TRANSACTIONS OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 363, Number 4, April 2011, Pages 1805–1819
S 0002-9947(2010)05065-2
Article electronically published on November 17, 2010

STEINER PROBLEMS IN OPTIMAL TRANSPORT

JONATHAN DAHL

Abstract. We study the Steiner problem of finding a minimal spanning network in the setting of a space of probability measures with metric defined by the cost of optimal transport between measures. The existence of a solution is shown for the Wasserstein space $P_p(\mathcal{X})$ over any base space $\mathcal{X}$ which is a separable, locally compact Hadamard space. Structural results are given for the case $P_2(\mathbb{R}^n)$.

1. Introduction

The problem of optimal transport, originally proposed by Monge, has a long history of investigation and application ([12] is an extensive reference). Roughly stated, the problem involves one who has an initial configuration of mass and would like to transport it to a terminal configuration of mass, doing so at least cost. For instance, one might have a set of water towers and a region of drought that one would like to relieve as quickly as possible. Abstractly, this becomes a constrained optimization problem in a space $\mathcal{P}$ of probability measures over the base space. Unfortunately, the mere existence of a solution is difficult to come by, due to the nonlinear nature of the problem. It was over two hundred years before Kantorovich [8, 7] provided serious progress by formulating and solving a weak version of the problem. We will focus on this Monge-Kantorovich problem, defined in detail below.

Returning to our drought problem, suppose that the drought and even the construction of the water towers have yet to occur. The question becomes where to build the water towers to best prepare for possible droughts. If there are multiple possible droughts one wishes to protect against, but one can only afford enough water towers to combat a single drought at a time, one wishes to find a configuration of water towers which is nicely balanced amongst the possible droughts. We will investigate this by solving Steiner-type problems in the space $\mathcal{P}$ of probabilities. A Steiner problem is a search for a length minimizing network, usually satisfying some boundary conditions, in a metric space.

As Steiner solutions can be considered generalized geodesics [10], we cannot reasonably hope to solve the classical Steiner problem if $\mathcal{P}$ is not a geodesic space. Therefore, we will also define and solve weak solutions of a Monge-Kantorovich-Steiner problem. The main idea in our definition of the weak problem is that, in a geodesic space, the edges of a Steiner solution are always geodesic segments, so the problem only sees distances of a finite point configuration. The weak problem

Received by the editors July 10, 2008.
2010 *Mathematics Subject Classification.* Primary 49Q20; Secondary 90C35, 49J10.
The author would like to thank Chikako Mese for suggesting the problem and for many helpful discussions, as well as the referee for recommendations on an earlier draft.

is then to minimize the sum of these distances in place of the sum of the lengths of the connecting paths. One may then discuss weak solutions of the problem in any metric space.

Steiner problems are traditionally solved via local compactness arguments; however, as we cannot expect local compactness from our space $\mathcal{P}$, we will instead need to argue using the geometry of the base space. In particular, we show:

Main Theorem. *Suppose $\mathcal{X}$ is a separable, locally compact Hadamard space or a compact complete metric space. Then for any $p > 1$, the parameterized and general versions of the Steiner problem are solvable in $(P_p(\mathcal{X}), W_p)$ for arbitrary boundary data.*

Here $(P_p(\mathcal{X}), W_p)$ is the p-Wasserstein space, whose definition and basic properties are recalled in Section 2.2. The argument actually gives a technically more general result, listed precisely as Theorem 21 in Section 3.3.

We prove the Main Theorem by first considering the weak Monge-Kantorovich-Steiner problem only for boundary data of compact support. The boundary data is then supported on a convex compact subset H of the base space. By examining the orthogonal projection onto H, we see that the problem reduces to finding a solution with measures having support restricted to remain in H. We may then apply Prokhorov's Theorem to show that the direct method of minimization gives the existence of a solution. The Main Theorem will then follow by a suitable approximation argument.

We conclude by using the geometry of Wasserstein spaces of order 2 to study the structure of the Steiner solutions. The standard variational argument for showing solutions of planar Steiner problems having vertex degrees of at most 3 is partially generalized to Steiner problems in $P_2(\mathbb{R}^n)$. We show that for boundary data of compact support, vertices of degree greater than 3 in a Steiner solution cannot be absolutely continuous with respect to Lebesgue measure.

2. Statement of the problem

2.1. **Steiner problems.** Steiner problems are concerned with finding minimal networks between a fixed set of points in a metric space. More precisely, a network Γ is a continuous map $\phi : G \to \mathcal{X}$ where $\mathcal{X}$ is a metric space and G is a graph, topologized in the standard way, called the parametric graph of Γ. ϕ may be decomposed into a union of curves, allowing one to compute the total length $l(\Gamma) = l(\phi)$ by working on each curve separately. The Steiner problem is to find a network of minimal length in some set of networks. We will focus on two cases.

Definition 1. The parameterized Steiner problem for a graph G and a metric space $\mathcal{X}$ is: given k vertices $v_1, \ldots, v_k \in G$ and k points $p_1, \ldots, p_k \in \mathcal{X}$, find a network of minimal length in the set of all networks in $\mathcal{X}$ with parametric graph G that sends v_i to p_i for $1 \leq i \leq k$.

Definition 2. The general Steiner problem for a metric space $\mathcal{X}$ is: given k points $p_1, \ldots, p_k \in \mathcal{X}$, find a network of minimal length in the set of all networks $\phi : G \to \mathcal{X}$ such that G is a connected graph and $p_1, \ldots, p_k$ are contained in the image of the vertex set.

We call the points $p_1, \ldots, p_k$ the boundary points of the problem. If $\mathcal{X}$ is a complete, locally compact, geodesic space, then the parameterized and general

Steiner problems are solvable for any boundary points (see [6]). Conversely, since for boundary points p_1, p_2 these Steiner problems are equivalent to the geodesic problem, $\mathcal{X}$ must be a geodesic space in order for solutions to exist for arbitrary boundary data. We will show by an explicit class of examples, however, that local compactness is not a necessary condition for the existence of solutions to arbitrary boundary data.

2.2. **Optimal transport.** Given two spaces $\mathcal{X}, \mathcal{Y}$ and two subsets of probability measures $\mathcal{P} \subset P(\mathcal{X})$ and $\mathcal{Q} \subset P(\mathcal{Y})$, we define the set of transport plans $\Pi(\mathcal{P}, \mathcal{Q})$ as the set of probability measures $\pi \in P(\mathcal{X} \times \mathcal{Y})$ such that $(\mathrm{proj}_{\mathcal{X}})_{\#}\pi \in \mathcal{P}$ and $(\mathrm{proj}_{\mathcal{Y}})_{\#}\pi \in \mathcal{Q}$. $(\mathrm{proj}_{\mathcal{X}})_{\#}\pi$ and $(\mathrm{proj}_{\mathcal{Y}})_{\#}\pi$ are called the marginals of π. Here $f_{\#}\mu$ denotes the push-forward of μ by f, defined by $f_{\#}\mu(A) = \mu(f^{-1}(A))$.

