Tapio Rajala⃝ID · Timo Schultz

Optimal transport maps on Alexandrov spaces revisited

Abstract. We give an alternative proof for the fact that in n-dimensional Alexandrov spaces with curvature bounded below there exists a unique optimal transport plan from any purely $(n-1)$-rectifiable starting measure, and that this plan is induced by an optimal map. Our proof does not rely on the full optimality of a given plan but rather on the c-monotonicity, thus we obtain the existence of transport maps for wider class of (possibly non-optimal) transport plans.

1. Introduction

The problem of optimal mass transportation has a long history, starting from the work of Monge [34] in the late 18th century. In the original formulation of the problem, nowadays called the Monge-formulation, the problem is to find the transport map T minimizing the transportation cost

$$\int_{\mathbb{R}^n} c(x, T(x)) \, d\mu_0(x), \tag{1.1}$$

among all Borel maps $T : \mathbb{R}^n \to \mathbb{R}^n$ transporting a given probability measure μ_0 to another given probability measure μ_1, that is, $T_\sharp \mu_0 = \mu_1$. In the original problem of Monge, the cost function $c(x, y)$ was the Euclidean distance. Later, other cost functions have been considered, in particular much of the study has involved the distance squared cost, $c(x, y) = |x - y|^2$, which is the cost studied also in this paper.

In the Monge-formulation (1.1) of the optimal mass transportation problem the class of admissible maps T that send μ_0 to μ_1 is in most cases not closed in any suitable topology. To overcome this problem, Kantorovich [26,27] considered a larger class of optimal transports, namely, measures π on $\mathbb{R}^n \times \mathbb{R}^n$ such that the first marginal of π is μ_0 and the second is μ_1. Such measures π are called transport

Tapio Rajala (✉)· Timo Schultz
Department of Mathematics and Statistics, University of Jyvaskyla, P.O. Box 35 (MaD), 40014 Jyvaskyla, Finland.
e-mail: tapio.m.rajala@jyu.fi; tschultz@math.uni-bielefeld.de

plans. Kantorovich's relaxation leads to the so-called Kantorovich-formulation of the problem,

$$\inf_{\pi} \int_{\mathbb{R}^n \times \mathbb{R}^n} c(x, y) \, \mathrm{d}\pi(x, y). \tag{1.2}$$

Due to the closedness of the admissible transport plans and the lower semi-continuity of the cost, minimizers exist in the Kantorovich-formulation under very mild assumptions on the underlying space and the cost c.

For the quadratic cost in the Euclidean space, it was shown independently by Brenier [9] and Smith and Knott [42] that having μ_0 absolutely continuous with respect to the Lebesgue measure guarantees that the optimal transport plans (minimizer of (1.2)) are unique and given by a transport map. Moreover, the optimal transport map is given by a gradient of a convex function.

The results of Brenier and of Smith and Knott have been generalized in many ways. The most important directions of generalization have been: going from the underlying space $\mathbb{R}^n$ to other metric spaces, considering other cost functions, and relaxing the assumption of the starting measure being absolutely continuous with respect to the reference measure (here the Lebesgue measure). In this paper, we study the direction of relaxing the absolute continuity in a more general metric space setting, the Alexandrov spaces. We note that one should be able to generalize our proof for more general costs, such as the distance to a power $p \in (1, \infty)$. In order to keep the presentation simpler, we concentrate here on the distance squared cost.

The existence of optimal transportation maps in Alexandrov spaces with curvature bounded below for starting measures that are absolutely continuous with respect to the reference Hausdorff measure was proven by Bertrand [6]. Later Bertrand improved this result [7] by relaxing the assumption on the starting measure to give zero measure to $c - c$-hypersurfaces. Here we provide an alternative proof for the result of Bertrand under the slightly stronger assumption on the starting measure of pure $(n-1)$-unrectifiability (see Definition 2.1 for the definition of pure $(n-1)$-unrectifiability).

Theorem 1.1. *Let (X, d) be an n-dimensional Alexandrov space with curvature bounded below. Then for any pair of measures $\mu_0, \mu_1 \in \mathcal{P}_2(X)$ such that μ_0 is purely $(n-1)$-unrectifiable, every c-monotone plan π from μ_0 to μ_1 is induced by a map.*

In particular, there exists a unique optimal transport plan from μ_0 to μ_1 and this transport plan is induced by a map.

Remark 1.2. (1) The uniqueness of the optimal transport plan follows from the fact that $\mathrm{Opt}(\mu_0, \mu_1)$ is a convex subset of the set of c-monotone plans.
(2) While the motivation for the above formulation of Theorem 1.1 arises from the optimal mass transportation theory, it could be restated in the spirit of regularity of monotone operators, cf. [46,48].

The contribution of this paper is to provide a different approach to showing the existence and uniqueness of optimal transport maps than what was used by

Bertrand in [6,7]. In [6], Bertrand used the local $(1 + \varepsilon)$-biLipschitz maps to $\mathbb{R}^n$ on the regular set of X, and the general existence of Kantorovich potentials and their Lipschitzness. Since the singular set of X is at most $(n - 1)$-dimensional, and the Rademacher's theorem on $\mathbb{R}^n$ can be restated in X via the biLipschitz maps, Bertrand concluded that the optimal transport is concentrated on a graph that is given by applying the exponential map to the gradient of the Kantorovich potential. In [7], Bertrand considered the problem in boundaryless Alexandrov spaces. He used Perelman's DC calculus to translate the problem to differentiability of convex functions on Euclidean spaces. Then the result follows from the characterization of nondifferentiability points of convex functions due to Zajíček [47].

In this paper, we translate a contradiction argument (Lemma 2.11) from the Euclidean space (which uses just monotonicity in certain geometric configurations) to the space X via the $(1 + \varepsilon)$-biLipschitz charts. In order to use the contradiction argument, we need to get all the used distances to be comparable. For this we use the fact that the directions of geodesics are well-defined in the biLipschitz charts (Theorem 2.7) and thus we can contract along the geodesics without changing the geometric configuration too much. Finally, the geometric configurations that result in the contradiction via cyclical monotonicity are given by the pure $(n - 1)$-unrectifiability (Lemma 2.2). Let us briefly describe the contradiction argument in the Euclidean case $X = \mathbb{R}^2$ under the assumption that the starting measure μ_0 is absolutely continuous with respect to the Lebesgue measure. Suppose towards a contradiction that we have an optimal transport π transporting μ_0 to μ_1 so that π is not induced by a map. Then, after some discretizations, we find a positive measure set A of points where π transports measure to two different directions that are roughly some directions v_1 and v_2. Then, since A has positive μ_0-measure and μ_0 is absolutely continuous, in a Lebesgue point X of A there is another point y nearby roughly in the direction $v_1 - v_2$ from x. But now, the lines from x to the direction v_1 and from y to the direction v_2 cross. Such crossing violates the optimality of π because by interchanging the endpoints of the transports corresponding to the intersecting lines, we would decrease the cost of π.

To the best of our knowledge, in the context of optimal transportation this contradiction argument was first used by Champion, De Pascale and Juutinen [16] to prove the existence of optimal maps for the ∞-transportation distance. Similar idea was also used by Champion and De Pascale [14] to solve the Monge problem in $\mathbb{R}^d$. The limits of the contradiction argument were later pushed further by Champion and De Pascale [15] and by Jylhä [25].

