where $\tilde{\rho}_t$ is the density of $(e_t)_\# \tilde{\pi}$ with respect to \mathbf{m} .

Lemma 3.2. *Let (X, d) be a separable metric space. Then for any $\sigma \in \mathcal{P}(X \times X)$ for which $\sigma(\{(x, x) : x \in X\}) = 0$ there exists $E \subset X$ so that $\sigma(E \times (X \setminus E)) > 1/5$.*

The proof of Lemma 3.1 is the same as the proof of [11, Lemma 3.2] and the proof of Lemma 3.2 can be found in [19]. We will also use the following simple lemma.

Lemma 3.3. *Let $\mu_0, \mu_1 \in \mathcal{P}_2(X)$ be absolutely continuous with respect to \mathbf{m} with densities ρ_0 and ρ_1 such that $\text{Ent}_\infty(\mu_0), \text{Ent}_\infty(\mu_1) \in \mathbb{R}$ and $\rho_0, \rho_1 < C$, and let $\pi \in \text{OptGeo}(\mu_0, \mu_1)$ be an optimal plan concentrated on a set of geodesics with length bounded by some constant L such that the convexity inequality (3) holds for all restrictions of π . Then there exists a constant $M < \infty$ such that for all $t \in [0, 1]$ we have $\rho_t \leq M$ μ_t -almost everywhere, where $\mu_t := (e_t)_\# \pi$ with density ρ_t .*

Proof. We argue by contradiction. Assume that for all $M > 0$ there exists $t_M \in (0, 1)$ so that $N_M := \mu_{t_M}(\{\rho_{t_M} > M\}) > 0$. Define

$$\hat{\pi} := \pi|_{e_{t_M}^{-1}(\{\rho_{t_M} > M\})}$$

and denote $\hat{\mu}_t := (e_t)_\# \hat{\pi}$ and the density of $\hat{\mu}_t$ by $\hat{\rho}_t$. Then we have that

$$\begin{aligned} & \int \frac{\hat{\rho}_{t_M}}{N_M} \log \frac{\hat{\rho}_{t_M}}{N_M} \, d\mathbf{m} - (1 - t_M) \int \frac{\hat{\rho}_0}{N_M} \log \frac{\hat{\rho}_0}{N_M} \, d\mathbf{m} - t_M \int \frac{\hat{\rho}_1}{N_M} \log \frac{\hat{\rho}_1}{N_M} \, d\mathbf{m} \\ & \geq \log \frac{M}{N_M} - (1 - t_M) \log \frac{C}{N_M} - t_M \log \frac{C}{N_M} = \log M - \log C \rightarrow \infty, \end{aligned}$$

when

$$M \rightarrow \infty.$$

On the other hand by the K-convexity we have for all M that

$$\begin{aligned} & \int \frac{\hat{\rho}_{t_M}}{N_M} \log \frac{\hat{\rho}_{t_M}}{N_M} \, d\mathbf{m} - (1 - t_M) \int \frac{\hat{\rho}_0}{N_M} \log \frac{\hat{\rho}_0}{N_M} \, d\mathbf{m} - t_M \int \frac{\hat{\rho}_1}{N_M} \log \frac{\hat{\rho}_1}{N_M} \, d\mathbf{m} \\ & \leq -\frac{K}{2} t_M (1 - t_M) W_2^2\left(\frac{\hat{\mu}_0}{N_M}, \frac{\hat{\mu}_1}{N_M}\right) \leq \frac{|K|}{2} L^2 < \infty. \end{aligned}$$

which is a contradiction. Hence there exists M so that $\rho_t \leq M$ for all $t \in [0, 1]$. \square

Proof of Theorem 1.1. Let $\mu_0, \mu_1 \in \mathcal{P}_2(X)$ be measures that are absolutely continuous with respect to \mathbf{m} , and let ρ_0 and ρ_1 be densities of μ_0 and μ_1 with respect to \mathbf{m} . We will prove that the measure $\pi \in \text{OptGeo}(\mu_0, \mu_1)$ given by the definition of very strict $CD(K, \infty)$ -space is induced by a map. We will argue by contradiction. Assume that π is not induced by a map. As in [19], we may assume that $\rho_0, \rho_1 < C < \infty$ and that the space X is bounded.