If we suppose that the cost of implementing a transport plan depends only on the structure of the spaces $\mathcal{X}$ and $\mathcal{Y}$, we might suppose that for some cost function $c : \mathcal{X} \times \mathcal{Y} \to \mathbb{R}$, the total cost of the transport plan is

$$\int_{\mathcal{X}\times\mathcal{Y}} c(x, y)\, d\pi(x, y).$$

If for some $\pi_0 \in P(\mathcal{X} \times \mathcal{Y})$ with marginals μ and ν, we have

$$\int_{\mathcal{X}\times\mathcal{Y}} c(x, y)\, d\pi_0(x, y) = \inf_{\pi\in\Pi(\{\mu\},\{\nu\})} \int_{\mathcal{X}\times\mathcal{Y}} c(x, y)\, d\pi(x, y),$$

then we say that π_0 is an optimal transference plan. In this case,

$$\left[\int_{\mathcal{X}\times\mathcal{Y}} c(x, y)\, d\pi_0(x, y)\right]^{\alpha}$$

is called the α-optimal cost between μ and ν. (We include the parameter α so that we can allow the Wasserstein distance in this framework.)

Definition 3. The Monge-Kantorovich problem for spaces $\mathcal{X}, \mathcal{Y}$ is: given a cost function c, and measures $\mu \in P(\mathcal{X})$ and $\nu \in P(\mathcal{Y})$, find an optimal transference plan between μ and ν.

The Monge-Kantorovich problem is solvable in very general settings.

Theorem 4 (Theorem 4.1 of [12])**.** *Let $(\mathcal{X}, \mu)$ and $(\mathcal{Y}, \nu)$ be two Polish probability spaces. Let $a : \mathcal{X} \to \mathbb{R} \cup \{-\infty\}$ and $b : \mathcal{Y} \to \mathbb{R} \cup \{-\infty\}$ be upper semicontinuous functions such that $a \in L^1(\mu)$ and $b \in L^1(\nu)$. Suppose that $c : \mathcal{X}\times\mathcal{Y} \to \mathbb{R}\cup\{+\infty\}$ is a lower semicontinuous function such that $c(x, y) \geq a(x)+b(y)$ for all $x \in \mathcal{X}, y \in \mathcal{Y}$. Then there exists an optimal transference plan between μ and ν.*

An important special case occurs when $\mathcal{X} = \mathcal{Y}$ is a Polish metric space with distance function d and $c = d^p$ for some $p \in [1, \infty)$. The $1/p$-optimal cost is then the Wasserstein distance of order p, given by

$$W_p(\nu, \mu) = \left(\inf_{\pi\in\Pi(\{\mu\},\{\nu\})} \int_{\mathcal{X}\times\mathcal{X}} d^p(x, y)\, d\pi(x, y)\right)^{1/p}.$$

For any arbitrary $x_0 \in \mathcal{X}$, we define

$$P_p(\mathcal{X}) = \left\{\mu \in P(\mathcal{X}) \mid \int_{\mathcal{X}} d^p(x, x_0)\, d\mu(x) < \infty\right\}.$$

W_p is then a metric on $P_p(\mathcal{X})$. Furthermore, if $\mathcal{X}$ is a complete, separable and locally compact length space and $p > 1$, then $P_p(\mathcal{X})$ is a geodesic space (see Chapters 6 and 7 of [12]).

We wish to solve the Steiner problem in a probability space $\mathcal{P} \subset P(\mathcal{X})$. It is useful however to first consider a hybrid Monge-Kantorovich-Steiner problem. In a geodesic space, the solutions of Steiner problems map edges to geodesics, so this becomes equivalent to minimizing a certain sum of distances. In a more general metric space, this correspondence need not hold, but one could view a solution which minimizes the sum of distances as a weak solution to the Steiner problem. Accordingly, the α-optimal cost of a network $\phi : G \to P(\mathcal{X})$ is defined as the sum over all edges $\{v_i, v_j\}$ of G of the α-optimal cost between $\phi(v_i)$ and $\phi(v_j)$. The optimal costs are achieved by solutions of the Monge-Kantorovich problem, so we consider the following Monge-Kantorovich-Steiner problems:

Definition 5. The parameterized Monge-Kantorovich-Steiner problem for a graph G, a metric space $\mathcal{X}$, $\alpha > 0$ and a cost function c is: given k vertices $v_1, \ldots, v_k \in G$ and k probability measures $\mu_1, \ldots, \mu_k \in \mathcal{P} \subset P(\mathcal{X})$, find a network of minimal α-optimal cost in the set of all networks in $\mathcal{P}$ with parametric graph G that send v_i to μ_i for $1 \leq i \leq k$.

Definition 6. The general Monge-Kantorovich-Steiner problem for a metric space $\mathcal{X}$, $\alpha > 0$ and a cost function c is: given a subset $\mathcal{P} \subset P(\mathcal{X})$ and k probability measures $\mu_1, \ldots, \mu_k \in \mathcal{P}$, find a network of minimal α-optimal cost in the set of all networks $\phi : G \to \mathcal{P}$ such that G is a connected graph and $\mu_1, \ldots, \mu_k$ are contained in the image of the vertex set.

If for some $\pi \in \Pi(P(\mathcal{X}), P(\mathcal{Y}))$ with marginals $\mu \in P(\mathcal{X}), \nu \in P(\mathcal{Y})$ there exists a measurable map $T : \mathcal{X} \to \mathcal{Y}$ such that $\pi = (\mathrm{id}, T)_{\#}\mu$, then π is said to be deterministic, and T is called the transport map. The classical Monge problem is to look for an optimal deterministic transference plan. We may thus consider the following two problems as well:

Definition 7. The parameterized Monge-Steiner problem for a graph G, a metric space $\mathcal{X}$, $\alpha > 0$ and a cost function c is: given k vertices $v_1, \ldots, v_k \in G$ and k probability measures $\mu_1, \ldots, \mu_k \in \mathcal{P} \subset P(\mathcal{X})$, find a network of minimal α-optimal cost in the set of all networks in $\mathcal{P}$ with parametric graph G that send v_i to μ_i for $1 \leq i \leq k$, and each α-optimal cost is achieved by a deterministic transference plan.

Definition 8. The general Monge-Steiner problem for a metric space $\mathcal{X}$, $\alpha > 0$ and a cost function c is: given k probability measures $\mu_1, \ldots, \mu_k \in \mathcal{P} \subset P(\mathcal{X})$, find a network of minimal α-optimal cost in the set of all networks $\phi : G \to \mathcal{P}$ such that G is a connected graph and $\mu_1, \ldots, \mu_k$ are contained in the image of the vertex set, and each α-optimal cost is achieved by a deterministic transference plan.

Solvability of the Monge problem holds far less generally than solvability of the Monge-Kantorovich problem. For example, a transport map can only send a Dirac mass to another Dirac mass. We will therefore see the strongest results for the Monge-Kantorovich-Steiner problems and the classical Steiner problems.

