Canad. J. Math. Vol. **69** (5), 2017 pp. 1087–1108
http://dx.doi.org/10.4153/CJM-2016-035-8

Absolute Continuity of Wasserstein Barycenters Over Alexandrov Spaces

Yin Jiang

Abstract. In this paper, we prove that on a compact, n-dimensional Alexandrov space with curvature at least −1, the Wasserstein barycenter of Borel probability measures $\mu_1, \dots, \mu_m$ is absolutely continuous with respect to the n-dimensional Hausdorff measure if one of them is.

1 Introduction

In this paper, we study the barycenters in Wasserstein space over a compact Alexandrov space.

Let M be a compact, n-dimensional Alexandrov space with curvature at least −1. Denote by $P(M)$ the set of Borel probability measures on M, and by $P_{\text{ac}}(M)$ the set of absolutely continuous Borel probability measures on M.

Definition 1.1 A *Wasserstein barycenter* (with equal weights) of $\mu_1, \dots, \mu_m \in P(M)$ is defined as the Borel probability measure on M that minimizes

$$\mu \longmapsto \sum_{i=1}^{m} W_2^2(\mu_i, \mu),$$

where $W_2(\mu_i, \mu)$ denotes the quadratic Wasserstein distance from μ_i to μ.

The existence and uniqueness (under mild conditions) of Wasserstein barycenters are not difficult to establish; see Theorem 4.1. When $m = 2$, for μ_0 and μ_1, the barycenter $\mu_{\frac{1}{2}}$ is equivalent to the displacement interpolation [16], which was introduced in R. McCann's PhD thesis [15]. Proving that displacement interpolants are absolutely continuous (with respect to Hausdorff measure of the appropriate dimension) plays a key role in studying the behavior of functionals along these interpolants. Absolute continuity of displacement interpolants was proved in $\mathbb{R}^n$ in McCann's thesis, on Riemannian manifolds in [9], and on Alexandrov spaces in [10].

In the multi-marginal case, the Wasserstein barycenters were considered previously by Agueh–Carlier [1] when the underlying space is Euclidean. They proved that the barycenter is absolutely continuous with an L^∞ density if one of the marginals is. See also [7, 20, 21] for other results. Recently, for compact Riemannian manifolds, Y.-H. Kim and B. Pass [13] proved the absolute continuity of the Wasserstein barycenter for probability measures when one of $\mu_1, \dots, \mu_m$ is.

Received by the editors March 16, 2016; revised September 28, 2016.
Published electronically December 6, 2016.
AMS subject classification: **51K05**, 49K30.
Keywords: Alexandrov space, Wasserstein barycenter, multi-marginal optimal transport.

For our purpose, first of all we study the multi-marginal optimal transport problem. The existence and uniqueness of the solution is fundamental for our proof.

Let M be an n-dimensional Alexandrov space with curvature at least -1, possibly non-compact. Given Borel probability measures $\mu_1, \mu_2, \dots, \mu_m$ on M and a cost function $c\colon M^m \to \mathbb{R}$, we consider the multi-marginal optimal transportation problem of Monge. That is, minimize

$$\textbf{(M)} \qquad \int_M c\big(x_1, F_2(x_1), \dots, F_m(x_1)\big)\, \mathrm{d}\mu_1(x_1)$$

among $(m-1)$-tuples of mappings $(F_2, \dots, F_m)$ such that for each i, $F_i\colon M \to M$ pushes μ_1 forward to μ_i (often denoted by $F_\# \mu_1 = \mu_i$). That is, for any Borel subset $A \subset M$, $\mu_i(A) = \mu_1(F_i^{-1}(A))$.

Let

$$\pi_i\colon (x_1, \dots, x_m) \in M^m \longmapsto x_i \in M$$

be the projection operator. The set of probability measures on M^m, which project to μ_i for $i = 1, \dots, m$, is denoted by

$$\Gamma(\mu_1, \dots, \mu_m) := \{\gamma | \pi_\# \gamma = \mu_i\}.$$

The corresponding Kantorovich formulation of the multi-marginal optimal transport problem is to minimize

$$\textbf{(K)} \qquad \int_{M^m} c(x_1, x_2, \dots, x_m) \mathrm{d}\gamma(x_1, x_2, \dots, x_m)$$

among $\gamma \in \Gamma(\mu_1, \dots, \mu_m)$. When $c(x_1, \dots, x_m)$ is lower semi-continuous, an optimal plan for (K) always exists; see [2]. Note that the minimum may be $+\infty$ when M is non-compact.

When $m = 2$, this problem has been studied extensively over the past 25 years. In recent years, the multi-marginal case $m \geqslant 3$ of the Monge problem has attracted increasing attention. However, the structure of solutions for general cost functions is not well understood. Gangbo and Swiech [11] proved the existence and uniqueness of an optimal map for Monge problem for the cost function $\sum_{i \neq j} |x_i - x_j|^2$ on Euclidean space. Kim and Pass [14] generalized this result to compact Riemannian manifolds for cost

$$c(x_1, x_2, \dots, x_m) = \inf_{y \in M} \sum_{i=1}^{m} \frac{d^2}{2}(x_i, y).$$

Our first result is the following existence and uniqueness theorem for the cost

$$(1.1) \qquad c(x_1, x_2, \dots, x_m) = \inf_{y \in M} \sum_{i=1}^{m} f_i\big(d(x_i, y)\big),$$

where $f_i\colon [0, \infty) \to \mathbb{R}$ are C^1, strictly increasing, strictly convex functions, and the right derivative $f_i^+(0) = 0$ for each $i = 1, \dots, m$. This is a generalization of Kim and Pass's result.[1]

[1]Kim and Pass [14] proved an existence and uniqueness theorem for costs as in (1.1) in the case of a compact Riemannian manifold and when the $f_i{}'s$ are C^2.

https://doi.org/10.4153/CJM-2016-035-8 Published online by Cambridge University Press

Theorem 1.2 *Let M be an n-dimensional Alexandrov space (not necessarily compact) with curvature at least -1. Denote by $\mathcal{H}^n$ the n-dimensional Hausdorff measure. Assume μ_1 is absolutely continuous with respect to $\mathcal{H}^n$ and*

$$\inf_{\gamma}\Big\{\int_{M^m} c(x_1,\dots,x_m)\mathrm{d}\gamma : \gamma\in\Gamma(\mu_1,\dots,\mu_m)\Big\}<\infty;$$

then the solution γ to $(\boldsymbol{K})$ is concentrated on the graph of a mapping $(F_2,\dots,F_m)$ over the first variable. This mapping is a solution of Monge's problem $(\boldsymbol{M})$. And the solutions of both $(\boldsymbol{K})$ and $(\boldsymbol{M})$ are unique.

If $y\in M$ attains the infimum of $z\mapsto\sum_{i=1}^m f_i(d(x_i,z))$, then we say that y is a *mean* of $x_1,\dots,x_m$. If in addtion $f_i=t^2/2$, y is called a *barycenter* of $x_1,\dots,x_m$.

The approach of [14] consists of two parts. Let γ be an optimal measure in $(\mathbf{K})$. In the first part, Y.-H. Kim and B. Pass showed that if x_1 is μ_1-a.e., then those m-tuples $(x_1,\dots,x_m)\in\operatorname{spt}(\gamma)$ share the same means, which is a single point y. A key lemma in [14] is that any mean y of $x_1,\dots,x_m$ is not in the cut locus of x_i for each i, so each function $d^2/2(x_i,\cdot)$ is differentiable at y (in fact C^2) and $\sum_{i=1}^m\nabla_y\frac{d^2}{2}(x_i,y)=0$. In the second part, by this lemma, they showed that for each y there is at most one corresponding point $(x_1,\dots,x_m)$ in the support of γ. The combination of these two facts implies the existence and uniqueness of the solution of Monge problem. Our proof is basically along the lines of [14]. However, on Alexandrov spaces, the means might not even be regular points (see Example 3.9). To overcome this difficulty, we use the result by Ohta on barycenters on Alexandrov spaces (see Theorem 3.5).