We may also assume that $\mathbf{m}(X)$ is finite. The argument goes as follows. Since X is complete and separable, so is $\text{Geo}(X)$ (as a closed subset of a complete and separable

space of continuous curves). Therefore π is a Radon measure. In particular there exists an increasing sequence of compact subsets $K_i \subset \text{Geo}(X)$ for which $\pi(K_i) \geq 1 - \frac{1}{i}$. Since π is not induced by a map, there exists $n \in \mathbb{N}$ so that $\pi|_{K_n}$ is not induced by a map. Furthermore $\text{spt}(e_t)_\# \pi|_{K_n} \subset e_t(K_n) \subset \bigcup_{t \in [0,1]} e_t(K_t)$. Thus the relevant part of the space X lives inside the compact set $\bigcup_{t \in [0,1]} e_t(K_t)$, which by local finiteness of \mathfrak{m} has finite measure. In particular, by using Jensen's inequality we may also assume that $\text{Ent}_\infty(\mu_0), \text{Ent}_\infty(\mu_1) \in \mathbb{R}$.

As in the proof of the essential non-branching of strong $CD(K, \infty)$ -spaces in [19], we want to make the square of the Wasserstein distance $W_2^2(\mu_0, \mu_1)$ arbitrary small, so that we can basically consider only the convexity of the entropy and forget the K dependent error term in (3). This is done by the following lemma

Lemma 3.4. *If $\pi \in \text{OptGeo}(\mu_0, \mu_1)$ is not induced by a map, then for every $k \in \mathbb{N}$ there exists an interval $[i/k, (i+1)/k]$ so that $(\text{restr}_{i/k}^{(i+1)/k})_\# \pi$ is not induced by a map i.e. $(\text{restr}_{i/k}^{(i+1)/k})_\# \pi \neq T_\#((e_{i/k})_\# \pi)$.*

We will postpone the proof of Lemma 3.4 to the end of the paper. Using the above lemma we may restrict the plan π so that we have the inequality

$$L^2 \leq \frac{\log 2}{6|K| + 1},$$

where

$$L := \text{ess sup}_{\gamma \in \text{Geo}(X)} l(\gamma)$$

and the essential supremum is taken with respect to the (restricted) measure π .

Step 1: As in the proof of essential non-branching of strong $CD(K, \infty)$ -spaces, we lift the measure π to a measure in $\mathcal{P}(\text{Geo}(X)^2)$. Let $\{\pi_x\}$ be a disintegration of the measure π with respect to the evaluation map $e_0: \text{Geo}(X) \rightarrow X$. We define $\sigma \in \mathcal{P}(\text{Geo}(X)^2)$ by defining the integral of any Borel function $f: \text{Geo}(X)^2 \rightarrow [0, \infty]$ with respect to σ as

$$\int_{\text{Geo}(X)^2} f d\sigma := \int_X \int_{e_0^{-1}(x) \times e_0^{-1}(x)} f d(\pi_x \times \pi_x) d\mu_0.$$

Note that σ is well defined since the map $x \mapsto \pi_x \times \pi_x(A)$ is Souslin measurable for every Borel set $A \subset \text{Geo}(X)^2$. This can be seen by applying Dynkin's $\pi - \lambda$ theorem to a π -system $\mathcal{B}(\text{Geo}(X)) \times \mathcal{B}(\text{Geo}(X))$ and a λ -system $\{A \in \mathcal{B}(\text{Geo}(X)^2) : x \mapsto \pi_x \times \pi_x(A) \text{ is Souslin measurable}\}$.