Let us comment also on the history of the sufficient assumptions on μ_0. The assumption of pure $(n - 1)$-unrectifiability was shown by McCann [33] to be sufficient for the existence of optimal maps in the case of Riemannian manifolds. A sharper condition based on the characterization by Zajíček [47] of the set of non-differentiability points of convex functions was first used in the Euclidean context by Gangbo and McCann [20] when they showed that having an initial measure that gives zero mass to $c - c$-hypersurfaces is sufficient to give the existence of optimal maps. It was then shown by Gigli [21] that even in the Riemannian manifold context the sharp requirement for the starting measure to have optimal maps for any target measure is indeed that it gives zero measure to $c - c$-hypersurfaces.

It still remains open whether zero measure of $c - c$ -hypersurfaces also gives a full characterization in the case of Alexandrov spaces. One of the directions, the sufficiency, was obtained by Bertrand [7].

The existence of optimal maps has been studied in wider classes of metric measure spaces that satisfy some form of Ricci curvature lower bounds or weak versions of measure contraction property. These classes include $CD(K, N)$-spaces that were introduced by Lott and Villani [32], and by Sturm [43,44], $MCP(K, N)$-spaces (see Ohta [35]), and $RCD(K, N)$ spaces that were first introduced in the case $N = \infty$ by Ambrosio, Gigli and Savaré [3] and then for general N by Gigli [23] (see also the improvements and later work by Ambrosio, Gigli, Mondino and Rajala [1], Erbar, Kuwada and Sturm [18] and Ambrosio, Mondino and Savaré [4]). All of these classes contain Alexandrov spaces with curvature lower bounds, see Petrunin [37].

It was first shown by Gigli [22], that in nonbranching $CD(K, N)$-spaces you do have the existence of optimal maps provided that the starting measure is absolutely continuous with respect to the reference measure. In all the subsequent work, the assumption has been the same for the starting measure, and it would be interesting to see if it can be relaxed also in the more general context of metric measure spaces with Ricci curvature lower bounds.

Also a metric version of Brenier's theorem was studied by Ambrosio, Gigli and Savaré [2]. They did not obtain the existence of optimal maps, but showed that at least the transportation distance is given by the Kantorovich potential. Later, Ambrosio and Rajala [5] showed that under sufficiently strong nonbranching assumptions one can conclude the existence of optimal maps.

Rajala and Sturm [39] noticed that strong $CD(K, \infty)$ spaces, and hence $RCD(K, \infty)$ spaces are at least essentially nonbranching, and that this weaker form of nonbranching is sufficient for carrying out Gigli's proof. This result was later improved by Gigli, Rajala and Sturm [24]. Essential nonbranching was then studied together with the measure contraction property $MCP(K, N)$ by Cavalletti and Mondino [13] (see also Cavalletti and Huesmann [12] where the case of non-branching and a weaker version of $MCP(K, N)$ was considered), and finally it was shown by Kell [29] that under a weak type measure contraction property, the essential nonbranching characterizes the uniqueness of optimal transports and that the unique optimal transport is given by a map for absolutely continuous starting measures. The role of nonbranching and measure contraction type properties was also studied by De Pascale and Rigot [17] in connection with their sollution of the Monge problem in the Heisenberg group. See also the work of Bianchini and Cavalletti [8] on the Monge problem in nonbranching geodesic spaces.

The existence of optimal transport maps in $CD(K, N)$ spaces without any extra assumption on nonbranching is still an open problem. An intermediate definition between $CD(K, N)$ and essentially nonbranching $CD(K, N)$, called very strict $CD(K, N)$, was studied by Schultz [40]. He showed that in these spaces one still has optimal transport maps even if the space could be highly branching and the optimal plans non-unique. It is also worth noting that if one drops the assumption of essential nonbranching for $MCP(K, N)$, then optimal transport maps need not exist. This is seen from the examples by Ketterer and Rajala [30].

The paper is organized as follows. In Sect. 2 we recall basic things about rectifiability, Alexandrov spaces and optimal mass transportation. While doing this, we also present a few facts that easily follow from well-known results: purely $n-1$-unrectifiable measures have mass in all directions (Lemma 2.2), the singular set in an Alexandrov space is $(n-1)$-rectifiable (Theorem 2.5), gradients of geodesics exist in charts in Alexandrov spaces (Theorem 2.7) and the failure of cyclical monotonicity persists after small perturbations (Lemma 2.11). In Sect. 3 we then put these things together and prove Theorem 1.1.

2. Preliminaries

In this paper (X, d) always refers to a complete and locally compact length space. By a length space we mean a metric space where the distance between any two points x and y is equal to the infimum of lengths of curves connecting x and y. By the Hopf-Rinow-Cohn-Vossen Theorem, our spaces (X, d) are then geodesic, proper and, in particular, separable. A space is called geodesic, if any two points in the space can be connected by a geodesic. By a geodesic we mean a constant speed length minimizing curve $\gamma : [0, 1] \to X$. Notice that we parametrize all the geodesics by the unit interval. We denote the space of geodesics of X by $\mathrm{Geo}(X)$ and equip it with the supremum-distance. By a (geodesic) triangle $\Delta(x, y, z)$ we mean points $x, y, z \in X$ and any choice of geodesics $[x, y]$, $[y, z]$ and $[x, z]$ pairwise connecting them.

2.1. Rectifiability

For our Theorem 1.1 the starting measure μ_0 is diffused enough if it is purely $n-1$-unrectifiable. Let us recall this notion.

Definition 2.1. A set $A \subset X$ is called *(countably) k-rectifiable* if there exist Lipschitz maps $f_i : E_i \to X$ from Borel sets $E_i \subset \mathbb{R}^k$ for $i \in \mathbb{N}$, such that $A \subset \bigcup_{i \in \mathbb{N}} f_i(E_i)$.

A measure μ is called *purely k-unrectifiable*, if $\mu(A) = 0$ for every k-rectifiable set A.

The property of purely unrectifiable measures that we use is that they have mass in all directions. This is made precise using (one-sided) cones that are defined as follows. Given $x \in \mathbb{R}^n$, $\theta \in^{n-1}$, $\alpha > 0$ and $r > 0$, we denote the open cone at x in direction θ with opening angle α, by

$$C(x, \theta, \alpha) := \left\{ y \in \mathbb{R}^n \ : \ \langle y - x, \theta \rangle > \cos(\alpha)|y - x| \right\}.$$

Lemma 2.2. *Let μ be a purely $(n-1)$-unrectifiable measure on $\mathbb{R}^n$ and let $E \subset \mathbb{R}^n$ with $\mu(E) > 0$. Then at μ-almost every $x \in E$ we have $C(x, \theta, \alpha) \cap B(x, r) \cap E \neq \emptyset$ for all $\theta \in^{n-1}$, $\alpha > 0$ and $r > 0$.*

Proof. Suppose that there is a subset $E_0 \subset E$ with $\mu(E_0) > 0$ such that the conclusion fails, i.e. for every $x \in E_0$ there exist $\theta_x \in^{n-1}, \alpha_x > 0$ and $r_x > 0$ such that $C(x, \theta_x, \alpha_x) \cap B(x, r_x) \cap E = \emptyset$. Since

$$C(x, \theta, \alpha) \cap B(x, r) \subset C(x, \theta, \alpha') \cap B(x, r')$$

if $\alpha' \geq \alpha$ and $r' \geq r$, there exist $r > 0$ and $\alpha > 0$ such that the subset

$$\{x \in E_0 \,:\, C(x, \theta_x, \alpha) \cap B(x, r) \cap E = \emptyset\}$$

has positive μ-measure. By considering a countable dense set of directions $\{\theta_i\}_{i \in \mathbb{N}}$, we have that there exists one fixed direction θ_i such that the set

$$E_1 := \{x \in E_0 \,:\, C(x, \theta_i, \alpha/2) \cap B(x, r) \cap E = \emptyset\}$$

has positive μ-measure. But now, for evey $x \in \mathbb{R}^n$, the set $E_1 \cap B(x, r/2)$ is contained in a Lipschitz graph and hence E_1 is an $(n - 1)$-rectifiable set, giving a contradiction with the pure $(n - 1)$-unrectifiability of μ.   □

2.2. *Alexandrov spaces*

Let us recall some basics about Alexandrov spaces. Unless we provide another source, all the following definitions and results can be found in [10].