3. Existence of solutions

3.1. **The direct method of minimization.** In the existence proof for a solution of the classical Steiner problem on $\mathbb{R}^n$, one shows that any minimizing sequence

must eventually remain in a bounded set and applies a compactness argument. Unfortunately, bounded sets are no longer precompact in a probability space $P(\mathcal{X})$. We must therefore use another criterion for precompactness, which is given by the following:

Theorem 9 (Prokhorov's Theorem [11]). *If $\mathcal{X}$ is a Polish space, then a set $\mathcal{P} \subset P(\mathcal{X})$ is precompact for the weak topology if and only if it is tight, i.e. for any $\epsilon \geq 0$ there exists a compact set $K_\epsilon \subset \mathcal{X}$ such that $\mu(\mathcal{X} \setminus K_\epsilon) \leq \epsilon$ if $\mu \in \mathcal{P}$.*

We now show that the direct method works if a tightness bound is assumed.

Proposition 10 (cf. Theorem 4). *Let $\mathcal{X}$ be a Polish space, and suppose $\alpha > 0$ and $c : \mathcal{X} \times \mathcal{X} \to [0, +\infty]$ is a lower semicontinuous cost function. Fix boundary points $\mu_1, \ldots, \mu_k \in P(\mathcal{X})$ and let $\lambda^{1,1}, \lambda^{1,2}, \ldots, \lambda^{k+l,k+l} \in [0, \infty)$ be fixed edge coefficients. Let $\Phi : (P(\mathcal{X})^l) \to \mathbb{R}$ be the Monge-Kantorovich-Steiner functional*

$$\Phi(\mu_{k+1}, \ldots, \mu_{k+l}) = \sum_{i,j=1}^{k+l} \lambda^{i,j} \left[\inf_{\pi \in \Pi(\{\mu_i\},\{\mu_j\})} \int_{\mathcal{X}\times\mathcal{X}} c(x, y)\, d\pi(x, y) \right]^\alpha .$$

If $\mathcal{P} \subset P(\mathcal{X})$ is tight, then there exist $\nu_1, \ldots, \nu_l \in \overline{\mathcal{P}}$ such that

$$\Phi(\nu_1, \ldots, \nu_l) = \inf_{\mathcal{P}^l} \Phi.$$

Proof. Define $F : [\Pi(P(\mathcal{X}), P(\mathcal{X}))]^{(k+l)^2} \to \mathbb{R}$ by

$$F(\pi_{1,1}, \ldots, \pi_{k+l,k+l}) = \sum_{i,j=1}^{k+l} \lambda^{i,j} \left[\int_{\mathcal{X}\times\mathcal{X}} c(x, y)\, d\pi_{i,j}(x, y) \right]^\alpha ,$$

and note that

$$\inf_{\mathcal{P}^l} \Phi = \inf_{\mathcal{Q}} F,$$

where $\mathcal{Q} \subset [\Pi(P(\mathcal{X}), P(\mathcal{X}))]^{(k+l)^2}$ is the set of $(k + l)^2$-tuples of transference plans with marginals matching fixed points $\mu_1, \ldots, \mu_k$ where appropriate, with marginals in $\mathcal{P}$ otherwise and with internally consistent marginals. Since $\mathcal{X}$ is Polish, each set $\{\mu_i\}$ is tight. Also, $\mathcal{P}$ is tight by assumption. Thus $\mathcal{P}, \{\mu_1\}, \ldots, \{\mu_k\}$ is a finite collection of precompact sets by Prokhorov's Theorem and $\mathcal{P}' = \mathcal{P} \cup \{\mu_1, \ldots, \mu_k\}$ is precompact and tight. $\mathcal{Q} \subset [\Pi(\mathcal{P}', \mathcal{P}')]^{(k+l)^2}$ by construction. $\Pi(\mathcal{P}', \mathcal{P}')$ is tight, hence precompact, by the following lemma:

Lemma 11 (Lemma 4.4 of [12]). *Let $\mathcal{X}$ and $\mathcal{Y}$ be two Polish spaces. Let $\mathcal{P} \subset P(\mathcal{X})$ and $\mathcal{Q} \subset P(\mathcal{Y})$ be tight subsets. Then the set $\Pi(\mathcal{P}, \mathcal{Q})$ of all transference plans whose marginals lie in $\mathcal{P}$ and $\mathcal{Q}$, respectively, is itself tight in $P(\mathcal{X} \times \mathcal{Y})$.*

Let $\{\pi^n_{i,j}\} \subset \mathcal{Q}$ be an F-minimizing sequence. Since $\{\pi^n_{1,1}\} \subset \Pi(\mathcal{P}', \mathcal{P}')$, by taking a subsequence we may assume $\pi^n_{1,1}$ converges weakly to some $\pi_{1,1} \in \overline{\Pi(\mathcal{P}', \mathcal{P}')}$. Taking a subsequence $(k + l)^2 - 1$ more times, we may even assume that for all i, j, $\pi^n_{i,j}$ converges weakly to some $\pi_{i,j} \in \overline{\Pi(\mathcal{P}', \mathcal{P}')}$. $\mathcal{Q}$ is clearly closed, so $(\pi_{1,1}, \ldots, \pi_{k+l,k+l}) \in \mathcal{Q}$. We cite another lemma to show that $(\pi_{1,1}, \ldots, \pi_{k+l,k+l})$ is an F-minimizer.

Lemma 12 (Lemma 4.3 of [12]). *Let $\mathcal{X}$ and $\mathcal{Y}$ be two Polish spaces, and $c : \mathcal{X}\times\mathcal{Y} \to \mathbb{R} \cup \{+\infty\}$ a lower semicontinuous cost function. Let $h : \mathcal{X} \times \mathcal{Y} \to \mathbb{R} \cup \{-\infty\}$ be an upper semicontinuous function such that $c \geq h$. Let $(\pi_k)_{k\in\mathbb{N}}$ be a sequence in*

$P(\mathcal{X}\times\mathcal{Y})$, converging weakly to some $\pi \in P(\mathcal{X}\times\mathcal{Y})$, in such a way that $h \in L^1(\pi_k)$, $h \in L^1(\pi)$, and

$$\int_{\mathcal{X}\times\mathcal{Y}} h\, d\pi_k \to \int_{\mathcal{X}\times\mathcal{Y}} h\, d\pi.$$

Then

$$\int_{\mathcal{X}\times\mathcal{Y}} c\, d\pi \leq \liminf_{k\to\infty} \int_{\mathcal{X}\times\mathcal{Y}} c\, d\pi_k.$$

In particular, if $c \geq 0$, then $F : \pi \to \int c\pi$ is lower semicontinuous on $P(\mathcal{X}\times\mathcal{Y})$, equipped with the topology of weak convergence.

Applying Lemma 12 for $h \equiv 0$,

$$\inf_{\mathcal{Q}} F \leq F(\pi_{1,1}, \ldots, \pi_{k+l,k+l}) \leq \liminf_{n\to\infty} F(\pi^n_{1,1}, \ldots, \pi^n_{k+l,k+l}) = \inf_{\mathcal{Q}} F$$

and $(\pi_{1,1}, \ldots, \pi_{k+l,k+l})$ is F-minimizing. Taking marginals yields the desired Φ-minimizer. □

3.2. **Tightness from CAT(0).** Proposition 10 shows that the direct method for solving the Monge-Kantorovich-Steiner problem works as long as one has an a priori tightness estimate. We will now see how the geometry of the base space can provide this tightness estimate. We recall some basic notions of metric geometry.