Since π is not induced by a map, there exists $H \subset X$ with positive μ_0 -measure so that for any $x \in H$ the measure π_x is not a dirac mass. Therefore, there exists $F \subset \text{Geo}(X)^2$ for which $\sigma(F) > 0$, and it holds that for any $(\gamma^1, \gamma^2) \in F$ we have that $\gamma_0^1 = \gamma_0^2$ and $\gamma^1 \neq \gamma^2$. By Lemma 3.2 there exists $E \subset \text{Geo}(X)$ so that $\sigma((E \times (X \setminus E)) \cap F) > \sigma(F)/5$. Let $\eta > 0$ be such that

$$\sigma(A) > \frac{1}{10} \sigma(F)$$

for the set $A := \{(\gamma^1, \gamma^2) \in (E \times (X \setminus E)) \cap F : d(\gamma^1, \gamma^2) > \eta\}$. Let $m \in \mathbb{N}$ be large enough so that $\frac{1}{m} < \frac{\eta}{4L}$ and divide the interval $[0, 1]$ into m intervals $\{I_j\}_{j=1}^m$ of equal length. Then for every $(\gamma^1, \gamma^2) \in A$ there exist $j \in \{1, \dots, m\}$ and $t \in I_j$ so that $d(\gamma_t^1, \gamma_t^2) > \eta$ and thus for any $s \in I_j$, since $|t - s| \leq \frac{1}{m}$, we have that

$$d(\gamma_s^1, \gamma_s^2) \geq d(\gamma_t^1, \gamma_t^2) - d(\gamma_t^1, \gamma_s^1) - d(\gamma_t^2, \gamma_s^2) \geq \eta - 2\frac{\eta}{4L}L \geq \frac{\eta}{2}.$$

Hence there exists i so that $\sigma(A_i) > 0$ for $A_i := \{(\gamma^1, \gamma^2) \in A : d(\gamma_t^1, \gamma_t^2) > \frac{\eta}{2} \forall t \in I_i\}$. Take a countable partition $P := \{Q_j\}$ of X with $d(Q_j) < \frac{\eta}{4}$. Then for some $Q \in P$ we have that $\sigma((e_S^{-1}(Q) \times \text{Geo}(X)) \cap A_i) > 0$, where $S \in I_i$ is the mid-point of the interval I_i . Now for any $(\gamma^1, \gamma^2) \in (e_S^{-1}(Q) \times \text{Geo}(X)) \cap A_i$ it holds that $\gamma_S^1 \in Q$ and $\gamma_S^2 \notin Q$. Define $\tilde{\sigma}$ as the restriction $\tilde{\sigma} := \sigma|_{(e_S^{-1}(Q) \times \text{Geo}(X)) \cap A_i}$. Furthermore define $\pi^1 := P_{\#}^1 \sigma$ and $\pi^2 := P_{\#}^2 \sigma$. Then $\pi^1, \pi^2 \ll \pi$, $\mu_0^1 := (e_0)_{\#} \pi^1 = (e_0)_{\#} \pi^2 =: \mu_0^2$ and $\mu_S^1 \perp \mu_S^2$. By restricting the measures π^1 and π^2 we may assume that $\rho_j^i < C < \infty$ for some $C > 0$ and for all $i \in \{1, 2\}$ and $j \in \{0, 1\}$, where ρ_j^i is the density of $(e_j)_{\#} \mu^i$ with respect to \mathbf{m} .

Step 2: As in the proof of [11, Theorem 3.3] we will find a time $T \in (0, S)$ so that the intersection of the sets $\{\rho_T^1 > 0\}$ and $\{\rho_T^2 > 0\}$ has positive measure. We repeat the argument here. Since \mathbf{m} is locally finite and $\text{spt } \mu_0^1$ is compact, there exists a neighbourhood of $\text{spt } \mu_0^1$ with finite \mathbf{m} -measure. Furthermore there exists an ε -neighbourhood D of $\text{spt } \mu_0^1$ for which $\mathbf{m}(D) \leq \frac{3}{2}\mathbf{m}(\text{spt } \mu_0^1)$. By Lemma 3.1 and by the fact that $L < \infty$ there exists $T \in (0, S)$ so that $\mathbf{m}(\{\rho_T^i > 0\}) > \frac{3}{4}\mathbf{m}(\{\rho_0^i > 0\})$ and $\{\rho_T^i > 0\} \subset D$. Hence $\mathbf{m}(\{\rho_T^1 > 0\} \cap \{\rho_T^2 > 0\}) > 0$. Define now