Alexandrov spaces generalize sectional curvature bounds by means of comparison to constant curvature model spaces. Alexandrov spaces can be defined for instance by comparing geodesic triangles of a metric space to the corresponding ones in a model space. Let us next give precise definitions.

For each $k \in \mathbb{R}$, let M_k be a simply connected surface with constant sectional curvature equal to k, that is, for negative k, M_k is a scaled hyperbolic plane, for $k = 0$, M_k is the Euclidean plane, and for positive k, M_k is a (round) sphere. Let us denote the distance between two points $x, y \in M_k$ by $|x - y|$.

Let $k \in \mathbb{R}$. For a triplet $x, y, z \in X$, let $\tilde{x}, \tilde{y}, \tilde{z} \in M_k$ be points so that the triangles $\Delta(x, y, z)$ and $\Delta(\tilde{x}, \tilde{y}, \tilde{z})$ have the same side lengths, that is, $d(x, y) = |\tilde{x} - \tilde{y}|, d(y, z) = |\tilde{y} - \tilde{z}|, d(x, z) = |\tilde{x} - \tilde{z}|$. We call the triangle $\Delta(\tilde{x}, \tilde{y}, \tilde{z})$ a comparison triangle for $\Delta(x, y, z)$. For a triangle $\Delta(x, y, z)$ in X we denote by $\tilde{\angle}_k(y, x, z)$ the comparison angle at $\tilde{x}$ in the comparison triangle $\Delta(\tilde{x}, \tilde{y}, \tilde{z})$ in M_k.

Definition 2.3. (Alexandrov space) We say that (X, d) is an Alexandrov space (with curvature bounded below by k) if there exists $k \in \mathbb{R}$ so that for each point $p \in X$ there exists a neighbourhood U of p for which the following holds. If $\Delta(x, y, z) \subset U$, $\Delta(\tilde{x}, \tilde{y}, \tilde{z})$ its comparison triangle in M_k, and $w \in [x, y]$, $\tilde{w} \in [\tilde{x}, \tilde{y}]$ with $d(x, w) = |\tilde{x} - \tilde{w}|$, then $d(w, z) \geq |\tilde{w} - \tilde{z}|$.

An Alexandrov space might have infinite (Hausdorff) dimension. In this paper we study only finite dimensional Alexandrov spaces. Recall that in an Alexandrov space every open nonempty set has the same dimension, so the dimension of an Alexandrov space is always well defined. Moreover, the dimension is either an

integer or infinity. From now on, the space (X, d) is assumed to be an n-dimensional Alexandrov space with curvature bounded below by $k \in \mathbb{R}$ with $n \in \mathbb{N}$.

We will use the fact that our purely $(n-1)$-unrectifiable starting measures μ_0 live on the regular set of the space, that has nice charts. Let us recall the notion of regular and singular points.

Definition 2.4. A point $p \in X$ is called *regular*, if the space of directions Σ_p at p is isometric to the standard sphere $^{n-1}$, or equivalently, if the Gromov-Hausdorff tangent at p is the Euclidean $\mathbb{R}^n$. A point $p \in X$ that is not regular is called *singular*. The set of regular points of X is denoted by $\mathrm{Reg}(X)$ and the set of singular points by $\mathrm{Sing}(X)$.

The following result is from [36] (see also [11]). It implies that our starting measures μ_0 give zero measure to the singular set.

Theorem 2.5. *The set* $\mathrm{Sing}(X)$ *is* $(n-1)$-*rectifiable.*

Proof. Notice that [36, Theorem A] states that $\mathrm{Sing}(X)$ has Hausdorff dimension at most $n - 1$. However, the proof easily gives the stronger conclusion of $(n-1)$-rectifiability. Namely, observe that in the proof of [36, Theorem A] Otsu and Shioya show that $\mathrm{Sing}(X)$ is contained in Lipschitz images from subsets of the spaces of directions Σ_p for countably many points $p \in X$. Since the points p are only needed to locally form a maximal ε-discrete net in X, they can be chosen to be regular points of X. Thus, $\mathrm{Sing}(X)$ is contained in countably many Lipschitz images from subsets of $^{n-1}$ and is therefore $(n-1)$-rectifiable. $\qquad\square$

Let us then recall a well-known consequence of the nonbranching property of Alexandrov spaces. For its proof, we need the notion of an angle. Let $\alpha, \beta \colon [0, 1] \to X$ be two constant speed geodesics emanating from the same point $p = \alpha(0) = \beta(0)$. Let us denote by $\theta_k(t, s) := \tilde{\measuredangle}_k(\alpha(t), p, \beta(s))$ the angle at $\tilde{p}$ of the comparison triangle $\Delta(\tilde{p}, \tilde{\alpha}(t), \tilde{\beta}(s))$ in M_k of $\Delta(p, \alpha(t), \beta(s))$. In Alexandrov spaces the angle

$$\measuredangle(\alpha, \beta) := \lim_{t,s \searrow 0} \theta_k(t, s)$$

is well-defined for every pair of geodesics α, β emanating from the same point. Moreover, by Alexandrov convexity (see for instance [41, Sect. 2.2]) the quantity $\theta_k(t, s)$ is monotone non-increasing in both variables t and s.

Lemma 2.6. *Let* $\gamma_1, \gamma_2 \colon [0, 1] \to X$ *be be two constant speed geodesics with* $\gamma_1(0) = \gamma_2(0)$ *and* $\gamma_1(1) \neq \gamma_2(1)$. *Then*

$$\lim_{t \searrow 0} \frac{d(\gamma_1(t), \gamma_2(t))}{t} > 0.$$

Proof. We may assume $\ell(\gamma_1) \geq \ell(\gamma_2)$. If $\ell(\gamma_1) > \ell(\gamma_2)$, then by triangle inequality $d(\gamma_1(t), \gamma_2(t)) \geq t(\ell(\gamma_1) - \ell(\gamma_2))$, giving the claim. If $\ell(\gamma_1) = \ell(\gamma_2)$, then $\theta_k(1, 1) = \tilde{\measuredangle}_k(\gamma_1(1), x, \gamma_2(1)) > 0$. Then by Alexandrov convexity, $\measuredangle(\gamma_1, \gamma_2) \geq \theta_k(1, 1) > 0$, and thus by the cosine law

$$\frac{d(\gamma_1(t), \gamma_2(t))}{t} \to \ell(\gamma_1)\sqrt{2 - 2\cos(\measuredangle(\gamma_1, \gamma_2))} > 0,$$

as $t \to 0$.   □

Our aim is to arrive at a contradiction with cyclical monotonicity at a small scale near a regular point. We will transfer the Euclidean argument to the Alexandrov space X using the following standard charts φ. Since we need the existence of directions of geodesics in these charts, we write the existence down explicitly inside the following theorem.