Denote by γ the unique optimal measure in the multi-marginal problem $(\mathbf{K})$ and by $\operatorname{bc}(x_1,\dots,x_m)$ the set of barycenters of $x_1,\dots,x_m$ (which, by Lemma 3.3 and Theorem 3.10, is unique for γ almost all $(x_1,x_2,\dots,x_m)$). A result of Carlier and Ekeland [7] implies that $\nu:=\operatorname{bc}\#\gamma$ is the unique barycenter. This property plays an essential role in the proof of absolute continuity. Our main result of this paper is the following theorem, which generalizes a result of [13] on manifolds to Alexandrov spaces.

Theorem 1.3 *Let M be a compact, n-dimensional Alexandrov space with curvature at least -1. Let $\mu_1,\dots,\mu_m$ be Borel probability measures on M. If μ_1 is absolutely continuous with respect to $\mathcal{H}^n$, then the Wasserstein barycenter of $\mu_1,\dots,\mu_m$ is also absolutely continuous with respect to $\mathcal{H}^n$.*

In [13], the authors adapted an argument of Figalli–Juillet [10] (who studied the two measure case on the Heisenberg group and Alexandrov spaces). They first fix $x_2,\dots,x_m$ and prove that the map G from the barycenter y to x_1 is Lipschitz continuous. Then for $i=2,\dots,m$, they approximate μ_i by finite sum of Dirac measures, obtain uniform estimates for the approximating barycenters and pass to the limit. The Lipschitz continuity of G is essential, and their proof relies on the property that $\sum_{i=1}^m\frac{d^2}{2}(x_i,\cdot)$ are C^2 near y for all i. Our proof is basically along the same lines as theirs. However, for Alexandrov spaces, $d^2(x_i,\cdot)$ might not even be differentiable at y. To overcome this difficulty, we use Petrunin's perturbation method to perturb the function $\sum_{i=1}^m\frac{d^2}{2}(x_i,\cdot)$ in order to achieve the minimum at points where it is differentiable. Then we use the semi-concavity of $d^2(x_i,\cdot)$ to prove that G is Lipschitz

https://doi.org/10.4153/CJM-2016-035-8 Published online by Cambridge University Press

continuous. It seems to me that, at least, we cannot use the Lipschitz character of gradient curves relative to a semiconcave function directly; see Remark 4.4.

Our paper is organized as follows. In Section 2, we will recall the definition of Alexandrov spaces and some properties. In Section 3, we will prove Theorem 1.2 for compact Alexandrov spaces. In Section 4, we will prove Theorem 1.3. In the appendix, for the completeness of the theory of existence and uniqueness, we will prove Theorem 1.2 for non-compact Alexandrov spaces.

2 Preliminaries

In this section, we review the definition and some properties of Alexandrov spaces with curvature bounded below. These definitions and results are mainly taken from [5, 6, 18].

Let (M, d) be a metric space. A rectifiable curve γ connecting two points p, q is called a *geodesic* if its length is equal to $d(p, q)$ and it has unit speed. A metric space is called a *geodesic space* if any two points $p, q \in M$ can be connected by a geodesic. Denote by M_k^2 the simply connected 2-dimensional space form of constant curvature k. Given three points p, q, r in a geodesic space M, we can take a comparison triangle $\Delta\widetilde{p}\widetilde{q}\widetilde{r}$ in M_k^2 such that

$$d(\widetilde{p}, \widetilde{q}) = d(p, q), \quad d(\widetilde{p}, \widetilde{r}) = d(p, r), \quad d(\widetilde{q}, \widetilde{r}) = d(q, r).$$

If $k > 0$, we add the assumption $d(p, q) + d(p, r) + d(q, r) < 2\pi/\sqrt{k}$. The angle $\widetilde{\angle}_k pqr := \angle\widetilde{p}\widetilde{q}\widetilde{r}$ is called the comparison angle.

Definition 2.1 A geodesic space M is called an *Alexandrov space with curvature* at least k if it is locally compact, and for any point $x \in M$, there exists a neighborhood U_x such that, for any four different points p, a, b, c in U_x, we have

$$\widetilde{\angle}_k abp + \widetilde{\angle}_k bpc + \widetilde{\angle}_k cpa \leqslant 2\pi.$$

The Hausdorff dimension of an Alexandrov space is always an integer. Let M be an n-dimensional Alexandrov space with curvature $\geqslant k$. Denote the n-dimensional Hausdorff measure by $\mathcal{H}^n$. Given any two geodesics $\gamma(t)$ and $\eta(s)$ with $\gamma(0) = \eta(0) = p$, the angle

$$\angle\big(\gamma^+(0), \eta^+(0)\big) := \lim_{t,s\to 0} \widetilde{\angle}_k \gamma(t) p \eta(s)$$

is well defined.

We say $\eta(t)$ is equivalent to $\gamma(t)$ if $\angle(\gamma^+(0), \eta^+(0)) = 0$. Denote by Σ_p' the set of equivalence classes of geodesic $\gamma(t)$ with $\gamma(0) = p$. The space of directions Σ_p is the completion of metric space $(\Sigma_p', \angle)$.

The tangent cone at p, T_p, is the Euclidean cone over Σ_p; it is an Alexandrov space with curvature at least 0. For any two vectors $u, v \in T_p$. The "scalar product" (see [26, section 1]) is defined by

$$\langle u, v\rangle = |u||v| \cos\angle(u, v).$$

The distance $|uv|$ is defined by the law of cosines

$$|uv|^2 = |u|^2 + |v|^2 - 2|u||v| \cos\angle(u, v).$$

For each point $x \neq p$, we denote by $\Uparrow_p^x$ the set of directions at p corresponding to all geodesics connecting p to x. The symbol $\uparrow_p^x$ denotes the direction at p corresponding to some geodesic px. Given a direction $\xi \in \Sigma_p$, it is possible that there exists no geodesic $\gamma(t)$ starting at p with $\gamma^+(0) = \xi$. However, it was shown in [23] that for $p \in M$ and any direction $\xi \in \Sigma_p$, there exists a quasi-geodesic $\gamma : [0, +\infty) \to M$ with $\gamma(0) = p$ and $\gamma^+(0) = \xi$.

For $p \in M$, denote by

$$W_p := \{x \in M \backslash p \mid \text{ there exists } y \in M \text{ such that } y \neq x \text{ and } |py| = |px| + |xy|\}.$$

According to [18], W_p has full measure. Since geodesics do not branch in Alexandrov spaces [6], for any $x \in W_p$, there is a unique geodesic connecting p to x, $\Uparrow_p^x$ contains only one element, and the direction $\uparrow_p^x$ is uniquely determined. Recall that the map $\log_p \colon W_p \to T_p$ is defined by $\log_p(x) := |px| \cdot \uparrow_p^x$. Setting

$$\mathcal{W}_p := \log_p(W_p) \subset T_p,$$

the map $\log_p \colon W_p \to \mathcal{W}_p$ is one-to-one. The exponential map $\exp_p \colon T_p \to M$ is defined by Petrunin [25] as follows: $\exp_p(o_p) = p$ and for any $v \in T_p \backslash \{o_p\}$, $\exp_p(v)$ is a point on some quasi-geodesic of length $|v|$ starting from p along direction $v/|v| \in \Sigma_p$. If the quasi-geodesic is not unique, we fix some one of them as the definition of $\exp_p(v)$.