$$\hat{\pi}^i := \frac{1}{\int \min\{\rho_T^1, \rho_T^2\} d\mathbf{m}} \left(\frac{\min\{\rho_T^1, \rho_T^2\}}{\rho_T^i} \circ e_T \right) \pi^i$$

for $i \in \{1, 2\}$. Then we have that $\hat{\pi}^1, \hat{\pi}^2 \ll \pi$, $\hat{\mu}_T^1 = \hat{\mu}_T^2$, $\hat{\mu}_S^1 \perp \hat{\mu}_S^2$ and $\hat{\rho}_j^1, \hat{\rho}_j^2 < C < \infty$ for some $C > 0$ and for $j \in \{0, 1\}$. By Lemma 3.3 there exists $C > 0$ so that $\hat{\rho}_t^1, \hat{\rho}_t^2 < C < \infty$ for every $t \in [0, 1]$.

Step 3: Let $T = t_0 < t_1 < \dots < t_k = S$ be a partition of the interval $[T, S]$ into subintervals of equal length and define $f(i) := \int \min\{\hat{\rho}_{t_i}^1, \hat{\rho}_{t_i}^2\} d\mathbf{m}$. Then $f(0) = 1$ and $f(k) = 0$. Therefore there exists $i \in \{0, \dots, k\}$ so that $f(i) - f(i+1) \geq \frac{1}{k}$. Define now $\tilde{\pi}^1 := \hat{\pi}^1|_{e_{t_{i+1}}^{-1}(\{\hat{\rho}_{t_{i+1}}^1 > \hat{\rho}_{t_{i+1}}^2\})}$ and $\tilde{\pi}^2 := \hat{\pi}^2|_{e_{t_{i+1}}^{-1}(\{\hat{\rho}_{t_{i+1}}^1 \leq \hat{\rho}_{t_{i+1}}^2\})}$, and further for $j \in \{1, 2\}$

$$\bar{\pi}^j := \left(\frac{\min\{\tilde{\rho}_{t_i}^1, \tilde{\rho}_{t_i}^2\}}{\tilde{\rho}_{t_i}^j} \circ e_{t_i} \right) \tilde{\pi}^j.$$

Then $\bar{\mu}_{t_{i+1}}^1(\{\hat{\rho}_{t_{i+1}}^1 > \hat{\rho}_{t_{i+1}}^2\}) = \bar{\mu}_{t_{i+1}}^1(X)$ and $\bar{\mu}_{t_{i+1}}^2(\{\hat{\rho}_{t_{i+1}}^1 \leq \hat{\rho}_{t_{i+1}}^2\}) = \bar{\mu}_{t_{i+1}}^2(X)$. Thus we have that $\bar{\mu}_{t_{i+1}}^1 \perp \bar{\mu}_{t_{i+1}}^2$. By definition we also have that $\bar{\mu}_{t_i}^1 = \bar{\mu}_{t_i}^2$. Let us prove that

$$(4) \quad \bar{\mu}_{t_i}^1(X) \geq \frac{1}{k}.$$

By definition we have that $\bar{\rho}_{t_i}^j \leq \tilde{\rho}_{t_i}^j \leq \hat{\rho}_{t_i}^j$ for $j \in \{1, 2\}$. Also we have that

$$\begin{aligned} \tilde{\mu}_{t_i}^1(X) + \tilde{\mu}_{t_i}^2(X) &= \tilde{\mu}_{t_{i+1}}^1(X) + \tilde{\mu}_{t_{i+1}}^2(X) \\ &= \int \hat{\rho}_{t_{i+1}}^1 + \hat{\rho}_{t_{i+1}}^2 - \min\{\hat{\rho}_{t_{i+1}}^1, \hat{\rho}_{t_{i+1}}^2\} \, \mathrm{d}m \\ &= 2 - f(i+1). \end{aligned}$$