Theorem 2.7. *For every $p \in \mathrm{Reg}(X)$ and every $\varepsilon > 0$ there exist a neighborhood U of p and a $(1 + \varepsilon)$-biLipschitz map $\varphi \colon U \to \mathbb{R}^n$ with $\varphi(U)$ open so that for every constant speed geodesic $\gamma \colon [0, 1] \to U$ the limit*

$$\lim_{t \searrow 0} \frac{\varphi(\gamma(t)) - \varphi(\gamma(0))}{d(\gamma(t), \gamma(0))}$$

exists.

Proof. We recall (see [36] or [10, Theorem 10.8.4]) that the local $(1+\varepsilon)$-biLipschitz chart $\varphi \colon U \to \mathbb{R}^n$ can be obtained as

$$\varphi(x) = (d(a_1, x), d(a_2, x), \ldots, d(a_n, x)),$$

where $(a_i, b_i)_{i=1}^n$ is a δ-strainer for p, for some $\delta > 0$. Now, the first variation formula (see [36, Theorem 3.5] or [10, Theorem 4.5.6, Corollary 4.5.7]) implies that

$$\lim_{t \searrow 0} \frac{d(a_i, \gamma(t)) - d(a_i, \gamma(0))}{d(\gamma(t), \gamma(0))} = -\cos(\alpha),$$

where $\alpha = \angle(\gamma, \beta)$, with β a geodesic from $\gamma(0)$ to a_i. Thus, the required limit exists for each i.   □

2.3. Optimal mass transportation

In this section we recall a few basic things in optimal mass transportation.

The Monge–Kantorovich formulation of optimal mass transportation problem (with quadratic cost) is to investigate for two Borel probability measures μ_0 and μ_1 the following infimum

$$\inf \int_{X \times X} d^2(x, y) \, \mathrm{d}\pi(x, y),$$

where the infimum is taken over all Borel probability measures $\pi \in \mathcal{P}(X \times X)$ which has μ_0 and μ_1 as a marginals, that is, $\pi(A \times X) = \mu_0(A)$ and $\pi(X \times A) = \mu_1(A)$ for all Borel sets $A \in \mathcal{B}(X)$. In order to guarantee that the above infimum is finite, it is standard to assume the measures μ_0 and μ_1 to have finite second moments. The set of all Borel probability measures in X with finite second moments is denoted by $\mathcal{P}_2(X)$.

An admissible measure that minimizes the above infimum is called an optimal (transport) plan, and the set of optimal plans between μ_0 and μ_1 is denoted by

Opt(μ_0, μ_1). We say that an optimal plan π is induced by a map, if there exists a Borel measurable function $T : X \rightarrow X$ so that $\pi = (\text{id} \times T)_{\#}\mu_0$. Such a map is called an optimal (transport) map. While optimal plans exist under fairly general assumptions [45], the existence of optimal maps is not true in general.

Optimality of a given transport plan depends only on the c-cyclical monotonicity of the support of the plan. Let us recall this notion.

Definition 2.8. (cyclical monotonicity) A set $\Gamma \subset X \times X$ is called c-cyclically monotone, if for all finite sets of points $\{(x_i, y_i)\}_{i=1}^N \subset \Gamma$ the inequality

$$\sum_{i=1}^N d^2(x_i, y_i) \leq \sum_{i=1}^N d^2(x_{\sigma(i)}, y_i) \qquad (2.1)$$

holds for all permutations $\sigma \in S_N$ of $\{1, \ldots, N\}$.

If (2.1) is required to hold only in the case of $N = 2$ (i.e. for pairs of points), then the set Γ is called c-monotone.

A characterization of optimality using c-cyclical monotonicity of the support that is sufficient for us is the following result proven in [38] which holds for continuous cost functions.

Lemma 2.9. ([38, Theorem B]) *Let X be a Polish space and $\mu_0, \mu_1 \in \mathcal{P}_2(X)$. Then a transport plan π between μ_0 and μ_1 is optimal if and only if its support is c-cyclically monotone set.*

In the following lemma we recall a well-known fact which allows us to localize the problem.

Lemma 2.10. *Let (X, d) be a complete and separable geodesic metric space, and let $\Gamma \subset X \times X$ be a c-monotone set. Then, the set*

$$\Gamma_t := \{(\gamma(0), \gamma(t)) \in X \times X : \gamma \in \text{Geo}(X) \text{ with } (\gamma(0), \gamma(1)) \in \Gamma\}$$

is c-monotone for all $t \in [0, 1]$.

Proof. For a given pair of points $(\gamma_0^1, \gamma_t^1), (\gamma_0^2, \gamma_t^2) \in \Gamma_t$, the set $\{(\gamma_0^1, \gamma_t^1), (\gamma_0^2, \gamma_t^2)\}$ is c-cyclically monotone. Thus, the claim follows from the same statement for c-cyclically monotone sets which is turn can be deduced from the result of Lisini in [31] about Wasserstein geodesics and their lifts to the space of probability measures on geodesics of X, see [19]. $\square$

In order to arrive at a contradiction with monotonicity, we will use the following lemma.

Lemma 2.11. *For each $C > 1$ there exists $\delta > 0$ so that*

$$\frac{1}{2}|y_1 + y_2|^2 < (1 - \delta)(|y_1|^2 + |y_2|^2)$$

for all

$$y_1, y_2 \in K := \left\{ (y_1, y_2) \in \mathbb{R}^{2n} : |y_2| = 1 \text{ and } |y_2 - y_1| \in \left[\frac{1}{C}, C\right] \right\}.$$

Proof. Let us first observe that for $y_1, y_2 \in \mathbb{R}^n$, with $y_1 \neq y_2$ we have

$$0 < |y_1 - y_2|^2 = |y_1|^2 - 2\langle y_1, y_2 \rangle + |y_2|^2$$

and thus

$$|y_1 + y_2|^2 = |y_1|^2 + 2\langle y_1, y_2 \rangle + |y_2|^2 < 2(|y_1|^2 + |y_2|^2). \qquad (2.2)$$

The quantitative claim then follows by compactness of K: first of all notice that $K \subset \bar{B}(0, 2 + C)$ and thus K is bounded. The set K is also closed and hence it is compact. The function

$$(y_1, y_2) \mapsto \frac{|y_1 + y_2|^2}{|y_1|^2 + |y_2|^2}$$

is continuous as a function $K \to \mathbb{R}$. Therefore, the maximum of the above function is achieved in K. By (2.2), this maximum is strictly less than two and hence there exists $\delta > 0$ as in the claim. $\qquad \square$

3. Proof of Theorem 1.1

In order to prove the uniqueness of optimal transport plans it suffices to show that any optimal transport plan is induced by a map. Indeed, if there were two different optimal plans π_1 and π_2, then their convex combination $\frac{1}{2}(\pi_1 + \pi_2)$ would also be optimal and not given by a map. We will prove Theorem 1.1 by assuming that there exists a c-monotone plan that is not induced by a map, then localizing to a chart and using an Euclidean argument to find a contradiction.