Next, we introduce λ-concave functions and semi-concave functions. See [26, section 1].

Definition 2.2 Let M be an n-dimensional Alexandrov space without boundary and let $U \subset M$ be an open subset. A locally Lipschitz function $f \colon U \mapsto \mathbb{R}$ is called *λ-concave* if for any geodesic $\gamma(t)$ in U, the function $f \circ \gamma(t) - \lambda t^2/2$ is concave.

A function $f \colon M \mapsto \mathbb{R}$ is called *semiconcave* if for any point $x \in M$, there is a neighborhood $U_x \ni x$ and $\lambda \in \mathbb{R}$ such that the restriction $f|_{U_x}$ is λ-concave. Given a semiconcave function $f \colon M \mapsto \mathbb{R}$, its differential $d_p f$ is well defined for each point $p \in M$. Let $\varphi \colon \mathbb{R} \to \mathbb{R}$ be a continuous function. A function $f \colon M \mapsto \mathbb{R}$ is called *$\varphi(f)$-concave* if for any point $x \in M$ and $\epsilon > 0$, there is a neighborhood $U_x \ni x$ such that $f|_{U_x}$ is $(\varphi \circ f(x) + \epsilon)$-concave.

A point p in an n-dimensional Alexandrov space M is said to be *regular* if its tangent cone T_p is isometric to $\mathbb{R}^n$ with standard metric. Denote by $\mathrm{Reg}(\mathrm{M})$ the set of regular points.

Definition 2.3 We say that a function u is *differentiable* at $x \in \mathrm{Reg}(\mathrm{M})$, if there exists a vector in T_x, denoted by $\nabla u(x)$, such that for any geodesic $\gamma(t)$ with $\gamma(0) = x$

$$u(\gamma(t)) = u(x) + \left\langle \nabla u(x), \gamma^+(0) \right\rangle t + o(t).$$

The Rademacher theorem, in the framework of a metric measure space with a doubling measure and a Poincaré inequality for upper gradient, was proved by Cheeger [8]. In [4], Bertrand proved it in Alexandrov space via a simple argument: a locally Lipschitz function u is differentiable almost everywhere with respect to $\mathcal{H}^n$ in M. The points where a distance function is differentiable have the following property.

Lemma 2.4 *For any $p \in M$, if $f(\cdot) := d(p,\cdot)$ is differentiable at x, then there exists a unique geodesic connecting p to x.*

Proof Suppose not; then there exist two geodesics $\gamma_1(t), \gamma_2(t)$ connecting p to x with $\gamma_1(0) = \gamma_2(0) = x$. Since f is differentiable at x, for $i = 1, 2$ we have

$$-t = f \circ \gamma_i(t) - f(x) = \big\langle \nabla f(x), \gamma_i^+(0) \big\rangle t + o(t).$$

This means that

$$\cos \angle \Big(\frac{\nabla f(x)}{|\nabla f(x)|}, \gamma_i^+(0) \Big) = -1,$$

which is impossible. ∎

We will make frequent use of the following two lemmas.

Lemma 2.5 *For any $p \in M$, let $f(x) := d(p,x)$, then the function f is $\frac{\cosh f}{\sinh f}$-concave on $M \backslash p$ and $d^2(p,x)$ is $f\frac{\cosh f}{\sinh f}$-concave on M.*

Lemma 2.6 *Let $f\colon M \mapsto \mathbb{R}$ be a semiconcave function. If f achieves a local minimum at $x \in M$, then $d_x f = 0$. If we assume in addition that f is differentiable at x, then $\nabla f(x) = o_x$.*

Next we introduce the first variation formula, which is important in the proof of Theorem 1.2; see [5, Corollary 4.57 and Remark 4.5.12].

Theorem 2.7 (First variation formula of arc length [5]) *Let M be an n-dimensional Alexandrov space with curvature at least -1. For any geodesic γ and $p \in M$, $p \neq \gamma(0)$, the function $t \to l(t) = d(p,\gamma(t))$ has the right derivative and*

$$\lim_{t\to 0^+} \frac{l(t) - l(0)}{t} = -\cos \angle \alpha_{\min},$$

where $\alpha_{\min}$ is the infimum (in fact, minimum) of angles between γ and shortest geodesics connecting $\gamma(0)$ to p.

Set $\angle \alpha_{\min} \triangleq \angle(\Uparrow^p_{\gamma(0)}, \gamma^+(0))$. Note that for any $\uparrow^p_{\gamma(0)} \in \Uparrow^p_{\gamma(0)}$, since

$$\angle\big(\Uparrow^p_{\gamma(0)}, \gamma^+(0)\big) \leqslant \angle\big(\uparrow^p_{\gamma(0)}, \gamma^+(0)\big),$$

we have

$$\lim_{t\to 0^+} \frac{l(t) - l(0)}{t} = -\cos \angle\big(\Uparrow^p_{\gamma(0)}, \gamma^+(0)\big) \leqslant -\cos \angle\big(\uparrow^p_{\gamma(0)}, \gamma^+(0)\big) = -\langle \uparrow^p_{\gamma(0)}, \gamma^+(0)\rangle. \tag{2.1}$$

Next, we introduce Perelman's concave function.

Lemma 2.8 (Perelman's concave function [22, 24]) *For any $p \in M$, there exists a constant $r_1 > 0$ and a function $h\colon B_p(r_1) \mapsto \mathbb{R}$ such that*

(i) *h is -1-concave;*
(ii) *h is 2-Lipschitz (*i.e., *2 can be a Lipschitz constant).*

https://doi.org/10.4153/CJM-2016-035-8 Published online by Cambridge University Press

(iii) *For each $x \in B_p(r_1)$, we have*

$$\int_{\Sigma_x} d_x h(\xi) \mathrm{d}\xi \leqslant 0.$$

Moreover, if "=" holds, then x is regular.

The existence of such a concave function and (i), (ii) are due to Perelman [22]. Part (iii) is implicitly claimed in Petrunin's manuscript [24]. See [29, Lemma 3.3] for a detailed proof.

We now follow Petrunin in [24] to introduce a perturbation argument; we also refer the reader to [29]. The following lemma is used to perturb a non-regular point to a regular point; it is a particular case of one appeared in [24, p. 10].

Lemma 2.9 *Let h be as above. Suppose that u is a λ-concave function on $B_p(r_1)$. For any $\epsilon > 0$, if $x \in B_p(r_1)$ is a minimum point of function $u + \epsilon h$, then x has to be a regular point.*

Sketch of proof First, we state a property. Let $\Omega \subset M$ be a domain, suppose $f\colon \Omega \to \mathbb{R}$ is a semiconcave function, and $x \in \Omega$; then

$$\int_{\Sigma_x} d_x f(\xi) \mathrm{d}\xi \leqslant 0.$$

See, for example, [29, Proposition 3.1] for a proof. It follows that

$$\int_{\Sigma_x} d_x u(\xi) \mathrm{d}\xi \leqslant 0, \qquad \int_{\Sigma_x} d_x h(\xi) \mathrm{d}\xi \leqslant 0. \tag{2.2}$$

Since x is a minimum point of $u + \epsilon h$, by Lemma 2.6, for any $\xi \in \Sigma_x$, we have

$$0 = d_x(u + \epsilon h)(\xi) = d_x u(\xi) + \epsilon d_x h(\xi).$$

It follows that

$$\int_{\Sigma_x} d_x u(\xi) \mathrm{d}\xi + \epsilon \int_{\Sigma_x} d_x h(\xi) \mathrm{d}\xi = 0. \tag{2.3}$$