Definition 13. A shortest path in a length space is a rectifiable curve such that the distance between two points on the curve is equal to the length of the corresponding segment of the curve.

Definition 14. A space of nonpositive Alexandrov curvature is a length space which can be covered by a family of open sets $\{V_i\}$ such that for each V_i:

(1) There exists a shortest path in V_i connecting any two points in V_i.
(2) For any $a, b, c \in V_i$ with each pair of points connected by a shortest path and any point d in the shortest path ac, let $\Delta\tilde{a}\tilde{b}\tilde{c}$ be the comparison triangle for Δabc in $\mathbb{R}^2$, i.e. $|ab| = |\tilde{a}\tilde{b}|$, $|ac| = |\tilde{a}\tilde{c}|$ and $|bc| = |\tilde{b}\tilde{c}|$, and let $\tilde{d}$ be the point in $\tilde{a}\tilde{c}$ such that $|ad| = |\tilde{a}\tilde{d}|$. Then $|bd| \leq |\tilde{b}\tilde{d}|$. (Intuitively, Δabc is skinnier than $\Delta\tilde{a}\tilde{b}\tilde{c}$.)

This notion is equivalent to nonpositive sectional curvature in the setting of Riemannian manifolds.

Definition 15. A Hadamard space (or complete CAT(0) space) is a simply connected complete space of nonpositive Alexandrov curvature.

Hadamard spaces $\mathcal{X}$ are important because the curvature conditions hold for all triangles in $\mathcal{X}$, not just small triangles [2]. In our discussion, this allows us to ensure the convexity of the distance functions on $\mathcal{X}$, which will give us good control of convex hulls. In particular, $\mathcal{X}$ is locally convex. We will need $\mathcal{X}$ to be separable, so we note that this is always true for $\mathcal{X}$ of finite Hausdorff dimension.

In order to assure that the Monge-Kantorovich-Steiner problem is aware of our geometric assumptions, we will also assume that the cost function is based on the distance function.

Lemma 16. *Let $\mathcal{X}$ be a separable Hadamard space, $\alpha > 0$ and $c : \mathcal{X}\times\mathcal{X} \to [0, +\infty]$ a lower semicontinuous cost function of the form $c = \varphi \circ d$, where φ is a monotone nondecreasing function and d is the distance in $\mathcal{X}$. Let $\mu_1, \ldots, \mu_k \in P(\mathcal{X})$ be fixed*

boundary points and let $\lambda^{1,1}, \lambda^{1,2}, \ldots, \lambda^{k+l,k+l} \in [0,\infty)$ *be fixed edge coefficients. Suppose that* $\mu_1, \ldots, \mu_k$ *all have compact support. Let* $\Phi : (P(\mathcal{X})^l) \to \mathbb{R}$ *be the Monge-Kantorovich-Steiner functional*

$$\Phi(\mu_{k+1}, \ldots, \mu_{k+l}) = \sum_{i,j=1}^{k+l} \lambda^{i,j} \left[\inf_{\pi \in \Pi(\{\mu_i\},\{\mu_j\})} \int_{\mathcal{X}\times\mathcal{X}} c(x,y)\, d\pi(x,y) \right]^{\alpha}.$$

Then there exist $\nu_1, \ldots, \nu_l \in P(\mathcal{X})$ *(with compact support) such that*

$$\Phi(\nu_1, \ldots, \nu_l) = \inf_{(P(\mathcal{X}))^l} \Phi.$$

Proof. Let K be a large enough compact set so that $\mu_1, \ldots, \mu_k \in P(K)$. Since $\mathcal{X}$ is locally convex, $H = \overline{\mathrm{co}\,(K)}$ is compact by an exercise in [1].

As shown in [2], there exists a unique orthogonal projection map $\mathrm{proj}_H : \mathcal{X} \to H$ which is a distance nonincreasing retraction of $\mathcal{X}$ onto H. So given $\pi \in \Pi(\{\mu_i\}, P(\mathcal{X}))$, we have $(\mathrm{proj}_H \times \mathrm{proj}_H)_{\#}\pi \in \Pi(\{\mu_i\}, P(H))$ and

$$\int_{\mathcal{X}\times\mathcal{X}} c(x,y)\, d(\mathrm{proj}_H \times \mathrm{proj}_H)_{\#}\pi(x,y) \le \int_{\mathcal{X}\times\mathcal{X}} c(x,y)\, d\pi(x,y).$$

Thus $\inf_{(P(\mathcal{X}))^l} \Phi = \inf_{(P(H))^l} \Phi$.

Since H is compact, choosing $K_\epsilon = H$ yields that $P(H)$ is tight. $P(H)$ is also closed, so Proposition 10 implies that there exists $\nu_1, \ldots, \nu_l \in P(H)$ such that

$$\Phi(\nu_1, \ldots, \nu_l) = \inf_{(P(\mathcal{X}))^l} \Phi = \inf_{(P(H))^l} \Phi. \qquad \square$$

Theorem 17. *If* $\mathcal{X}$ *is a separable Hadamard space or a compact space, then the parameterized Monge-Kantorovich-Steiner problems are solvable for compactly supported boundary data with a lower semicontinuous cost function* $c : \mathcal{X}\times\mathcal{X} \to [0,+\infty]$ *of the form* $c = \varphi \circ d$*, where* φ *is a monotone nondecreasing function and* d *is the distance in* $\mathcal{X}$*.*

If $c = d^p$ *and* $\alpha = 1/p$ *for some* $p \in [1,\infty)$*, then the general Monge-Kantorovich-Steiner problem is solvable as well.*

Furthermore, the minimizing configuration in $P(\mathcal{X})$ *produced in both cases consists of measures with compact support.*

Proof. For the parameterized problem with graph G, label the boundary vertices as $v_1, \ldots, v_k$ and the remaining vertices as $v_{k+1}, \ldots, v_{k+l}$. Let $\lambda^{i,j}$ be $1/2$ if v_i and v_j are connected by an edge in G and 0 otherwise. Proposition 10 or Lemma 16 then provides the solution.

For the general problem, we follow the classical argument (see [6]). Let $\mathcal{K}$ be the set of finite connected graphs G with k distinguished boundary vertices, let Φ_G denote the corresponding Monge-Kantorovich-Steiner functional for each $G \in \mathcal{K}$. We wish to achieve

$$\inf_{G\in\mathcal{K}} \inf \Phi_G = \inf_{G\in\mathcal{K}} \Phi_G(\nu_1^G, \ldots, \nu_{|G|-k}^G),$$

where $(\nu_1^G, \ldots, \nu_{|G|-k}^G) \in (P(\mathcal{X}))^{|G|-k}$ is the solution of the parameterized problem. (Here $|G|$ denotes the number of vertices of G.)

Let $\mathcal{K}'$ be the subset of $\mathcal{K}$ consisting of trees, and let $\mathcal{K}''$ be the subset of $\mathcal{K}'$ where all vertices not in the distinguished boundary set have degree at least three.