Step 1: initial uniform bounds and measurable selections
Let $\mu_0, \mu_1 \in \mathcal{P}_2(X)$ with μ_0 purely $(n-1)$-unrectifiable. Let π be a c-monotone plan from μ_0 to μ_1. Towards a contradiction, we assume that π is not induced by a map, that is, there does not exist a Borel map $T: X \to X$ so that $\pi = (\mathrm{id}, T)_\sharp \mu_0$. Consider the set

$$A := \{x \in X : \text{ there exist } y^1, y^2 \in X \text{ such that } (x, y^1), (x, y^2) \in \mathrm{spt}(\pi)\}.$$

Since A is a projection of a Borel set

$$\{(x, y, z, w) \in \mathrm{spt}(\pi) \times \mathrm{spt}(\pi) : d(x, z) = 0, \ d(y, w) > 0\},$$

it is a Souslin set and thus μ_0-measurable. (Actually, as a projection of a σ-compact set, A is Borel.) We will show that A has positive μ_0 measure.

For that we will first show that there exists a Borel selection $T: \mathrm{p}_1(\mathrm{spt}(\pi)) \to X$ of $\mathrm{spt}(\pi)$, where $\mathrm{p}_1: X \times X \to X$ is the projection to the first coordinate. Define

$$(\mathrm{spt}(\pi))_x := \{y \in X : (x, y) \in \mathrm{spt}(\pi)\}.$$

Then $(\mathrm{spt}(\pi))_x = (\{x\} \times X) \cap \mathrm{spt}(\pi)$ and thus it is closed. Furthermore, as a proper space, X is also σ-compact, and thus so is $(\mathrm{spt}(\pi))_x$. Hence, by the Arsenin-Kunugui Theorem [28, Theorem 35.46] there exists a Borel selection of $\mathrm{spt}(\pi)$, in

other words, there exists a Borel map $T\colon \mathrm{p}_1(\mathrm{spt}(\pi)) \to X$ with $\mathrm{p}_1(\mathrm{spt}(\pi))$ Borel so that $T(x) \in (\mathrm{spt}(\pi))_x$ for all $x \in \mathrm{p}_1(\mathrm{spt}(\pi))$.

Suppose now that $\mu_0(A) = 0$. We will show that in this case π would be induced by the map T. Indeed, for $E \subset X \times X$ we have that

$$(\mathrm{id}, T)_{\#}\mu_0(E) = \mu_0((\mathrm{id}, T)^{-1}(E) \setminus A) = \mu_0(\mathrm{p}_1(E \cap \mathrm{Graph}(T)))$$
$$= \pi((\mathrm{p}_1(E \cap \mathrm{Graph}(T)) \times X) \cap \mathrm{spt}(\pi))$$
$$= \pi((E \cap \mathrm{spt}(\pi)) \setminus (A \times X)) = \pi(E \cap \mathrm{spt}(\pi)) = \pi(E).$$

Thus $\mu_0(A) > 0$.

Since X is geodesic, for all $x \in A$ there exist $\gamma_x^1, \gamma_x^2 \in \mathrm{Geo}(X)$ such that $\gamma_x^1(0) = x = \gamma_x^2(0)$, $\gamma_x^1(1) \neq \gamma_x^2(1)$, and $(\gamma_x^i(0), \gamma_x^i(1)) \in \mathrm{spt}(\pi)$ for $i \in \{1, 2\}$. We will need to choose the geodesics γ_x^1 and γ_x^2 in a measurable way. We will also make the selection so that

$$d(x, \gamma_x^1(1)) \leq d(x, \gamma_x^2(1)) \neq 0. \tag{3.1}$$

By now, we have a Borel selection T of $\mathrm{spt}(\pi)$. Since $\mathrm{p}_1(\mathrm{spt}(\pi))$ is a Borel set, we can extend T to a Borel map $T\colon X \to X$. Consider now the set $\mathrm{spt}(\pi) \setminus \mathrm{Graph(T)}$. Since T is a Borel map, the graph of T is a Borel set and thus the set $\mathrm{spt}(\pi) \setminus \mathrm{Graph(T)}$ is a Borel set. Since $X \setminus T(x)$ is σ-compact by the properness and separability of X, we have that $(\mathrm{spt}(\pi) \setminus \mathrm{Graph(T)})_x$ is σ-compact as a closed subset of $X \setminus T(x)$. Thus again by the Arsenin-Kunugui Theorem there exists a Borel selection $S\colon \mathrm{p}_1(\mathrm{spt}(\pi) \setminus \mathrm{Graph(T)}) \to X$ that we can further extend to a Borel map $S\colon X \to X$ for which we have that $T(x) = S(x)$ for $x \notin A$, and $T(x) \neq S(x)$ for $x \in A$.

To have (3.1) we will define two auxiliary maps $\tilde{T}^1, \tilde{T}^2\colon X \to X \times X$ as

$$\tilde{T}^1(x) := \begin{cases} (x, T(x)), & x \in h^{-1}(-\infty, 0) \\ x, S(x)), & x \in h^{-1}(0, \infty), \end{cases}$$

where $h(x) := d(x, T(x)) - d(x, S(x))$, and similarly

$$\tilde{T}^2(x) := \begin{cases} (x, S(x)), & x \in h^{-1}(-\infty, 0) \\ x, T(x)), & x \in h^{-1}(0, \infty). \end{cases}$$

The maps $\tilde{T}^1$ and $\tilde{T}^2$ are Borel maps since T, S and h are Borel maps.

It remains to select the geodesics between points x and $T^i(x)$. For that, we consider the set

$$G := \{(x, y, \gamma) \in X \times X \times \mathrm{Geo}(X) : \gamma(0) = x, \gamma(1) = y\}.$$

The set G is Borel as the preimage of zero under the Borel map

$$(x, y, \gamma) \mapsto \sup\{d(x, \gamma(0)), d(y, \gamma(1))\}.$$

Furthermore, we have by the Arzelà-Ascoli Theorem that

$$G_{(x,y)} := \{\gamma \in \mathrm{Geo}(X) : \gamma(0) = x, \gamma(1) = y\}.$$

is compact. Thus, by the Arsenin-Kunugui Theorem there exists a Borel selection $F \colon X \times X \to G_{(x,y)}$. With this we may finally define $T^1, T^2 \colon X \to \mathrm{Geo}(X)$ as

$$T^1 := F \circ \tilde{T}^1 \quad \text{and}$$

$$T^2 := F \circ \tilde{T}^2.$$

From now on, we will denote $\gamma_x^1 = T^1(x)$ and $\gamma_x^2 = T^2(x)$ for all $x \in A$. Notice that γ_x^1 and γ_x^2 satisfy (3.1).

By Lemma 2.6, we have for all $x \in A$ that

$$\lim_{t \searrow 0} \frac{d(\gamma_x^1(t), \gamma_x^2(t))}{d(x, \gamma_x^2(t))} \in (0, \infty).$$

Thus, we may write A as a countable union of sets

$$A_i := \left\{ x \in A \,:\, d(x, \gamma_x^2(1)) \in [1/i, i] \text{ and } \frac{d(\gamma_x^1(t), \gamma_x^2(t))}{d(x, \gamma^2(t))} \in [1/i, i] \text{ for all } t \le \frac{1}{i} \right\},$$

and therefore there exists $k \in \mathbb{N}$ so that $\mu_0(A_k) > 0$. Notice that the sets A_i are measurable, since we can write A_i as the intersection of

$$\left\{ x \in A : d(x, \gamma_x^2(1)) \in [1/i, i] \right\}$$

and

$$\bigcap_{\substack{t \le \frac{1}{i} \\ t \in \mathbb{Q}}} \left\{ x \in X : \frac{d(\gamma_x^1(t), \gamma_x^2(t))}{d(x, \gamma_x^2(t))} \in [1/i, i] \right\}.$$

We now consider $k \in \mathbb{N}$ fixed so that $\mu_0(A_k) > 0$.