Since $\epsilon > 0$, by combining (2.2) and (2.3), we obtain

$$\int_{\Sigma_x} d_x h(\xi) \mathrm{d}\xi = 0.$$

By Lemma 2.8, x is a regular point. ∎

Let u be a λ-concave function on a bounded domain U. Suppose x_0 is the unique minimum point of u on U and $u(x_0) < \min_{x \in \partial U} u$. It is easy to see that $\lambda > 0$. Otherwise, by Lemma 2.6, we have $u(x) \leqslant u(x_0)$ for $x \in U$, which contradicts the fact that x_0 is the unique minimum point. Suppose also that x_0 is regular. By [6, theorem 9.4], there exist n points $a_1, \dots, a_n \in U$ such that

$$g = \big(d(a_1, \cdot), \dots, d(a_n, \cdot) \big)$$

https://doi.org/10.4153/CJM-2016-035-8 Published online by Cambridge University Press

maps a small neighborhood $B_{R_0}(x_0)(\subset U)$ almost isometrically onto a domain in $\mathbb{R}^n$. That is, there exists a sufficiently small number $\kappa > 0$ such that

$$\left| \frac{\|g(x) - g(y)\|}{d(x,y)} - 1 \right| \leqslant \kappa \quad \text{for all } x, y \in B_{R_0}(x_0), x \neq y. \tag{2.4}$$

For $1 \leqslant i \leqslant n$, denote $g_i(x) = d(a_i, x)$. By Lemma 2.5, there exists a positive constant λ_0 depending only on $\min_{\{1\leqslant i\leqslant n\}} d(a_i, B_{R_0}(x_0))$, such that g_i are λ_0-concave on $B_{R_0}(x_0)$ for $1 \leqslant i \leqslant n$. There exists $\epsilon_0 > 0$ such that for each vector $V = (v^1, v^2, \dots, v^n)$ with $|v^i| \leqslant \epsilon_0$ for all $1 \leqslant i \leqslant n$, the function

$$G(V,x) := u(x) + \sum_{i=1}^{n} v^i g_i(x)$$

has a minimum point in the interior of U. We can define $\rho: [0,\epsilon_0]^n \subset \mathbb{R}^n \mapsto U$ by letting $\rho(V)$ be one of the minimum point of $G(V,x)$. The following lemma is used to perturb a regular point to a nearby point; compare the lemma on [24, p. 8].

Lemma 2.10 *Let $u, x_0, \{g_i\}_{i=1}^n$ and let ρ be as above. Then there exists some $\epsilon \in (0,\epsilon_0)$ and $\delta > 0$ such that*

$$d\big(\rho(V), \rho(W)\big) \geqslant \delta\|V - W\|, \quad \forall V, W \in [0,\epsilon]^n.$$

In particular, for arbitrary $\epsilon' \in (0,\epsilon)$, the image $\rho([0,\epsilon']^n)$ has nonzero Hausdorff measure.

Remark 2.11 In Lemmas 2.9 and 2.10, we assume that u is λ-concave, which is stronger than Petrunin's assumption. In [24], Petrunin used a chart with concave components near a regular point, while in Lemma 2.10, we do not need the components to be concave.

Proof As $V \to 0$, $G(V,x) \to u$. Since x_0 is the unique minimum point of u, we have that $\rho(V) \to x_0$. We fix a small positive number $\epsilon > 0$ such that when $|v_i| \leqslant \epsilon$ for all $1 \leqslant i \leqslant n$, $\rho(V) \in B_{\frac{R_0}{4}}(x_0)$. Since $\rho(V)$ is the minimum point of $G(V,x)$, we have that

$$\begin{aligned} G\big(W,\rho(W)\big) &- G\big(W,\rho(V)\big) \\ &= G\big(V,\rho(W)\big) - G\big(V,\rho(V)\big) + (W - V)\big(g(\rho(W)) - g(\rho(V))\big) \\ &\geqslant (W - V)\big(g(\rho(W)) - g(\rho(V))\big) \\ &\geqslant -2\|W - V\| d\big(\rho(W),\rho(V)\big), \end{aligned} \tag{2.5}$$

the last inequality holding since g is an almost isometry. Denote $\overline{\lambda} := \lambda + n\epsilon\lambda_0 > 0$. Since g_i are λ_0-concave for all $1 \leqslant i \leqslant n$, $G(V,x)$ is $\overline{\lambda}$-concave. Since $\rho(V)$ is the minimum point of $g(V,x)$, by Lemma 2.6, for any $x \in B_{R_0}(x_0)$, we have

$$\begin{aligned} G(W,x) &- G\big(W,\rho(V)\big) \\ &= G(V,x) - G\big(V,\rho(V)\big) + (W - V)\big(g(x) - g(\rho(V))\big) \\ &\leqslant \frac{\overline{\lambda}}{2}|x\rho(V)|^2 + (W - V)\big(g(x) - g(\rho(V))\big). \end{aligned} \tag{2.6}$$

https://doi.org/10.4153/CJM-2016-035-8 Published online by Cambridge University Press

For any $0 < R < \frac{R_0}{2}$, there exists $x \in B_{R_0}(x_0)$ with $d(x, \rho(V)) = R$ and

$$(W - V)\big(g(x) - g(\rho(V))\big) \leqslant -\frac{\|W - V\| \|g(x) - g(\rho(V))\|}{2} \leqslant -\frac{\|W - V\| R}{4}. \tag{2.7}$$

Since $G(W, x) \geqslant G(W, \rho(W))$ for any $x \in B_{R_0}(x_0)$, by combining (2.5), (2.6), and (2.7), we have

$$\frac{\overline{\lambda}}{2} R^2 - \frac{R\|W - V\|}{4} \geqslant -2d\big(\rho(W), \rho(V)\big)\, \|W - V\|.$$

Set $N = \frac{4n\epsilon}{R_0 \overline{\lambda}} + 1$, choose

$$R = \frac{\|W - V\|}{10\overline{\lambda} N} \leqslant \frac{n\epsilon}{10\overline{\lambda} N} \leqslant \frac{R_0}{10}.$$

We get that

$$d(\rho(W), \rho(V)) \geqslant \frac{1}{80\overline{\lambda} N}(1 - \frac{1}{5N}) \|W - V\|.$$

Let $\delta = \frac{1}{80\overline{\lambda} N}(1 - \frac{1}{5N})$, and we complete the proof. ∎

3 Existence and Uniqueness of the Solution of Multi-marginal Optimal Transport Problem

The dual problem to (K) is to maximize

$$\sum_{i=1}^{m} \int_M u_i(x_i) \mathrm{d}\mu_i(x_i) \tag{D}$$

among all m tuples $(u_1, \ldots, u_m)$ of functions, with $u_i \in L^1(\mu_i)$ and

$$\sum_{i=1}^{m} u_i(x_i) \leqslant c(x_1, \ldots, x_m)$$

for μ_1-a.e. $x_1 \in M$,..., μ_m-a.e. $x_m \in M$

We say that an m-tuple $(u_1, \ldots, u_m)$ is c-conjugate if, for all $i = 1, \ldots, m$, we have

$$u_i(x_i) = \inf_{x_j \in M, j \neq i} \Big[c(x_1, \ldots, x_m) - \sum_{j \neq i} u_j(x_j) \Big]. \tag{3.1}$$

The following theorem is well known. In fact, it holds on more general settings; see [12, 19, 27, 28] for further discussion.

Theorem 3.1 *If*

$$\inf_{\gamma} \Big\{ \int_{M^m} c(x_1, \ldots, x_m) \mathrm{d}\gamma : \gamma \in \Gamma(\mu_1, \ldots, \mu_m) \Big\} < \infty,$$

then there exists a c-conjugate solution $(u_1, \ldots, u_m)$ to (D) and the maximum value in (D) is finite. If γ is an optimal measure in the Kantorovich problem, we have

$$\sum_{i=1}^{m} u_i(x_i) = c(x_1, \ldots, x_m) \tag{3.2}$$

https://doi.org/10.4153/CJM-2016-035-8 Published online by Cambridge University Press

γ-almost everywhere.