Therefore we get that

$$\begin{aligned} \bar{\mu}_{t_i}^1(X) &= \int \min\{\bar{\rho}_{t_i}^1, \bar{\rho}_{t_i}^2\} \, \mathrm{d}m = \tilde{\mu}_{t_i}^1(X) + \tilde{\mu}_{t_i}^2(X) - \int \max\{\bar{\rho}_{t_i}^1, \bar{\rho}_{t_i}^2\} \, \mathrm{d}m \\ &\geq 2 - f(i+1) - \int \max\{\hat{\rho}_{t_i}^1, \hat{\rho}_{t_i}^2\} = f(i) - f(i+1) \geq \frac{1}{k}. \end{aligned}$$

Final step: Now we are ready to arrive to a contradiction with the convexity of the entropy. As in [19] we consider three measures $\bar{\pi}^1/M$, $\bar{\pi}^2/M$ and $(\bar{\pi}^1 + \bar{\pi}^2)/(2M)$ along which the entropy is K -convex. Here $M := \bar{\pi}^i(\text{Geo}(X))$. For these measures we have that $\bar{\rho}_t^i < C$ for all $t \in [0, 1]$ and $i \in \{1, 2\}$, $M > 1/k$, $\bar{\mu}_{t_i}^1 = \bar{\mu}_{t_i}^2$ and $\bar{\mu}_{t_{i+1}}^1 \perp \bar{\mu}_{t_{i+1}}^2$. From these facts we get that

$$\begin{aligned} &\int \frac{\bar{\rho}_{t_i}^1}{M} \log \frac{\bar{\rho}_{t_i}^1}{M} \, \mathrm{d}m \\ &\leq \frac{1}{kt_{i+1}} \int \frac{\bar{\rho}_0^1 + \bar{\rho}_0^2}{2M} \log \frac{\bar{\rho}_0^1 + \bar{\rho}_0^2}{2M} \, \mathrm{d}m + \frac{t_i}{t_{i+1}} \int \frac{\bar{\rho}_{t_{i+1}}^1 + \bar{\rho}_{t_{i+1}}^2}{2M} \log \frac{\bar{\rho}_{t_{i+1}}^1 + \bar{\rho}_{t_{i+1}}^2}{2M} \, \mathrm{d}m \\ &\quad + \frac{|K|}{2} \frac{t_i}{t_{i+1}} W_2^2(\bar{\mu}_0, \bar{\mu}_1) \\ &\leq \frac{\log \frac{C}{M}}{kt_{i+1}} - \frac{t_i}{t_{i+1}} \log 2 + \frac{t_i}{2t_{i+1}} \left(\int \frac{\bar{\rho}_{t_{i+1}}^1}{M} \log \frac{\bar{\rho}_{t_{i+1}}^1}{M} \, \mathrm{d}m + \int \frac{\bar{\rho}_{t_{i+1}}^2}{M} \log \frac{\bar{\rho}_{t_{i+1}}^2}{M} \, \mathrm{d}m \right) \\ &\quad + \frac{|K|}{2} \frac{t_i}{t_{i+1}} L^2 \\ &\leq \frac{\log \frac{C}{M}}{kt_{i+1}} - \frac{t_i}{t_{i+1}} \log 2 + \frac{t_i}{2t_{i+1}} \left[\frac{2(1-t_{i+1})}{1-t_i} \int \frac{\bar{\rho}_{t_i}^1}{M} \log \frac{\bar{\rho}_{t_i}^1}{M} \, \mathrm{d}m \right. \\ &\quad \left. + \frac{1}{k(1-t_i)} \left(\int \frac{\bar{\rho}_1^1}{M} \log \frac{\bar{\rho}_1^1}{M} \, \mathrm{d}m + \int \frac{\bar{\rho}_1^2}{M} \log \frac{\bar{\rho}_1^2}{M} \, \mathrm{d}m \right) \right] + |K| \frac{t_i}{t_{i+1}} L^2 \\ &\leq \left(\frac{1}{kt_{i+1}} + \frac{t_i}{t_{i+1}k(1-t_{i+1})} \right) \log \frac{C}{M} - \frac{5}{6} \frac{t_i}{t_{i+1}} \log 2 + \frac{t_i(1-t_{i+1})}{t_{i+1}(1-t_i)} \int \frac{\bar{\rho}_{t_i}^1}{M} \log \frac{\bar{\rho}_{t_i}^1}{M} \, \mathrm{d}m \end{aligned}$$