Step 2: localization to a chart

Now we are ready to localize the problem so that we may use properties of the Euclidean space to arrive to the contradiction. We will need to choose $\varepsilon > 0$ sufficiently small to arrive to a contradiction with c-monotonicity in a $(1+\varepsilon)$-chart given by Theorem 2.7. We define

$$\varepsilon := \frac{\delta}{100} \in (0, 1/200),$$

where $\delta = \delta(2k) \in (0, 1/2)$ is the constant given by Lemma 2.11 for the k fixed above. Since μ_0 is purely $(n-1)$-unrectifiable and $\mathrm{Sing}(X)$ is $(n-1)$-rectifiable by Theorem 2.5, we have $\mu_0(A_k \cap \mathrm{Reg}(X)) = \mu_0(A_k)$. By Theorem 2.7 we can cover the set $\mathrm{Reg}(X)$ with open sets U for which the associated maps $\varphi \colon U \to \mathbb{R}^n$ are $(1+\varepsilon)$-biLipschitz, and the limit

$$\lim_{t \searrow 0} \frac{\varphi(\gamma(t)) - \varphi(\gamma(0))}{d(\gamma(t), \gamma(0))}$$

exists for all geodesics $\gamma \subset U$. Since X is a proper metric space, it is in particular hereditarily Lindelöf. Therefore, there exists a countable subcover $\mathcal{F}$ of such open

sets U. Hence, there exists $U \in \mathcal{F}$ for which $\mu_0(U \cap A_k) > 0$. Let $\varphi \colon U \to \mathbb{R}^n$ be as in Theorem 2.7.

Step 3: discretization and choice of points for the contradiction
Next we take a subset of $A_k \cap U$ where the direction of the two selected geodesics is independent of the point, up to a small error

$$\hat{\varepsilon} := \frac{\varepsilon}{80k^4} > 0. \tag{3.2}$$

This is done by covering the set $\mathbb{R}^n$ by sets $\{B(y_i, \hat{\varepsilon})\}_{i \in \mathbb{N}}$. Then there exist i, j and $t_0 > 0$ so that the set

$$B := \left\{ x \in A_k \cap U \ : \ \frac{\varphi(\gamma_x^1(t)) - \varphi(x)}{t} \in B(y_i, \hat{\varepsilon}), \ \frac{\varphi(\gamma_x^2(t)) - \varphi(x)}{t} \in B(y_j, \hat{\varepsilon}), \right.$$
$$\left. \varphi(\gamma_x^1(t)), \varphi(\gamma_x^2(t)) \in U \text{ for all } t \le t_0 \right\}$$

has positive μ_0-measure. Notice that B is seen to be measurable by a similar argument than A_i. By relabeling, we may assume that $i = 1$ and $j = 2$.

Since φ is biLipschitz, the measure $\varphi_{\#}\mu_0$ is purely $(n-1)$-unrectifiable on $\mathbb{R}^n$. Hence, by Lemma 2.2 there exist points $x_1, x_2 \in B$ such that

$$\varphi(x_2) \in C\left(\varphi(x_1), \frac{y_2 - y_1}{|y_2 - y_1|}, \hat{\varepsilon} \right) \cap B(\varphi(x_1), r), \tag{3.3}$$

where $r \le \hat{\varepsilon}$ is such that $r \le \frac{t_0}{2}|y_2 - y_1|$. Now that we have selected the initial points x_1 and x_2 for the contradiction argument, we still need to bring the target points close enough to x_1 and x_2 by contracting along the geodesics $\gamma_{x_1}^2$ and $\gamma_{x_2}^1$. Since $|\varphi(x_1) - \varphi(x_2)| < r$, there exists the desired contraction parameter $t \le t_0$ for which

$$2|\varphi(x_2) - \varphi(x_1)| = |ty_2 - ty_1|. \tag{3.4}$$

We will now use as target points the points $\gamma_{x_1}^2(t)$ and $\gamma_{x_2}^1(t)$.

Step 4: verifying the bounds for Lemma 2.11
In the remainder of the proof we verify that the four selected points $x_2, x_1, \gamma_{x_1}^2(t)$ and $\gamma_{x_2}^1(t)$ give a contradiction with c-monotonicity. Towards this goal we first check that we may apply Lemma 2.11 with the selected δ.

First of all, we have by the definition of A_k that

$$\frac{|\varphi(\gamma_{x_1}^2(t)) - \varphi(\gamma_{x_1}^1(t))|}{|\varphi(\gamma_{x_1}^2(t)) - \varphi(x_1)|} \in \left[\frac{1}{(1+\varepsilon)^2 k}, (1+\varepsilon)^2 k \right].$$

Since

$$\hat{\varepsilon} \le \frac{\varepsilon}{2(1+\varepsilon)k^2},$$

we have by the fact that $x_1 \in A_k$ and φ is $(1+\varepsilon)$-biLipschitz, that

$$2t\hat{\varepsilon} \le \frac{\varepsilon}{(1+\varepsilon)} \frac{d(\gamma_{x_1}^2(t), x_1)}{k} \le \frac{\varepsilon d(\gamma_{x_1}^2(t), \gamma_{x_1}^1(t))}{(1+\varepsilon)} \le \varepsilon |\varphi(\gamma_{x_1}^2(t)) - \varphi(\gamma_{x_1}^1(t))|.$$

Similarly, since $\hat{\varepsilon} \leq \frac{\varepsilon}{(1+\varepsilon)k}$, we have that

$$t\hat{\varepsilon} \leq \varepsilon |\varphi(\gamma_{x_1}^2(t)) - \varphi(x_1)|.$$

Therefore, we have by the fact that $x_1 \in B$, the triangle inequality and the choice of ε and $\hat{\varepsilon}$ that

$$
\begin{aligned}
\frac{|ty_2 - ty_1|}{|ty_2|} &\leq \frac{|\varphi(\gamma_{x_1}^2(t)) - \varphi(\gamma_{x_1}^1(t))| + 2t\hat{\varepsilon}}{|\varphi(\gamma_{x_1}^2(t)) - \varphi(x_1)| - t\hat{\varepsilon}} \\
&\leq \frac{(1+\varepsilon)}{(1-\varepsilon)} \frac{|\varphi(\gamma_{x_1}^2(t)) - \varphi(\gamma_{x_1}^1(t))|}{|\varphi(\gamma_{x_1}^2(t)) - \varphi(x_1)|} \\
&\leq \frac{(1+\varepsilon)}{(1-\varepsilon)}(1+\varepsilon)^2 k < 2k.
\end{aligned}
$$

By similar arguments, we have that

$$\frac{|ty_2 - ty_1|}{|ty_2|} > \frac{1}{2k}.$$

Thus, by Lemma 2.11 with the $\delta = \delta(2k)$ already chosen accordingly, we have

$$\frac{\frac{1}{2}|t(y_1 + y_2)|^2}{|ty_2|^2} < (1-\delta)\frac{(|ty_2|^2 + |ty_1|^2)}{|ty_2|^2},$$

that is,

$$\frac{1}{2}|t(y_1 + y_2)|^2 < (1-\delta)(|ty_2|^2 + |ty_1|^2). \tag{3.5}$$

Step 5: the contradiction

We will then use the inequality (3.5) to get to a contradiction with the c-monotonicity guaranteed by Lemma 2.10. Let us first estimate the terms on the right-hand side of (3.5).