In this section, in order to make our argument more clear and understandable, in this section, we only prove Theorem 1.2 for compact Alexandrov spaces, which is enough for our use. The non-compact case, the proof of which follows the same steps as that of the compact case, is left to the appendix.

So in this section, we suppose that M is a compact, n-dimensional Alexandrov space with curvature $\geqslant -1$. The following two lemmas are standard.

Lemma 3.2 *The cost function $c(x_1,\dots,x_m)$ is Lipchitz. Suppose u_1 is given by (3.1). If u_1 is not identically infinity, then u_1 is Lipschitz.*

Proof For any $(x_1,\dots,x_m),(x_1',\dots,x_m') \in M^m$. Let y be a mean of $x_1,\dots,x_m$. Since f_i' is increasing for each i, any $x_i, y \in M$,

$$f_i'(d(x_i,y)) \leqslant f_i'(\operatorname{diam} M) \triangleq L_i < \infty.$$

By the definition of c, we have

$$\begin{aligned} c(x_1',\dots,x_m') - c(x_1,\dots,x_m) &\leqslant \sum_{i=1}^m f_i\big(d(x_i',y)\big) - \sum_{i=1}^m f_i\big(d(x_i,y)\big) \\ &\leqslant \sum_{i=1}^m \big| f_i\big(d(x_i',y)\big) - f_i\big(d(x_i,y)\big)\big| \\ &\leqslant \sum_{i=1}^m L_i d(x_i,x_i') \leqslant m\sum_{i=1}^m L_i \sqrt{\sum_{i=1}^m d^2(x_i,x_i')}. \end{aligned}$$

Thus, c is $m(L_1+\cdots+L_m)$-Lipschitz.

Since c is bounded on M^m, if there exists a point $z\in M$ such that $u_1(z) = -\infty$, then there exists a sequence $(x_2^k,\dots,x_m^k)\in M^{m-1}$ such that $\sum_{j=2}^m u_j(x_j^k)\to +\infty$ as $k\to\infty$. For any $x_1\in M$, choose this sequence in (3.1), and we can get $u_1\equiv -\infty$. If there exists a point $z\in M$ such that $u_1(z) = +\infty$, then for any $m-1$ tuple $(x_2,\dots,x_m)$,

$$-\sum_{j=2}^m u_j(x_j) = +\infty;$$

thus, $u_1\equiv +\infty$. So if u_1 is not identically infinity, it is finite.

In the case where u_1 is finite, for all $\epsilon>0$, there exists $(x_{2,\epsilon},\dots,x_{m,\epsilon})\in M^{m-1}$ such that

$$u_1(x_1) \geqslant c(x_1,x_{2,\epsilon},\dots,x_{m,\epsilon}) - \sum_{j=2}^m u_j(x_j) - \epsilon.$$

For $x_1'\neq x_1$,

$$\begin{aligned} u_1(x_1') - u_1(x_1) &\leqslant c(x_1',x_{2,\epsilon},\dots,x_{m,\epsilon}) - c(x_1,x_{2,\epsilon},\dots,x_{m,\epsilon}) + \epsilon \\ &\leqslant L_1 d(x_1,x_1') + \epsilon. \end{aligned}$$

By the arbitrariness of ϵ, we have that u_1 is L_1-Lipschitz. ∎

Since $u_1\in L^1(\mu_1)$, we have that u_1 is not identically infinity; then u_1 is Lipschitz, hence μ_1-a.e. differentiable. For any $x_1\in M$, by the compactness of M, there exists

$(x_2, \dots, x_m) \in M^{m-1}$ such that $\sum_{i=1}^m u_i(x_i) = c(x_1, \dots, x_m)$. The next lemma implies that for μ_1-a.e. x_1, those m-tuples $(x_1, \dots, x_m) \in \operatorname{spt}(\gamma)$ share the same means, which is a single point.

Lemma 3.3 *If $x_1 \in \operatorname{Reg}(M)$ is a point where u_1 is differentiable, then for any $(x_2, \dots, x_m) \in M^{m-1}$ such that $\sum_{i=1}^m u_i(x_i) = c(x_1, \dots, x_m)$, c is differentiable with respect to x_1 at $(x_1, \dots, x_m)$ and*

$$\nabla_{x_1} c(x_1, \dots, x_m) = \nabla u_1(x_1). \tag{3.3}$$

The mean is uniquely determined by $\nabla u_1(x_1)$:

$$y = \begin{cases} x_1, & \text{if } \nabla u_1(x_1) = 0, \\ \exp_{x_1}\Big(-(f_1')^{-1}(|\nabla u_1(x_1)|)\frac{\nabla u_1(x_1)}{|\nabla u_1(x_1)|}\Big) & \text{if } \nabla u_1(x_1) \neq 0. \end{cases} \tag{3.4}$$

Proof Let $\gamma(t)$ be a geodesic with $\gamma(0) = x_1$. For any $(x_2, \dots, x_m) \in M^{m-1}$ such that $\sum_{i=1}^m u_i(x_i) = c(x_1, \dots, x_m)$,

$$\begin{aligned} c\big(\gamma(t), x_2, \dots, x_m\big) - c(x_1, \dots, x_m) &\geqslant u_1\big(\gamma(t)\big) + \sum_{i=2}^m u_i(x_i) - \sum_{i=1}^m u_i(x_i) \\ &= u_1(\gamma(t)) - u_1(x_1) \\ &= \big\langle \nabla u_1(x_1), \gamma^+(0)\big\rangle t + o(t). \end{aligned} \tag{3.5}$$

Let y be a mean of $x_1, \dots, x_m$. Then

$$\begin{aligned} & c\big(\gamma(t), x_2, \dots, x_m\big) - c(x_1, \dots, x_m) \\ &\quad \leqslant f_1\big(d(\gamma(t), y)\big) + \sum_{i=2}^m f_i\big(d(x_i, y)\big) - \sum_{i=1}^m f_i\big(d(x_i, y)\big) \\ &\quad = f_1\big(d(\gamma(t), y)\big) - f_1\big(d(x_1, y)\big). \end{aligned} \tag{3.6}$$

If $y \neq x_1$, for any $\uparrow_{x_1}^y \in \Uparrow_{x_1}^y$, by (2.1), we have

$$f_1\big(d(\gamma(t), y)\big) - f_1\big(d(x_1, y)\big) \leqslant -\big\langle f_1'\big(d(x_1, y)\big)\uparrow_{x_1}^y, \gamma^+(0)\big\rangle t + o(t).$$

If $y = x_1$, since $f^+(0) = 0$, then

$$f_1\big(d(\gamma(t), y)\big) - f\big(d(x_1, y)\big) = f(t) - f(0) = o(t),$$

which coincides with the above inequality.

Thus, we have

$$\big\langle \nabla u_1(x_1) + f_1'(d(x_1, y))\uparrow_{x_1}^y, \gamma^+(0)\big\rangle \leqslant 0.$$

Applying the above argument to a sequence of geodesics starting at x, whose directions converge to $-\gamma^+(0)$, we get

$$f_1'(d(x_1, y))\uparrow_{x_1}^y = -\nabla u_1(x_1).$$

Together with (3.5) and (3.6), we can get that $c(\cdot, x_2, \dots, x_m)$ is differentiable at x_1 and (3.3).