from which we obtain

$$\frac{1}{kt_{i+1}(1-t_i)} \int \frac{\bar{\rho}_{t_i}^1}{M} \log \frac{\bar{\rho}_{t_i}^1}{M} \, \mathrm{d}m \leq \frac{1}{kt_{i+1}(1-t_i)} \log \frac{C}{M} - \frac{5}{6} \frac{t_i}{t_{i+1}} \log 2$$

and further

$$\int \frac{\bar{\rho}_{t_i}^1}{M} \log \frac{\bar{\rho}_{t_i}^1}{M} d\mathbf{m} \leq \log \frac{C}{M} - \frac{5k}{6}(1-t_i)t_i \log 2.$$

Choosing k large enough so that $T, (1-S)^2 \geq \frac{1}{k}$, and using the above inequality together with the convexity of the entropy along $\bar{\pi}^1$ we get that

$$\begin{aligned} & \int \frac{\bar{\rho}_{t_{i+1}}^1}{M} \log \frac{\bar{\rho}_{t_{i+1}}^1}{M} d\mathbf{m} \leq \frac{t_{i+1}-t_i}{1-t_i} \log \frac{C}{M} + \frac{1-t_{i+1}}{1-t_i} \int \frac{\rho_{t_i}^1}{M} \log \frac{\rho_{t_i}^1}{M} \\ & \quad + \frac{|K|}{2} \frac{(1-t_{i+1})}{k(1-t_i)^2} L^2 \\ & \leq \left(\frac{t_{i+1}-t_i}{1-t_i} + \frac{1-t_{i+1}}{1-t_i} \right) \log \frac{C}{M} - \frac{5k}{6} t_i(1-t_{i+1}) \log 2 \\ & \quad + \frac{|K|}{2} k t_i(1-t_{i+1}) L^2 \\ & \leq \log \frac{C}{M} - \frac{k}{3} t_i(1-t_{i+1}) \log 2 \end{aligned}$$

from which by (4) and Jensen's inequality we obtain

$$\begin{aligned} \log \frac{1}{k\mathbf{m}(X)} - \log M & \leq \int \frac{\bar{\rho}_{t_{i+1}}^1}{M} \log \bar{\rho}_{t_{i+1}}^1 d\mathbf{m} - \log M = \int \frac{\bar{\rho}_{t_{i+1}}^1}{M} \log \frac{\bar{\rho}_{t_{i+1}}^1}{M} d\mathbf{m} \\ & \leq \log \frac{C}{M} - \frac{k}{3} t_i(1-t_{i+1}) \log 2. \end{aligned}$$

Hence we arrive to a contradiction

$$\frac{3}{k} \left(\log \frac{1}{k\mathbf{m}(X)} - \log C \right) \leq -t_i(1-t_{i+1}) \log 2 \leq -T(1-S) \log 2$$

since the left hand side goes to zero when k goes to infinity while the right hand side is negative and bounded away from zero. \square

Proof of Lemma 3.4. By an induction argument it suffices to prove that if $\pi \in \text{OptGeo}(\mu_0, \mu_1)$ is not induced by a map, then for each $t \in (0, 1)$ we have that $(\text{restr}_0^t)_\# \pi$ or $(\text{restr}_t^1)_\# \pi$ is not induced by a map.

Suppose that there is $t \in (0, 1)$ so that both $\pi^1 := (\text{restr}_0^t)_\# \pi$ and $\pi^2 := (\text{restr}_t^1)_\# \pi$ are induced by maps T^1 and T^2 respectively. Denote $\{\pi_x\}, \{\pi_x^1\}, \{\pi_x^2\}$ the disintegrations of π, π^1 and π^2 with respect to e_0 . It suffices to prove that $\text{spt } \pi_x$ is a singleton for μ_0 -a.e. $x \in X$. We will do this by proving that actually

$$\text{spt } \pi_x = \text{spt } \pi_x^1 * \text{spt } \pi_{\hat{T}(x)}^2,$$

where $\hat{T} := e_1 \circ T^1$ is the optimal map from μ_0 to μ_t , and $*$ is the concatenation of paths.