Since $\mathcal{K}''$ is a finite set, the infimum over $\mathcal{K}''$ is trivially achieved. It remains to show that

$$\inf_{G\in\mathcal{K}} \inf \Phi_G = \inf_{G\in\mathcal{K}'} \inf \Phi_G = \inf_{G\in\mathcal{K}''} \inf \Phi_G.$$

For any $G \in \mathcal{K}$ and any $(\nu_1, \ldots, \nu_{|G|-k}) \in (P(\mathcal{X}))^{|G|-k}$, we may take a tree $G' \in \mathcal{K}'$ such that G' is a subgraph of G and $\Phi_{G'}(\nu_1, \ldots, \nu_{|G|-k}) \leq \Phi_G(\nu_1, \ldots, \nu_{|G|-k})$ since

$$\lambda^{i,j} W_p(\mu_i, \mu_j)$$

is always nonnegative and we are only possibly setting some of the $\lambda^{i,j}$ to zero. So

$$\inf_{G\in\mathcal{K}} \inf \Phi_G = \inf_{G\in\mathcal{K}'} \inf \Phi_G.$$

Similarly, for any $G' \in \mathcal{K}'$, removing interior vertices of degree one will keep us in $\mathcal{K}'$ and will not increase $\Phi_{G'}(\nu_1, \ldots, \nu_{|G|-k})$. We may also replace any interior vertex of degree two by an edge between its neighbors to obtain a graph $G'' \in \mathcal{K}''$. It follows from the triangle inequality for Wasserstein distances that $\Phi_{G''}(\nu_1, \ldots, \nu_{|G|-k}) \leq \Phi_{G'}(\nu_1, \ldots, \nu_{|G|-k})$. Thus

$$\inf_{G\in\mathcal{K}} \inf \Phi_G = \min_{G\in\mathcal{K}''} \Phi_G(\nu_1^G, \ldots, \nu_{|G|-k}^G).$$

The compactness of supports also follows from Lemma 16. □

We note that in the $c = d^2$ and $\alpha = 1/2$ case, if $\mathcal{X}$ is a Riemannian manifold, G is a star and the boundary data is absolutely continuous, the optimal couplings given by Theorem 4 may be taken to be deterministic, i.e. the Monge problem is solvable (Theorem 10.40 of [12]). Thus Theorem 17 gives:

Corollary 18. *If M is a Riemannian manifold, M is compact or has nonpositive sectional curvature, G is a star, $c = d^2$, $\alpha = 1/2$ and each fixed boundary point μ_i is compactly supported and absolutely continuous with respect to the volume measure of M, then the parameterized Monge-Steiner problem is solvable.*

We may say quite a bit more for the classical Steiner problem.

Corollary 19. *If $\mathcal{X}$ is a compact space and $p > 1$, then the parameterized and general Steiner problems are solvable for arbitrary boundary data on $(P_p(\mathcal{X}), W_p)$. If $\mathcal{X}$ is a separable locally compact Hadamard space and $p > 1$, then the parameterized and general Steiner problems are solvable for compactly supported boundary data on $(P_p(\mathcal{X}), W_p)$.*

Proof. In either case, we have Monge-Kantorovich-Steiner solutions for the cost function d^p with $\alpha = 1/p$. The associated cost functional is W_p, which metrizes $P_p(\mathcal{X})$ as a geodesic space. Thus the adjacent measures may be joined by geodesics, forming a minimal network. □

Note that the geodesics in the Steiner network may be assumed to stay in the set $P_2(H)$ of measures of compact support, as can be seen by considering the geodesic problem as a parametric Steiner problem.

The curvature assumption on $\mathcal{X}$ was only used above to control properties of convex hulls. In particular, the arguments carry through for any complete separable length space $\mathcal{X}$ satisfying:

(1) If $K \subset \mathcal{X}$ is compact, then $\overline{\mathrm{co}\,(K)}$ is compact.

(2) There exists a compact set $K_0 \subset \mathcal{X}$ such that if $K \subset M$ is a compact set with $K_0 \subset K$, then there exists a distance nonincreasing retraction $\mathrm{proj}_H : \mathcal{X} \to H$, where H is compact and $K \subset H$.

As mentioned above, condition (1) holds for any complete metric space which is locally convex; thus condition (1) holds for Riemannian manifolds (by the existence of strongly convex neighborhoods [4]) and for Alexandrov spaces of curvature bounded above (see Proposition II.1.4 of [2]). If $\mathcal{X} = \mathcal{S} \times \mathcal{N}$, where $\mathcal{X}$ has the product metric, $\mathcal{S}$ is compact and $\mathcal{N}$ is a Hadamard space, then it is easy to show that condition (2) holds for $K_0 = \emptyset$ and $H = \mathcal{S} \times \overline{\mathrm{co}\,(\mathrm{proj}_{\mathcal{N}} K)}$ by setting

$$\mathrm{proj}_H = \mathrm{id} \times \mathrm{proj}_{\overline{\mathrm{co}(\mathrm{proj}_{\mathcal{N}} K)}},$$

where the second component function is the orthogonal projection in $\mathcal{N}$. We therefore have:

Proposition 20. *If $\mathcal{X} = \mathcal{S} \times \mathcal{N}$, where $\mathcal{X}$ is a complete, separable, locally convex length space endowed with the product metric, $\mathcal{S}$ is compact, $\mathcal{N}$ is a Hadamard space and $p > 1$, then both the parameterized and general versions of the Steiner problem are solvable in $(P_p(\mathcal{X}), W_p)$ for arbitrary boundary data of compact support.*

As a particular case, the hypothesis of Proposition 20 is satisfied for any Riemannian manifold M which splits isometrically as $M = S \times \mathbb{R}^n$ with S compact. One may think of this condition as a strong version of the Soul Theorem of Cheeger and Gromoll.

3.3. **Proof of Main Theorem.** We also assumed compact support for the boundary data in order to say that the supports of the minimizing sequence could be assumed to be compact. For general boundary data $\mu_1, \ldots, \mu_k$, one only has tightness of the set $\{\mu_1, \ldots, \mu_k\}$. We will thus approximate the solution for boundary data $\mu_1, \ldots, \mu_k$ by a sequence of solutions for compact boundary data $\mu_1^n, \ldots, \mu_k^n$ and show that the approximate solutions converge to a solution of the original problem.