By the definition of y_1 and A_k we have that

$$
\begin{aligned}
|ty_1| &\leq |ty_1 - \varphi(\gamma_{x_1}^1(t)) + \varphi(x_1)| + |\varphi(\gamma_{x_1}^1(t)) - \varphi(x_1)| \\
&\leq t\hat{\varepsilon} + (1+\varepsilon)d(\gamma_{x_1}(t), x_1) \leq tk + (1+\varepsilon)tk \leq 3tk.
\end{aligned}
$$

Similarly,

$$|ty_2| \leq 3tk.$$

Therefore, we have that

$$\left|\frac{1}{2}t(y_1 + y_2)\right|, \left|\frac{1}{2}t(y_2 - y_1)\right| \leq 3tk. \tag{3.6}$$

Using the definition of the set B, and (3.4), (3.3) and (3.6), we have

$$\frac{1}{(1+\varepsilon)^2} d^2(x_2, \gamma_{x_1}^2(t))$$

$$\leq |\varphi(\gamma_{x_1}^2(t)) - \varphi(x_2)|^2$$

$$= |\frac{1}{2}t(y_1 + y_2) + (\varphi(\gamma_{x_1}^2(t)) - \varphi(x_1) - ty_2) - (\varphi(x_2) - \varphi(x_1) - \frac{1}{2}t(y_2 - y_1))|^2$$

$$\leq \left(|\frac{1}{2}t(y_1 + y_2)| + |\varphi(\gamma_{x_1}^2(t)) - \varphi(x_1) - ty_2| + |\varphi(x_2) - \varphi(x_1) - \frac{1}{2}t(y_2 - y_1)|\right)^2$$

$$\leq \left(|\frac{1}{2}t(y_1 + y_2)| + t\hat{\varepsilon} + \frac{1}{2}|t(y_2 - y_1)|\hat{\varepsilon}\right)^2 \leq \left(|\frac{1}{2}t(y_1 + y_2)| + (3k + 1)t\hat{\varepsilon}\right)^2$$

$$\leq |\frac{1}{2}t(y_1 + y_2)|^2 + 6tk(3k + 1)t\hat{\varepsilon} + ((3k + 1)t\hat{\varepsilon})^2 \leq |\frac{1}{2}t(y_1 + y_2)|^2 + 40t^2k^2\hat{\varepsilon}$$

and similarly

$$\frac{1}{(1+\varepsilon)^2} d^2(x_1, \gamma_{x_2}^2(t)) \leq |\frac{1}{2}t(y_1 + y_2)|^2 + 40t^2k^2\hat{\varepsilon}.$$

Thus, by summing the two terms, using (3.2) and the fact that $x_1 \in A_k$,

$$\frac{1}{(1+\varepsilon)^2}[d^2(x_2, \gamma_{x_1}^2(t)) + d^2(x_1, \gamma_{x_2}^1(t))]$$

$$\leq 2|\frac{1}{2}t(y_1 + y_2)|^2 + 80t^2k^2\hat{\varepsilon} \tag{3.7}$$

$$\leq \frac{1}{2}|t(y_1 + y_2)|^2 + \frac{t^2}{k^2}\varepsilon \leq \frac{1}{2}|t(y_1 + y_2)|^2 + \varepsilon d^2(\gamma_{x_1}^2(t), x_1).$$

Again, by the definition of the set B and the choice of $\hat{\varepsilon}$

$$|ty_1|^2 \leq \left((1+\varepsilon)d(\gamma_{x_2}^1(t), x_2) + t\hat{\varepsilon}\right)^2 \leq \left((1+\varepsilon)d(\gamma_{x_2}^1(t), x_2) + \varepsilon d(\gamma_{x_1}^2(t), x_1)\right)^2$$

$$\leq ((1+\varepsilon)^2 + 2(1+\varepsilon)\varepsilon)d^2(\gamma_{x_2}^1(t), x_2) + (\varepsilon^2 + 2(1+\varepsilon)\varepsilon)d^2(\gamma_{x_1}^2(t), x_1)$$

$$\leq (1 + 7\varepsilon)d^2(\gamma_{x_2}^1(t), x_2) + 5\varepsilon d^2(\gamma_{x_1}^2(t), x_1) \tag{3.8}$$

and

$$|ty_2|^2 \leq \left((1+\varepsilon)d(\gamma_{x_1}^2(t), x_1) + t\hat{\varepsilon}\right)^2$$

$$\leq (1 + 2\varepsilon)^2 d^2(\gamma_{x_1}^2(t), x_1) \leq (1 + 8\varepsilon)d^2(\gamma_{x_1}^2(t), x_1). \tag{3.9}$$

Using the inequalities (3.5), (3.8) and (3.9), we get that

$$\frac{1}{2}|t(y_1 + y_2)|^2 < (1 - \delta)(|ty_2|^2 + |ty_1|^2)$$

$$\leq (1 - \delta)(1 + 13\varepsilon)(d^2(\gamma_{x_1(t)}^2, x_1) + d^2(\gamma_{x_2(t)}^1, x_2)). \tag{3.10}$$

Hence, by (3.7), (3.10), the fact that $\delta \leq \frac{1}{2}$ and the choice of ε, we have that

$$d^2(x_2, \gamma_{x_1}^2(t)) + d^2(x_1, \gamma_{x_2}^1(t))$$
$$\leq (1+\varepsilon)^2 \left(\frac{1}{2}|t(y_1 + y_2)|^2 + \varepsilon d^2(\gamma_{x_1}^2(t), x_1) \right)$$
$$\leq (1+\varepsilon)^2(1-\delta)(1+15\varepsilon)(d^2(\gamma_{x_1}^2(t), x_1) + d^2(\gamma_{x_2}^1(t), x_2))$$
$$\leq (1-\delta)(1+100\varepsilon)(d^2(\gamma_{x_1}^2(t), x_1) + d^2(\gamma_{x_2}^1(t), x_2))$$
$$< d^2(x_2, \gamma_{x_2}^1(t)) + d^2(x_1, \gamma_{x_1}^2(t)).$$

However, since $(x_2, \gamma_{x_2}^1(1)), (x_1, \gamma_{x_1}^2(1)) \in \operatorname{spt}(\pi)$ we have by Lemma 2.10 that

$$d^2(x_2, \gamma_{x_2}^1(t)) + d^2(x_1, \gamma_{x_1}^2(t)) \leq d^2(x_2, \gamma_{x_1}^2(t)) + d^2(x_1, \gamma_{x_2}^1(t)).$$

which is a contradiction. Therefore, the plan π is induced by a map.