Thus, we have

$$f_1'\big(d(x_1, y)\big) = |\nabla u_1(x_1)|. \tag{3.7}$$

If $\nabla u_1(x_1) = 0$, since f_1' is strictly increasing from 0, then $y = x_1$. If $\nabla u_1(x_1) \neq 0$, then

$$\uparrow_{x_1}^{y} = -\frac{\nabla u_1(x_1)}{f_1'(d(x_1, y))} = -\frac{\nabla u_1(x_1)}{|\nabla u_1(x_1)|}. \tag{3.8}$$

Thus, we have (3.4). ∎

Remark 3.4 If $y \neq x_1$, from (3.7) and (3.8), we can see that there exists a unique geodesic connecting x_1 to y.

In [17], Ohta obtained a crucial property of the means; see [17, Lemma 4.6 and Theorem 4.11]. We adapt the original statement for our use.

Theorem 3.5 (Ohta [17]) *Let M be an Alexandrov space and let $y \in M$ be a mean of $x_1, \ldots, x_m \in M$. Then for each i, there exists a unique geodesic connecting y to x_i. If $y = x_i$ for some i, we just let $\uparrow_y^{x_i} = o_y$. Moreover, $\uparrow_y^{x_1}, \ldots, \uparrow_y^{x_m}$ are contained in a subset $\mathcal{H} \subset T_y$ which is isometric to a Hilbert space, and*

$$\sum_{i=1}^{m} f_i'(d(x_i, y)) \uparrow_y^{x_i} = 0. \tag{3.9}$$

Remark 3.6 Ohta's original theorem corresponds to the case where $f_i(t) = \frac{t^2}{2}$. Following the proof of [17, Lemma 4.6 and Theorem 4.11], it is not hard to prove the above theorem.

Corollary 3.7 ([17, Corollary 4.12]) *Suppose that, at a point $y \in M$, no pair of directions $\xi, \eta \in \Sigma_y$ satisfies $\angle(\xi, \eta) = \pi$. Then y cannot be a mean.*

Next we list two typical examples.

Example 3.8 Let M be the Euclidean cone over a circle of length $l \in (0, 2\pi)$. Then the origin of the cone cannot be a barycenter.

Example 3.9 Let S be the spherical suspension over a circle of length $l \in (0, 2\pi)$. Let M be the Euclidean cone over S with origin p. Then $\Sigma_p = S$, so p is not a regular point. Let $\xi, \eta \in S$ with $\angle(\xi, \eta) = \pi$. Then the two rays starting at p along the directions ξ and η form a geodesic through p. Thus, p is a barycenter of any two points on this geodesic with the same distance to p.

Next we show that given a mean $y \in M$, the m tuple $(x_1, \ldots, x_m) \in \operatorname{spt}(\gamma)$ can be uniquely determined.

Theorem 3.10 *Suppose $x = (x_1, \ldots, x_m)$ and $\overline{x} = (\overline{x}_1, \ldots, \overline{x}_m)$ are both in $\operatorname{spt}(\gamma)$ and there exists $y \in M$ that is both the mean of $x_1, \ldots, x_m$ and $\overline{x}_1, \ldots, \overline{x}_m$. Then $x = \overline{x}$.*

Proof First, we recall a basic fact about multi-marginal optimal transport, known as c-monotonicity. If $x, x' \in \operatorname{spt}\gamma$ for an optimal γ, set $x' = (x_1, \ldots, \overline{x}_i, \ldots, x_m)$ and

https://doi.org/10.4153/CJM-2016-035-8 Published online by Cambridge University Press

$\overline{x}' = (\overline{x}_1, \dots, x_i, \dots, \overline{x}_m)$. Then

$$c(x') + c(\overline{x}') \geqslant c(x) + x(\overline{x}). \tag{3.10}$$

We will show that $x_i = \overline{x}_i$. We have

$$\begin{aligned} c(x') + c(\overline{x}') &\leqslant \sum_{j \neq i} f_j\big(d(x_j, y)\big) + f_i\big(d(\overline{x}_i, y)\big) + \sum_{j \neq i} f_j\big(d(\overline{x}_j, y)\big) + f_i\big(d(x_i, y)\big) \\ &= \sum_{j=1}^{m} f_j\big(d(x_j, y)\big) + \sum_{j=1}^{m} f_j\big(d(\overline{x}_j, y)\big) = c(x) + c(\overline{x}). \end{aligned}$$

In view of (3.10), the above inequalities are equalities, and we have

$$c(x') = \sum_{j \neq i} f_j(d(x_j, y)) + f_i(d(\overline{x}_i, y)).$$

That is, y is also a mean of $x_1, \dots, \overline{x_i}, \dots, x_m$. By (3.9), we have

$$\sum_{i=1}^{m} f_i'\big(d(x_i, y)\big) \uparrow_y^{x_i} = 0 = \sum_{j \neq i} f_j'\big(d(x_i, y)\big) + f_i'\big(d(\overline{x}_i, y)\big) \uparrow_y^{\overline{x}_i}.$$

Hence, we obtain

$$f_i'\big(d(x_i, y)\big) \uparrow_y^{x_i} = f_i'\big(d(\overline{x}_i, y)\big) \uparrow_y^{\overline{x}_i}.$$

If $y = x_i$, then $\uparrow_y^{\overline{x}_i} = o_y$ or $d(\overline{x}_i, y) = 0$; in either case, $y = \overline{x}_i$. If $y = \overline{x}_i$, we have $y = x_i$. If $y \neq x_i$ and $y \neq \overline{x}_i$, then

$$\uparrow_Y^{x_i} = \uparrow_y^{\overline{x}_i} \quad \text{and} \quad f_i'\big(d(x_i, y)\big) = f_i'\big(d(\overline{x}_i, y)\big).$$

Since f_i' is strictly increasing, we have $d(x_i, y) = d(\overline{x}_i, y)$. So $x_i = \overline{x}_i$. ∎

Next, we prove Theorem 1.2 for the compact case.

Proof For any optimal measure γ, we need to show that for μ_1-a.e. x_1, there is a unique $(x_1, x_2, \dots, x_m) \in \operatorname{spt}(\gamma)$. If not, then there exists a Borel subset $A \subset M$ with $\mu_1(A) > 0$, such that for $x_1 \in A$, there are no $(x_1, x_2, \dots, x_m) \in \operatorname{spt}(\gamma)$. That is, $A \times M^{m-1} \cap \operatorname{spt}(\gamma) = \varnothing$. However, $\gamma(A \times M^{m-1}) = \mu_1(A) > 0$, contradiction. So we have existence.

Since u_1 is Lipschitz, it is differentiable $\mathcal{H}^n$-a.e. Since μ_1 is absolutely continuous with respect to $\mathcal{H}^n$, we have u_1 is differentiable μ_1-a.e. Thus, at μ_1-a.e. x_1, u_1 is differentiable and there exists $(x_1, x_2, \dots, x_m) \in \operatorname{spt}(\gamma)$.

For such a point, by Lemma 3.3, all the m-tuples $(x_1, x_2, \dots, x_m) \in \operatorname{spt}(\gamma)$ have the same unique mean y. While by Theorem 3.10, there is only one $(x_1, x_2, \dots, x_m) \in \operatorname{spt}(\gamma)$ such that y is the mean of $(x_1, x_2, \dots, x_m)$. Thus, we have the uniqueness. We define the map $(x_2, \dots, x_m) = (F_2(x_1), \dots, F_m(x_1))$, it is well defined μ_1-a.e.

It remains to prove the uniqueness of the optimal measure γ. Suppose there exists another optimal minimizer $\overline{\gamma}$. By the above argument, it is concentrated on a graph, denoted by $(\overline{F}_2, \dots, \overline{F}_m)$. By linearity of the Kantorovich functional, $\frac{1}{2}\gamma + \frac{1}{2}\overline{\gamma}$ is also optimal and must be concentrated on a graph. This implies that

$$(F_2, \dots, F_m) = (\overline{F}_2, \dots, \overline{F}_m)$$

for μ_1-a.e. $x \in M$. Thus, we complete the proof. ∎

Remark 3.11 In [14], by assuming f_i to be C^2, the authors proved that the mean y is not in the cutlocus of $x_1, \dots, x_m$ We only assume that $f_i \in C^1$, since our argument does not rely on this property.