Theorem 21. *Suppose $\mathcal{X} = \mathcal{S} \times \mathcal{N}$, where $\mathcal{X}$ is a complete, separable, locally compact, locally convex length space endowed with the product metric, $\mathcal{S}$ is compact, $\mathcal{N}$ is a Hadamard space and $p > 1$. Then the parameterized and general versions of the Monge-Kantorovich-Steiner problem and the Steiner problem are solvable in $(P_p(\mathcal{X}), W_p)$ for arbitrary boundary data. (For $p = 1$, the Monge-Kantorovich-Steiner problems are solvable.)*

Proof. First consider the G-parameterized problem where each vertex of G is adjacent to the boundary. Let $\mu_1, \ldots, \mu_k \in P_p(\mathcal{X})$ denote the boundary data and let $\Phi : (P_p(\mathcal{X}))^l \to \mathbb{R}$ be the Monge-Kantorovich-Steiner functional. Since $\{\mu_1, \ldots, \mu_k\}$ is tight, we may choose $r_n > 0$ such that $\mu_i(\mathcal{X} \setminus B_{r_n}) \leq 1/n$ for all i, n, where B_{r_n} is the intrinsic closed ball of radius r_n about some fixed base point x_0. In particular, $\mu_i(B_{r_n}) \geq (n-1)/n$. B_{r_n} is compact since $\mathcal{X}$ is complete and locally compact. Define the cutoff measures

$$\mu_i^n = \frac{\mu_i|_{B_{r_n}}}{\mu_i|_{B_{r_n}}(\mathcal{X})} = \frac{\mu_i|_{B_{r_n}}}{\mu_i(B_{r_n})}.$$

Note that for $m \geq 2$, we have the tightness estimate

$$\mu_i^m(\mathcal{X} \setminus B_{r_n}) = \frac{\mu_i|_{B_{r_m}}(\mathcal{X} \setminus B_{r_n})}{\mu_i(B_{r_m})} \leq \frac{\mu_i(\mathcal{X} \setminus B_{r_n})}{\mu_i(B_{r_m})} \leq \frac{m}{m-1}\frac{1}{n} \leq \frac{2}{n}.$$

By Theorem 6.9 in [12], a sequence μ^m converges to μ in the W_p metric if and only if $\mu^m \to \mu$ weakly and

$$\limsup_{m\to\infty} \int d^p(x,x_0)\, d\mu^m(x) \le \int d^p(x,x_0)\, d\mu(x).$$

Thus the weak convergence of μ_i^m to μ_i and the bound

$$\int d^p(x,x_0)\, d\mu_i^m(x) = W_p^p(\mu_i^m, \delta_{x_0}) \le W_p^p(\mu_i, \delta_{x_0}) = \int d^p(x,x_0)\, d\mu_i(x)$$

imply that $\mu_i^m \to \mu_i$ in W_p.

Let Φ^m denote the Monge-Kantorovich-Steiner functional for the boundary data $\mu_1^m, \dots, \mu_k^m$. Since

$$b := \Phi(\delta_{x_0}, \dots, \delta_{x_0}) \ge \Phi^m(\delta_{x_0}, \dots, \delta_{x_0}),$$

b gives an upper bound on the infima for Φ and Φ^m.

Let $\nu_1^m, \dots, \nu_l^m$ solve the G-parameterized problem for the compact boundary data $\mu_1^m, \dots, \mu_k^m$. We now show that $\{\nu_1^m, \dots, \nu_l^m\}$ is tight. Let

$$\psi(n) = (b+1)\left(\frac{n}{2}\right)^{1/p} + 2r_n.$$

Suppose that there exists a pair (j,m) such that for some n,

$$\nu_j^m(\mathcal{X} \setminus B_{\psi(n)}) > \frac{4}{n}.$$

Choose i such that ν_j^m and μ_i^m are G-adjacent. Since $\mu_i^m(\mathcal{X} \setminus B_{r_n}) < 2/n$, transporting from ν_j^m to μ_i^m must move at least $2/n$ of the mass from outside $B_{\psi(n)}$ to inside B_{r_n}. More precisely, if $\pi \in \Pi(\{\nu_j^m\}, \{\mu_i^m\})$, then

$$\pi((\mathcal{X} \setminus B_{\psi(n)}) \times B_{r_n}) + \pi((\mathcal{X} \setminus B_{\psi(n)}) \times (\mathcal{X} \setminus B_{r_n})) = \nu_j^m(\mathcal{X} \setminus B_{\psi(n)}) > \frac{4}{n},$$

and similarly

$$\pi((\mathcal{X} \setminus B_{\psi(n)}) \times B_{r_n}) + \pi(B_{\psi(n)} \times B_{r_n}) = 1 - \mu_i^m(\mathcal{X} \setminus B_{\psi(n)}) > 1 - \frac{2}{n}.$$

Adding inequalities we find

$$\begin{aligned}
1 + \frac{2}{n} &< 2\pi((\mathcal{X} \setminus B_{\psi(n)}) \times B_{r_n}) + \pi(B_{\psi(n)} \times B_{r_n}) \\
&\quad + \pi((\mathcal{X} \setminus B_{\psi(n)}) \times (\mathcal{X} \setminus B_{r_n})) \\
&\le 2\pi((\mathcal{X} \setminus B_{\psi(n)}) \times B_{r_n}) + \pi(B_{\psi(n)} \times B_{r_n}) \\
&\quad + \pi((\mathcal{X} \setminus B_{\psi(n)}) \times (\mathcal{X} \setminus B_{r_n})) + \pi(B_{\psi(n)} \times (\mathcal{X} \setminus B_{r_n})) \\
&= \pi((\mathcal{X} \setminus B_{\psi(n)}) \times B_{r_n}) + \pi(\mathcal{X} \times \mathcal{X}) \\
&= \pi((\mathcal{X} \setminus B_{\psi(n)}) \times B_{r_n}) + 1
\end{aligned}$$

since $\pi(\mathcal{X}\times\mathcal{X})=1$, and therefore $\pi((\mathcal{X}\setminus B_{\psi(n)})\times B_{r_n})>2/n$. Thus

$$\begin{aligned}\int_{\mathcal{X}\times\mathcal{X}} d^p(x,y)\,d\pi(x,y) &\geq \int_{(\mathcal{X}\setminus B_{\psi(n)})\times B_{r_n}} d^p(x,y)\,d\pi(x,y)\\ &\geq \int_{(\mathcal{X}\setminus B_{\psi(n)})\times B_{r_n}} (\psi(n)-r_n)^p\,d\pi(x,y)\\ &= (\psi(n)-r_n)^p\pi((\mathcal{X}\setminus B_{\psi(n)})\times B_{r_n})\\ &> \frac{2}{n}(\psi(n)-r_n)^p,\end{aligned}$$

and we see that

$$W_p(\nu_j^m,\mu_i^m)\geq\left(\frac{2}{n}\right)^{1/p}(\psi(n)-r_n)\geq b+1.$$

In particular, $\Phi^m(\nu_1^m,\dots,\nu_l^m)>b$, which contradicts the minimality of Φ^m at $(\nu_1^m,\dots,\nu_l^m)$. Therefore

$$\nu_j^m(\mathcal{X}\setminus B_{\psi(n)})\leq\frac{4}{n}$$

and $\{\nu_1^m,\dots,\nu_l^m\}$ is tight.

For general graphs G, we may inductively show tightness of $\{\nu_1^m,\dots,\nu_l^m\}$ for vertices $k+1$ edges away from the boundary by assuming tightness of such sequences for vertices k edges away from the boundary by the above argument. Since G is a finite graph, we cover all vertices in a finite number of iterations of the argument.

Taking subsequences, we may assume that for all j, $\nu_j^m\to\nu_j$ weakly. Since $\mu_i^m\to\mu_i$ weakly as well, we may use Lemma 11 to take subsequences and assume for all i,j that the optimal transference plans $\pi_{i,j}^m$ converge weakly to some $\hat{\pi}_{i,j}$ with the appropriate marginals.