References

[1] Ambrosio, L., Gigli, N., Mondino, A., Rajala, T.: Riemannian Ricci curvature lower bounds in metric measure spaces with σ-finite measure. Trans. Am. Math. Soc. **367**(7), 4661–4701 (2015)

[2] Ambrosio, L., Gigli, N., Savaré, G.: Calculus and heat flow in metric measure spaces and applications to spaces with Ricci bounds from below. Invent. Math. **195**(2), 289–391 (2014)

[3] Ambrosio, L., Gigli, N., Savaré, G.: Metric measure spaces with Riemannian Ricci curvature bounded from below. Duke Math. J. **163**(7), 1405–1490 (2014)

[4] Ambrosio, Luigi, Mondino, Andrea, Savaré, Giuseppe: Nonlinear diffusion equations and curvature conditions in metric measure spaces, Mem. Amer. Math. Soc. **262** (2019), no.1270, v+121

[5] Ambrosio, L., Rajala, T.: Slopes of Kantorovich potentials and existence of optimal transport maps in metric measure spaces. Ann. Mat. Pura Appl. (4) **193**(1), 71–87 (2014)

[6] Bertrand, J.: Existence and uniqueness of optimal maps on Alexandrov spaces. Adv. Math. **219**(3), 838–851 (2008)

[7] Bertrand, J.: Alexandrov, Kantorovitch et quelques autres. Exemples d'interactions entre transport optimal et géométrie d'Alexandrov, Manuscrit présenté pour l'obtention de l'Habilitation à Diriger des Recherches (2015)

[8] Bianchini, S., Cavalletti, F.: The Monge problem for distance cost in geodesic spaces. Commun. Math. Phys. **318**(3), 615–673 (2013)

[9] Brenier, Y.: Polar factorization and monotone rearrangement of vector-valued functions. Commun. Pure Appl. Math. **44**(4), 375–417 (1991)

[10] Burago, D., Burago, Y., Ivanov, S.: A Course in Metric Geometry. Graduate Studies in Mathematics, vol. 33. American Mathematical Society, Providence (2001)

[11] Burago, Y., Gromov, M., Perel'man, G.: A. D. Aleksandrov spaces with curvatures bounded below. Uspekhi Mat. Nauk **47**, no. 2 (284), 3–51, 222 (1992)

[12] Cavalletti, F., Huesmann, M.: Existence and uniqueness of optimal transport maps. Ann. Inst. H. Poincaré Anal. Non Linéaire **32**(6), 1367–1377 (2015)

[13] Cavalletti, F., Mondino, A.: Optimal maps in essentially non-branching spaces. Commun. Contemp. Math. **19**(6), 1750007 (2017)

[14] Champion, T., De Pascale, L.: The Monge problem in $\mathbb{R}^d$. Duke Math. J. **157**(3), 551–572 (2011)

[15] Champion, T., De Pascale, L.: On the twist condition and c-monotone transport plans. Discrete Contin. Dyn. Syst. **34**(4), 1339–1353 (2014)

[16] Champion, T., De Pascale, L., Juutinen, P.: The ∞-Wasserstein distance: local solutions and existence of optimal transport maps. SIAM J. Math. Anal. **40**(1), 1–20 (2008)

[17] De Pascale, L., Rigot, S.: Monge's transport problem in the Heisenberg group. Adv. Calc. Var. **4**(2), 195–227 (2011)

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

[19] Galaz-García, F., Kell, M., Mondino, A., Sosa, G.: On quotients of spaces with Ricci curvature bounded below. J. Funct. Anal. **275**(6), 1368–1446 (2018)

[20] Gangbo, W., McCann, R.J.: The geometry of optimal transportation. Acta Math. **177**(2), 113–161 (1996)

[21] Gigli, N.: On the inverse implication of Brenier-McCann theorems and the structure of $(\mathcal{P}_2(M), W_2)$. Methods Appl. Anal. **18**(2), 127–158 (2011)

[22] Gigli, N.: Optimal maps in non branching spaces with Ricci curvature bounded from below. Geom. Funct. Anal. **22**(4), 990–999 (2012)

[23] Gigli, N.: On the differential structure of metric measure spaces and applications. Mem. Am. Math. Soc. **236**(1113), vi+91 (2015)

[24] Gigli, N., Rajala, T., Sturm, K.-T.: Optimal maps and exponentiation on finite-dimensional spaces with Ricci curvature bounded from below. J. Geom. Anal. **26**(4), 2914–2929 (2016)

[25] Jylhä, H.: The L^∞ optimal transport: infinite cyclical monotonicity and the existence of optimal transport maps. Calc. Var. Partial Differ. Equ. **52**(1–2), 303–326 (2015)

[26] Kantorovich, L.: On a problem of monge (in russian). Uspekhi Mat. Nauk. **3**, 225–226 (1948)

[27] Kantorovitch, L.: On the translocation of masses. C. R. (Doklady) Acad. Sci. URSS (N.S.) **37**, 199–201 (1942)

[28] Kechris, S.A.: Classical Descriptive Set Theory. Graduate Texts in Mathematics, vol. 156. Springer-Verlag, New York (1995)

[29] Kell, M.: Transport maps, non-branching sets of geodesics and measure rigidity. Adv. Math. **320**, 520–573 (2017)

[30] Ketterer, C., Rajala, T.: Failure of topological rigidity results for the measure contraction property. Potential Anal. **42**(3), 645–655 (2015)

[31] Lisini, S.: Characterization of absolutely continuous curves in Wasserstein spaces. Calc. Var. Partial Differ. Equ. **28**(1), 85–120 (2007)

[32] Lott, J., Villani, C.: Ricci curvature for metric-measure spaces via optimal transport. Ann. Math. (2) **169**(3), 903–991 (2009)

[33] McCann, R.J.: Polar factorization of maps on Riemannian manifolds. Geom. Funct. Anal. **11**(3), 589–608 (2001)

[34] Monge, G.: Mémoire sur la théorie des déblais et remblais, Histoire de l'Académie Royale des Sciences de Paris (1781), 666–704

[35] Ohta, S.: On the measure contraction property of metric measure spaces. Comment. Math. Helv. **82**(4), 805–828 (2007)

[36] Otsu, Y., Shioya, T.: The Riemannian structure of Alexandrov spaces. J. Differ. Geom. **39**(3), 629–658 (1994)

[37] Petrunin, A.: Alexandrov meets Lott–Villani–Sturm. Münster J. Math. **4**, 53–64 (2011)

[38] Pratelli, A.: On the sufficiency of c-cyclical monotonicity for optimality of transport plans. Math. Z. **258**(3), 677–690 (2008)

[39] Rajala, T., Sturm, K.-T.: Non-branching geodesics and optimal maps in strong $CD(K, \infty)$-spaces. Calc. Var. Partial Differ. Equ. **50**(3–4), 831–846 (2014)

[40] Schultz, T.: Existence of optimal transport maps in very strict $CD(K, \infty)$-spaces. Calc. Var. Partial Differ. Equ. **57**(5), Paper No. 139, 11 (2018)

[41] Shiohama, K.: An Introduction to the Geometry of Alexandrov Spaces. Lecture Notes Series, vol. 8. Seoul National University, Research Institute of Mathematics, Global Analysis Research Center, Seoul (1993)

[42] Smith, C.S., Knott, M.: Note on the optimal transportation of distributions. J. Optim. Theory Appl. **52**(2), 323–329 (1987)

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

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

[45] Villani, C.: Optimal transport, Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], vol. 338, Springer-Verlag, Berlin, Old and new (2009)

[46] Zajíček, L.: On the points of multiplicity of monotone operators. Comment. Math. Univ. Carolinae **19**(1), 179–189 (1978)

[47] Zajíček, Luděk: On the differentiation of convex functions in finite and infinite dimensional spaces, Czechoslovak Math. J. **29**(104) (1979), no. 3, 340–348

[48] Zarantonello, E.H.: Dense single-valuedness of monotone operators. Israel J. Math. **15**, 158–166 (1973)