We begin by observing that for μ_0 -a.e. x we have, by the definition of disintegration, that $\pi_x^1 = (\text{restr}_0^t)_\# \pi_x$. In particular we have that $\text{spt } \pi_x^1 = \text{restr}_0^t \text{spt } \pi_x$. Thus,

since both $\text{spt } \pi_x^1$ and $\text{spt } \pi_{\hat{T}(x)}^2$ are singletons for μ_0 -a.e. $x \in X$, it suffices to prove that $\pi_x((\text{restr}_t^1)^{-1}\text{spt } \pi_{\hat{T}(x)}^2) = 1$ for μ_0 -a.e. $x \in X$. Suppose this is not the case. Then the set

$$E := \{x : \pi_x((\text{restr}_t^1)^{-1}\text{spt } \pi_{\hat{T}(x)}^2) < 1\}$$

has positive μ_0 -measure. Consider the set $F := \cup_{x \in E} \text{spt } \pi_{\hat{T}(x)}^2$. Then

$$\begin{aligned} \pi^2(F) &= \pi((\text{restr}_t^1)^{-1}F) = \int \pi_x \left(\bigcup_{y \in E} (\text{restr}_t^1)^{-1}\text{spt } \pi_{\hat{T}(y)}^2 \right) d\mu_0(x) \\ &= \int_E \pi_x \left(\bigcup_{y \in E} (\text{restr}_t^1)^{-1}\text{spt } \pi_{\hat{T}(y)}^2 \right) d\mu_0(x) + \int_{X \setminus E} \pi_x \left(\bigcup_{y \in E} (\text{restr}_t^1)^{-1}\text{spt } \pi_{\hat{T}(y)}^2 \right) d\mu_0(x) \\ &= \int_E \pi_x \left((\text{restr}_t^1)^{-1}\text{spt } \pi_{\hat{T}(x)}^2 \right) d\mu_0(x) + \int_{X \setminus E} \pi_x \left(\bigcup_{y \in E} (\text{restr}_t^1)^{-1}\text{spt } \pi_{\hat{T}(y)}^2 \right) d\mu_0(x) \\ &< \mu_0(E) + \int_{X \setminus E} \pi_x \left(\bigcup_{y \in E} (\text{restr}_t^1)^{-1}\text{spt } \pi_{\hat{T}(y)}^2 \right) d\mu_0(x) \\ &= \mu_0(E) + \mu_0(\{x \in X \setminus E : \exists y \in E \text{ for which } \hat{T}(x) = \hat{T}(y)\}) \\ &= \int_E \pi_{\hat{T}(x)}^2 \left(\bigcup_{y \in E} \text{spt } \pi_{\hat{T}(y)}^2 \right) d\mu_0(x) + \int_{X \setminus E} \pi_{\hat{T}(x)}^2 \left(\bigcup_{y \in E} \text{spt } \pi_{\hat{T}(y)}^2 \right) d\mu_0(x) \\ &= \pi^2(F) \end{aligned}$$