Let $\epsilon>0$. Since $\mu_i^m\to\mu_i$ in W_p for all i, $\exists M$ such that $\forall m\geq M$,

$$\|\Phi^m-\Phi\|_{L^\infty((P_p(\mathcal{X}))^l)}\leq\epsilon.$$

In particular,

$$|\inf_{(P_p(\mathcal{X}))^l}\Phi^m-\inf_{(P_p(\mathcal{X}))^l}\Phi|\leq\epsilon.$$

As in the proof of Proposition 10, consider $F:[\Pi(P_p(\mathcal{X}),P_p(\mathcal{X}))]^{(k+l)^2}\to\mathbb{R}$ defined by

$$F(\pi_{1,1},\dots,\pi_{k+l,k+l})=\sum_{i,j=1}^{k+l}\lambda^{i,j}\left[\int_{\mathcal{X}\times\mathcal{X}} d^p(x,y)\,d\pi_{i,j}(x,y)\right]^{(1/p)},$$

where again $\lambda^{i,j}=1/2$ for edges of G and $\lambda^{i,j}=0$ otherwise. By construction,

$$F(\pi_{1,1}^m,\dots,\pi_{k+l,k+l}^m)=\Phi^m(\nu_1^m,\dots,\nu_l^m)=\inf_{(P_p(\mathcal{X}))^l}\Phi^m.$$

By Lemma 12, F is lower semicontinuous and

$$\begin{aligned}F(\hat{\pi}_{1,1},\dots,\hat{\pi}_{k+l,k+l}) &\leq \liminf_{m\to\infty}F(\pi_{1,1}^m,\dots,\pi_{k+l,k+l}^m)\\ &= \liminf_{m\to\infty}\inf_{(P_p(\mathcal{X}))^l}\Phi^m\\ &\leq\left(\inf_{(P_p(\mathcal{X}))^l}\Phi\right)+\epsilon.\end{aligned}$$

Sending $\epsilon \to 0$,

$$F(\hat{\pi}_{1,1}, \dots, \hat{\pi}_{k+l,k+l}) = \inf_{(P_p(\mathcal{X}))^l} \Phi,$$

so the $\hat{\pi}_{i,j}$ are in fact optimal transference plans and

$$\Phi(\nu_1, \dots, \nu_l) = F(\hat{\pi}_{1,1}, \dots, \hat{\pi}_{k+l,k+l}) = \inf_{(P_p(\mathcal{X}))^l} \Phi.$$

The solution of the general problem follows as in the proof of Theorem 17. □

4. Structure of Steiner trees in $(P_2(M), W_2)$

We now restrict our attention to the general Steiner problem on $(P_2(M), W_2)$ and investigate how the geometry of M can force structure on the parametric graph of a solution. We must first recall some notions of metric geometry.

Given three distinct points x, y, z in a length space $\mathcal{Y}$, the comparison angle $\tilde{\angle} xyz$ is defined as the corresponding angle in the triangle in $\mathbb{R}^2$ of sides $d(x,y), d(x,z), d(y,z)$. Explicitly,

$$\tilde{\angle} xyz = \arccos \frac{d^2(x,y) - d^2(x,z) + d^2(y,z)}{2d(x,y)d(y,z)}.$$

If $\beta : [0,\epsilon) \to \mathcal{Y}$ and $\gamma : [0,\epsilon) \to \mathcal{Y}$ are two paths in $\mathcal{Y}$ with $\beta(0) = \gamma(0) = p$, then we define the angle

$$\angle(\alpha, \beta) = \lim_{s,t \to 0} \tilde{\angle}(\alpha(s), p, \beta(t))$$

whenever the limit exists.

Definition 22. A length space $\mathcal{Y}$ is said to have nonnegative Alexandrov curvature if it has a covering by neighborhoods $\{V_i\}$ such that for any two shortest paths $\beta : [0,\epsilon) \to V_i$ and $\gamma : [0,\epsilon) \to V_i$ with $\beta(0) = p$ and $\gamma(0) = p$,

$$\tilde{\angle}(\alpha(s), p, \beta(t))$$

is nonincreasing in both s and t.

There are several equivalent definitions of nonnegative Alexandrov curvature. For instance, it is shown in [3] that for a locally compact length space $\mathcal{Y}$, nonnegative Alexandrov curvature is equivalent to having a covering by neighborhoods $\{V_i\}$ such that for any four distinct points $a, b, c, d \in V_i$,

$$\tilde{\angle} bac + \tilde{\angle} cad + \tilde{\angle} dab \leq 2\pi.$$

The following results of Lott and Villani will allow us to work geometrically on our space of probabilities.

Theorem 23 ([9])**.** *Suppose M is a compact Riemannian manifold with nonnegative sectional curvature. Then for all $\mu_0, \dots, \mu_3 \in P_2(M)$,*

$$\tilde{\angle} \mu_1\mu_0\mu_2 + \tilde{\angle} \mu_2\mu_0\mu_3 + \tilde{\angle} \mu_3\mu_0\mu_1 \leq 2\pi.$$

In particular, $P_2(M)$ has nonnegative Alexandrov curvature.

By passing to limits in the inequality, we obtain

$$\angle \gamma_1\mu_0\gamma_2 + \angle \gamma_2\mu_0\gamma_3 + \angle \gamma_3\mu_0\gamma_1 \leq 2\pi$$

for geodesics γ_i starting at μ_0.

Theorem 24 ([9]). *Suppose M is a compact Riemannian manifold with nonnegative sectional curvature. Then for each absolutely continuous measure $\mu \in P_2(M)$, the tangent cone K_μ of $P_2(M)$ at μ is a Hilbert space, under the inner product generated by angles of geodesics in the space of directions.*

We also recall a first variation formula for Alexandrov spaces.

Lemma 25 ([3]). *Let $\mathcal{Y}$ be a complete, locally compact length space of nonnegative Alexandrov curvature, $p \in \mathcal{Y}$ and $\gamma : [0, \epsilon) \to \mathcal{Y}$ a unit-speed shortest path. Let $l(t) = d(p, \gamma(t))$. Then*

$$\lim_{t \to 0^+} \frac{l(t_i) - l(0)}{t_i} = \min_{\sigma_0} \left[-\cos(\angle \sigma_0 \gamma) \right],$$

where the minimum is taken over all shortest paths from $\gamma(0)$ to p. (In particular, the limit exists and the minimum is achieved.)

We now assume that $M = S \times \mathbb{R}^n$, where S is compact with nonnegative sectional curvature, and $\mu_1, \ldots, \mu_k \in P_2(M)$ have compact support. By Proposition 20, there is a general Steiner solution which, by the proof of Theorem 17, may be represented by a network Γ whose parametric graph G is a tree and all interior vertices have degree at least three. This is known as the canonical representative. We will now show that if an interior vertex does not have degree three, then the corresponding measure is not absolutely continuous.

First, note that for some compact $K \subset M$ containing the supports of $\mu_1, \ldots, \mu_k$, we have that $\Gamma(G) \subset P_2(H)$ for $H = \mathcal{S} \times \overline{\mathrm{co}\,(\mathrm{proj}_{\mathcal{N}} K)}$ as above. Γ is thus trivially a Steiner solution for the restrained general Steiner problem on $P_2(H)$, where we can apply Theorems 23 and 24 and Lemma 25.