4 Barycenters in Wasserstein Space Over Alexandrov Spaces

In this section, we prove Theorem 1.3. Recall the basic setting: M is a compact, n-dimensional Alexandrov space with curvature at least -1, $\mu_1, \dots, \mu_m$ are Borel probability measures, μ_1 is absolutely continuous with respect to $\mathcal{H}^n$.

We present below an existence and uniqueness theorem, which is due to Pass and Kim.

Theorem 4.1 ([13, Theorem 3.1]) *Let M be a compact, n-dimensional Alexandrov space with curvature at least -1, let μ_i be Borel probability measures on M. If at least one of them is absolutely continuous with respect to $\mathcal{H}^n$, then there exists a unique Wasserstein barycenter.*

Since μ_1 is absolutely continuous with respect to $\mathcal{H}^n$, there exists a unique Wasserstein barycenter. The corresponding cost function we consider in this section is:

$$c(x_1, \dots, x_m) = \inf_{z \in M} \sum_{i=1}^{m} \frac{d^2}{2}(x_i, z).$$

In this case, the means are also called barycenters. We adopt some notation from [13]: $\operatorname{bc}(x_1, x_2, \dots, x_m)$ denotes the set of barycenters of $x_1, \dots, x_m$. For each Borel set $E \subset M$, and $m-1$ points $x_2, \dots, x_m$, let

$$\operatorname{bc}(E, x_2, \dots, x_m) := \bigcup_{x_1 \in E} \operatorname{bc}(x_1, x_2, \dots, x_m).$$

Let γ be the unique optimal measure in the multi-marginal problem (K). Since the barycenter is unique for γ-a.e. $(x_1, \dots, x_m)$, this gives a γ-a.e. one-to-one map $\operatorname{bc}: \operatorname{spt}\gamma \mapsto M$. A result of Calier and Ekeland [7, proof of Proposition 3] implies that $\nu := \operatorname{bc}\#\gamma$ is the unique Wasserstein barycenter.

Now fix $x_2, \dots, x_m$, set $\operatorname{bc}(E) := \operatorname{bc}(E, x_2, \dots, x_m)$; then $\operatorname{bc}(M)$ is the set of barycenters in M. For each $y \in \operatorname{bc}(M)$, there exists $x_1 \in M$ such that $y \in \operatorname{bc}(x_1)$. By (3.9), we know that x_1 is uniquely determined. Then we can define a map $G: \operatorname{bc}(M) \mapsto M$ by letting $G(y) = x_1$, which can be seen as the inverse of bc.

Lemma 4.2 *$G: \operatorname{bc}(M) \mapsto M$ is continuous.*

Proof For any $y \in \operatorname{bc}(M)$, let $u = G(y)$. For any sequence $\operatorname{bc}(M) \ni y_k \to y$, let $u_k = G(y_k)$. If $u_k \nrightarrow u$, by the compactness of M, there exists a subsequence of u_k (still denoted by u_k for simplicity) such that $u_k \to v \neq u$. Since $x_1 \mapsto c(x_1, \dots, x_m)$ is Lipschitz,

$$c(u_k, x_2, \dots, x_m) \longrightarrow c(v, x_2, \dots, x_m).$$

Note that

$$c(u_k, x_2, \dots, x_m) = \frac{d^2}{2}(u_k, y_k) + \sum_{i=2}^{m} \frac{d^2}{2}(x_i, y_k) \to \frac{d^2}{2}(v, y) + \sum_{i=2}^{m} \frac{d^2}{2}(x_i, y);$$

then $y \in \mathrm{bc}(\nu)$, which contradicts to the fact that $G(y)$ is unique. Thus, we have proved that G is continuous. ■

Denote by D the diameter of M. Our main result is the following theorem.

Theorem 4.3 *G is Lipschitz with a Lipschitz constant C_D only depending on D.*

To make our argument more understandable, we first explain the idea of our proof. Given two points $y_1, y_2 \in \mathrm{bc}(M)$, let $u = G(y_1)$ and $v = G(y_2)$. To simplify our explanation, we assume that y_1, y_2 are regular points and $d^2(u,\cdot)$, $d^2(v,\cdot)$, $d^2(x_2,\cdot),\dots,d^2(x_m,\cdot)$ are differentiable at y_1, y_2. Note that this assumption does not hold for all barycenters. We do not even know whether such a point exists. Let

$$f_1 := \frac{d^2}{2}(u,\cdot) + \sum_{i=2}^{m} \frac{d^2}{2}(x_i,\cdot), \qquad f_2 := \frac{d^2}{2}(v,\cdot) + \sum_{i=2}^{m} \frac{d^2}{2}(x_i,\cdot).$$

By Lemma 2.6, we have $\nabla f_1(y_1) = 0$, $\nabla f_2(y_2) = 0$.

Then on T_{y_2}, we obtain

$$\begin{aligned} \overrightarrow{uv} = \overrightarrow{y_2 v} - \overrightarrow{y_2 u} &= -\nabla_{y_2}\frac{d^2}{2}(v,\cdot) - \Big(-\nabla_{y_2}\frac{d^2}{2}(u,\cdot)\Big) \\ &= \nabla_{y_2}\frac{d^2}{2}(u,\cdot) + \nabla f_2(y_2) - \nabla_{y_2}\frac{d^2}{2}(v,\cdot) = \nabla f_1(y_2). \end{aligned} \tag{4.1}$$

Since $d(u,v) \leqslant L_D|\overrightarrow{uv}|$ (see Lemma 4.5 below), if we can prove

$$|\nabla f_1(y_2)| \leqslant C d(y_1,y_2)$$

for some constant C only depending on D, then we are done. To estimate $|\nabla f_1(y_2)|$, we use the λ-concavity of distance functions. By Lemma 2.5, f_i is λ_D-concave for each i, where $\lambda_D := mD\frac{\cosh D}{\sinh D}$. Let

$$z_2 = \exp_{y_2}\Big(-d(y_1,y_2)\frac{\nabla f_1(y_2)}{|\nabla f_1(y_2)|}\Big);$$

then we have

$$f_1(z_2) - f_1(y_2) \leqslant -d(y_1,y_2)|\nabla f_1(y_2)| + \frac{\lambda_D}{2}d^2(y_1,y_2). \tag{4.2}$$

Since $f_1(z_2) \geqslant f_1(y_1)$, we have

$$f_1(y_2) - f_1(z_2) \leqslant f_1(y_2) - f_1(y_1) \leqslant \frac{\lambda_D}{2}d^2(y_1,y_2). \tag{4.3}$$

The second inequality holds, since $\nabla f_1(y_1) = 0$. By combining (4.1) and (4.2) with (4.3), we have

$$|\overrightarrow{uv}| = |\nabla f_1(y_2)| \leqslant \lambda_D d(y_1,y_2).$$

Remark 4.4 One may think that we can approximate the barycenters by regular points. However, we do not know whether the subset $\mathrm{bc}(M) \cap \mathrm{Reg}(M)$ is dense in $\mathrm{bc}(M)$. One may wonder whether we can use the Lipschitz character of gradient curves relative to a semiconcave function (see [26]). However, $G(y)$ is not on the gradient curve of $\sum_{i=2}^{m} d^2/2(x_i,\cdot)$ starting from $y \in \mathrm{bc}(M)$. It seems to me that at least we cannot use gradient curve directly. So we use Petrunin's method to perturb f_1, f_2 such that they achieve a minimum at a "good" point.