which is a contradiction. Thus we have proven the lemma. \square

REFERENCES

1. Andrei Agrachev and Paul Lee, *Optimal transportation under nonholonomic constraints*, Trans. Amer. Math. Soc. **361** (2009), no. 11, 6019–6047. MR 2529923
2. L. Ambrosio and S. Rigot, *Optimal mass transportation in the Heisenberg group*, J. Funct. Anal. **208** (2004), no. 2, 261–301. MR 2035027
3. Luigi Ambrosio, Nicola Gigli, Andrea Mondino, and Tapio Rajala, *Riemannian Ricci curvature lower bounds in metric measure spaces with σ -finite measure*, Trans. Amer. Math. Soc. **367** (2015), no. 7, 4661–4701. MR 3335397
4. Luigi Ambrosio, Nicola Gigli, and Giuseppe Savaré, *Metric measure spaces with Riemannian Ricci curvature bounded from below*, Duke Math. J. **163** (2014), no. 7, 1405–1490. MR 3205729
5. Jérôme Bertrand, *Existence and uniqueness of optimal maps on Alexandrov spaces*, Adv. Math. **219** (2008), no. 3, 838–851. MR 2442054
6. Yann Brenier, *Décomposition polaire et réarrangement monotone des champs de vecteurs*, C. R. Acad. Sci. Paris Sér. I Math. **305** (1987), no. 19, 805–808. MR 923203
7. Dmitri Burago, Yuri Burago, and Sergei Ivanov, *A course in metric geometry*, Graduate Studies in Mathematics, vol. 33, American Mathematical Society, Providence, RI, 2001. MR 1835418
8. Fabio Cavalletti and Martin Huesmann, *Existence and uniqueness of optimal transport maps*, Ann. Inst. H. Poincaré Anal. Non Linéaire **32** (2015), no. 6, 1367–1377. MR 3425266

9. Fabio Cavalletti and Andrea Mondino, *Optimal maps in essentially non-branching spaces*, Commun. Contemp. Math. **19** (2017), no. 6, 1750007, 27. MR 3691502
10. Alessio Figalli and Ludovic Rifford, *Mass transportation on sub-Riemannian manifolds*, Geom. Funct. Anal. **20** (2010), no. 1, 124–159. MR 2647137
11. Nicola Gigli, *Optimal maps in non branching spaces with Ricci curvature bounded from below*, Geom. Funct. Anal. **22** (2012), no. 4, 990–999. MR 2984123
12. Nicola Gigli, Andrea Mondino, and Giuseppe Savaré, *Convergence of pointed non-compact metric measure spaces and stability of Ricci curvature bounds and heat flows*, Proc. Lond. Math. Soc. (3) **111** (2015), no. 5, 1071–1129. MR 3477230
13. Nicola Gigli, Tapio Rajala, and Karl-Theodor Sturm, *Optimal maps and exponentiation on finite-dimensional spaces with Ricci curvature bounded from below*, J. Geom. Anal. **26** (2016), no. 4, 2914–2929. MR 3544946
14. Martin Kell, *Transport maps, non-branching sets of geodesics and measure rigidity*, Adv. Math. **320** (2017), 520–573. MR 3709114
15. Christian Ketterer and Tapio Rajala, *Failure of topological rigidity results for the measure contraction property*, Potential Anal. **42** (2015), no. 3, 645–655. MR 3336992
16. Stefano Lisini, *Characterization of absolutely continuous curves in Wasserstein spaces*, Calc. Var. Partial Differential Equations **28** (2007), no. 1, 85–120. MR 2267755
17. John Lott and Cédric Villani, *Ricci curvature for metric-measure spaces via optimal transport*, Ann. of Math. (2) **169** (2009), no. 3, 903–991. MR 2480619
18. Robert J. McCann, *Polar factorization of maps on Riemannian manifolds*, Geom. Funct. Anal. **11** (2001), no. 3, 589–608. MR 1844080
19. Tapio Rajala and Karl-Theodor Sturm, *Non-branching geodesics and optimal maps in strong $CD(K, \infty)$ -spaces*, Calculus of Variations and Partial Differential Equations **50** (2014), no. 3, 831–846.
20. Karl-Theodor Sturm, *On the geometry of metric measure spaces. I*, Acta Math. **196** (2006), no. 1, 65–131. MR 2237206
21. Cédric Villani, *Optimal transport*, Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], vol. 338, Springer-Verlag, Berlin, 2009, Old and new. MR 2459454

UNIVERSITY OF JYVASKYLA, DEPARTMENT OF MATHEMATICS AND STATISTICS, P.O. BOX 35 (MAD),
FI-40014 UNIVERSITY OF JYVASKYLA, FINLAND

E-mail address: timo.m.schultz@student.jyu.fi