By Theorem 23, it suffices to show that the angle between any pair of adjacent geodesics γ_1, γ_2 connected at an absolutely continuous $\mu_0 \in P_2(H)$ is at least $2\pi/3$. Suppose that

$$\angle \gamma_1 \mu_0 \gamma_2 < 2\pi/3.$$

Let e_1, e_2 be the edges corresponding to γ_1, γ_2, and let v be the vertex corresponding to μ_0. Split the vertex v into v_1, v_2 and create a new graph G', where v_1 is incident to exactly e_1, e_2 and a new edge e, and v_2 is incident to e and the remaining edges originally incident to v. There is an obvious graph homomorphism $h : G' \to G$ identifying v_1 and v_2. Let $\Gamma' : G' \to P_2(H)$ denote the network $\Gamma \circ h$. Γ' is clearly a (noncanonical) Steiner network, so it is a global and local minimizer for length.

Let N_1, N_2 be the unit vector representatives of γ_1, γ_2 in the tangent cone K_{μ_0} at μ_0. Since K_{μ_0} is an inner product space, there is a (unit-speed) geodesic $\eta : [0, \epsilon) \to P_2(H)$ with $\eta(0) = \mu_0$ such that the angle between η and $N_1 + N_2$ is arbitrarily small. Let N be the unit vector representative of η and let $l(t)$ be the length of the network Γ'_t given by shifting v_1 to $\eta(t)$. By minimality of Γ',

$$\lim_{t \to 0^+} \frac{l(t_i) - l(0)}{t_i} \geq 0.$$

The edge e maps to $\eta([0,t])$ and thus has length t. The only other lengths changed are the images of e_1, e_2, so by Lemma 25,

$$\begin{aligned} 0 \leq \lim_{t\to 0^+} \frac{l(t_i)-l(0)}{t_i} &= 1 + \min_{\sigma_1}[-\cos(\angle\sigma_1\eta)] + \min_{\sigma_2}[-\cos(\angle\sigma_2\eta)] \\ &\leq 1 - \cos(\angle N_1\eta) - \cos(\angle N_2\eta) \\ &= 1 - \langle N_1, N\rangle - \langle N_2, N\rangle \\ &= 1 - \langle N_1 + N_2, N\rangle \\ &= 1 - \|N_1+N_2\|\cos(\angle(N_1+N_2), N) < 0, \end{aligned}$$

since $\|N_1+N_2\| > 1$ and $\cos(\angle(N_1+N_2), N)$ is arbitrarily close to 1. This contradiction implies that

$$\angle\gamma_1\mu_0\gamma_2 \geq 2\pi/3.$$

Summarizing, we have

Theorem 26. *Suppose M is a Riemannian manifold with isometric splitting $M = S\times\mathbb{R}^n$, where S is compact with nonnegative sectional curvature, and $\mu_1,\dots,\mu_k \in P_2(M)$ have compact support. Then there is a Steiner solution in $P_2(M)$ spanning $\mu_1,\dots,\mu_k$. Furthermore, this solution has a canonical representative $\Gamma : G \to P_2(M)$ such that*

1. *G is a tree.*
2. *Vertices in G not mapped to $\mu_1,\dots,\mu_k$ have degree at least three.*
3. *For any vertex v in $G\setminus\Gamma^{-1}(\{\mu_1,\dots,\mu_k\})$, if $\Gamma(v)$ is absolutely continuous with respect to the volume measure, then the degree of v is three and all pairs of geodesics in $\Gamma(G)$ meeting at $\Gamma(v)$ do so with an angle of $2\pi/3$.*

The method of proof for Theorem 26 also applies to Steiner trees in locally compact, finite-dimensional, nonnegatively curved Alexandrov space; one must simply replace the notion of absolute continuity of a measure with the notion of being a manifold point. This further illustrates the analogy between measures in $P_2(M)\setminus P_2^{ac}(M)$ and singular points in a finite-dimensional Alexandrov space mentioned in [9].

One may see that the absolute continuity assumption for the vertex v in Theorem 26 is only used to establish the existence of ϵ-almost midpoints in the tangent cone K_v. These ϵ-almost midpoints always exist for finite-dimensional Alexandrov spaces of nonnegative curvature; however, [5] has an infinite-dimensional counterexample.

References

1. Luigi Ambrosio and Paolo Tilli, *Topics on analysis in metric spaces*, Oxford Lecture Series in Mathematics and its Applications, vol. 25, Oxford University Press, Oxford, 2004. MR2039660 (2004k:28001)
2. Martin R. Bridson and André Haefliger, *Metric spaces of non-positive curvature*, Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], vol. 319, Springer-Verlag, Berlin, 1999. MR1744486 (2000k:53038)
3. Dmitri Burago, Yuri Burago, and Sergei Ivanov, *A course in metric geometry*, Graduate Studies in Mathematics, vol. 33, American Mathematical Society, Providence, RI, 2001. MR1835418 (2002e:53053)
4. Manfredo Perdigão do Carmo, *Riemannian geometry*, Mathematics: Theory & Applications, Birkhäuser Boston Inc., Boston, MA, 1992, Translated from the second Portuguese edition by Francis Flaherty. MR1138207 (92i:53001)
5. Stephanie Halbeisen, *On tangent cones of Alexandrov spaces with curvature bounded below*, Manuscripta Math. **103** (2000), no. 2, 169–182. MR1796313 (2001g:53078)

6. Alexandr O. Ivanov and Alexei A. Tuzhilin, *Minimal networks*, The Steiner problem and its generalizations, CRC Press, Boca Raton, FL, 1994. MR1271779 (95h:05050)
7. L. V. Kantorovich, *On a problem of Monge*, Zap. Nauchn. Sem. S.-Peterburg. Otdel. Mat. Inst. Steklov. (POMI) **312** (2004), no. Teor. Predst. Din. Sist. Komb. i Algoritm. Metody. 11, 15–16. MR2117877
8. ———, *On mass transportation*, Zap. Nauchn. Sem. S.-Peterburg. Otdel. Mat. Inst. Steklov. (POMI) **312** (2004), no. Teor. Predst. Din. Sist. Komb. i Algoritm. Metody. 11, 11–14. MR2117876
9. John Lott and Cedric Villani, *Ricci curvature for metric-measure spaces via optimal transport*, Ann. of Math. (2) **169** (2009), 903–991. MR2480619
10. Chikako Mese and Sumio Yamada, *The parameterized Steiner problem and the singular Plateau problem via energy*, Trans. Amer. Math. Soc. **358** (2006), no. 7, 2875–2895 (electronic). MR2216250 (2007a:58011)
11. Yu. V. Prokhorov, *Convergence of random processes and limit theorems in probability theory*, Teor. Veroyatnost. i Primenen. **1** (1956), 177–238. MR0084896 (18,943b)
12. Cédric Villani, *Optimal transport, old and new*, Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], vol. 338, Springer-Verlag, Berlin, 2009. MR2459454

Department of Mathematics, Johns Hopkins University, Baltimore, Maryland 21218
E-mail address: jdahl@math.jhu.edu

Current address: Department of Mathematics, University of California, Berkeley, California 94720
E-mail address: jdahl@math.berkeley.edu