Now we begin our proof.

Proof We divide the proof into two steps:

Step 1. Perturb the functions to achieve the minimums at points where the square of distance functions are differentiable.

Now given any two points $y_1, y_2 \in \mathrm{bc}(M)$, let $u = G(y_1)$ and $v = G(y_2)$. We suppose that $y_1, y_2 \neq u, v, x_2, \ldots, x_m$. For $i = 1, 2$, by Lemma 2.9, there exists a neighborhood $U_i \ni y_i$; we can choose Perelman's concave function h_i defined on U_i. Let

$$f_1(z) = \frac{d^2}{2}(u, z) + \sum_{i=2}^{m} \frac{d^2}{2}(x_i, z) \quad \text{and} \quad f_2(z) = \frac{d^2}{2}(v, z) + \sum_{i=2}^{m} \frac{d^2}{2}(x_i, z).$$

By Lemma 2.5, f_i is $\lambda_D := mD\frac{\cosh D}{\sinh D}$-concave for each i. Let $\overline{f}_i(z) = f_i(z) + d^2(y_i, z)$. Then for any $r > 0$ with $B_r(y_i) \subset U_i$, there exists $C_r > 0$, such that when $z \in U_i \backslash B_r(y_i)$, $\overline{f}_i(z) - \overline{f}_i(y_i) \geqslant C_r$. Since h_i is 2-Lipschitz, we can choose a sufficiently small $\epsilon > 0$ with $0 < \epsilon < \frac{r}{100}$, such that $\overline{f}_i + \epsilon h_i$ achieves a strict minimum at some point $\overline{y}_i \in B_r(y_i)$. By Lemma 2.9, $\overline{y}_i$ is a regular point.

Now we choose a coordinate system near $\overline{y}_1$ by semiconcave functions $g_1, g_2, \ldots, g_n$ and another coordinate system near $\overline{y}_2$ by semiconcave functions $g_{n+1}, \ldots, g_{2n}$. Denote by $Y \subset M$ the set of points where

$$\begin{aligned} &d(u,\cdot), d(v,\cdot), d^2(u,\cdot), d^2(v,\cdot), d^2(x_2,\cdot), \ldots \\ &\qquad\qquad d^2(x_m,\cdot), d^2(y_1,\cdot), d^2(y_2,\cdot), g_1, \ldots, g_{2n}, \end{aligned}$$

are all differentiable. By Rademacher's theorem, $\mathcal{H}^n(Y) = \mathcal{H}^n(M)$. By Lemma 2.10, there exist small positive numbers $a_1, \ldots, a_{2n}$ with $0 < a_i < \frac{r}{100n}$ such that H_i achieves a minimum at $y_i^* \subset Y \cap B_r(y_i)$, where

$$H_1 := \overline{f}_1 + \epsilon h_1 + \sum_{i=1}^{n} a_i g_i \quad \text{and} \quad H_2 := \overline{f}_2 + \epsilon h_2 + \sum_{i=n+1}^{2n} a_i g_i.$$

Then by Lemma 2.4, there exists a unique geodesic connecting y_2^* to u, a unique geodesic connecting y_2^* to v.

Note that

$$\begin{aligned} \nabla f_1(y_2^*) - \nabla f_2(y_2^*) &= \nabla_{y_2^*} \frac{d^2}{2}(u, y_2^*) - \nabla_{y_2^*} \frac{d^2}{2}(v, y_2^*) \\ &= -\uparrow_{y_2^*}^{u} - (-\uparrow_{y_2^*}^{v}) = \overrightarrow{uv}, \end{aligned} \tag{4.4}$$

where $\overrightarrow{uv} := \uparrow_{y_2^*}^{v} - \uparrow_{y_2^*}^{u} \in T_{y_2^*}$.

Step 2. Estimate $|\overrightarrow{uv}|$ and prove that G is Lipschitz.

https://doi.org/10.4153/CJM-2016-035-8 Published online by Cambridge University Press

Note that H_i is semi-concave in $B_r(y_i)$, by Lemma 2.6, we have $\nabla H_i(y_i^*) = 0$. For $i = 1, 2$, let

$$w_1 := -\nabla_{y_1^*} d^2(y_1, y_1^*) - \epsilon \nabla h_1(y_1^*) - \sum_{i=1}^{n} a_i \nabla g_i(y_1^*),$$
$$w_2 := -\nabla_{y_2^*} d^2(y_2, y_2^*) - \epsilon \nabla h_2(y_2^*) - \sum_{i=n+1}^{2n} a_i \nabla g_i(y_2^*).$$

Then we have

$$\nabla f_i(y_i^*) = -w_i. \tag{4.5}$$

By combining this and (4.4), we obtain

$$\overrightarrow{uv} = \nabla f_1(y_2^*) + w_2. \tag{4.6}$$

Note that by (2.4), we know that g_i is 2-Lipschitz for all $1 \leqslant i \leqslant 2n$. Then we have

$$|w_i| \leqslant 2r + 2\epsilon + 2n \frac{r}{100n} \leqslant 3r. \tag{4.7}$$

Next we estimate $|\nabla f_1(y_2^*)|$. Suppose $|\nabla f_1(y_2^*)| \neq 0$ and

$$z_2^* = \exp_{y_2^*}\Big(-|y_1^* y_2^*| \frac{\nabla f_1(y_2^*)}{|\nabla f_1(y_2^*)|} \Big).$$

Since f_1 achieves a local minimum at y_1, we have

$$f_1(z_2^*) - f_1(y_1^*) = f_1(z_2^*) - f_1(y_1) + f_1(y_1) - f_1(y_2^*) \geqslant f_1(y_1) - f_1(y_1^*). \tag{4.8}$$

Since $\overline{f}_1 + \epsilon h_1$ achieves a local minimum at $\overline{y}_1$, we have

$$f_1(\overline{y}_1) + d^2(y_1, \overline{y}_1) + \epsilon h_1(\overline{y}_1) \leqslant f_1(y_1) + \epsilon h_1(y_1). \tag{4.9}$$

Since H_1 achieves a local minimum at y_1^*, we have

$$f_1(y_1^*) + d^2(y_1, y_1^*) + \epsilon h_1(y_1^*) + \sum_{i=1}^{n} a_i g_i(y_1^*) \leqslant \\ f_1(\overline{y}_1) + d^2(y_1, \overline{y}_1) + \epsilon h_1(\overline{y}_1) + \sum_{i=1}^{n} a_i g_i(\overline{y}_1). \tag{4.10}$$

Let

$$c_1 = d^2(y_1, y_1^*) + \epsilon h_1(y_1^*) - \epsilon h_1(\overline{y}_1) + \sum_{i=1}^{n} a_i g_i(y_1^*) - \sum_{i=1}^{n} a_i g_i(\overline{y}_1);$$

then we have

$$|c_1| \leqslant r^2 + 2\epsilon r + 2n \frac{r}{100n} \leqslant r. \tag{4.11}$$

By combining (4.8), (4.9), and (4.10), we obtain

$$f_1(z_2^*) - f_1(y_1^*) \geqslant c_1. \tag{4.12}$$

Since f_1 is λ_D-concave, we have

$$\begin{aligned} f_1(y_2^*) - f_1(y_1^*) &\leqslant \big\langle \nabla f_1(y_1^*), \uparrow_{y_1^*}^{y_2^*} \big\rangle d(y_1^*, y_2^*) + \frac{\lambda_D}{2} d^2(y_1^*, y_2^*) \\ &= -\langle w_1, \uparrow_{y_1^*}^{y_2^*} \rangle d(y_1^*, y_2^*) + \frac{\lambda_D}{2} d^2(y_1^*, y_2^*), \end{aligned} \tag{4.13}$$

https://doi.org/10.4153/CJM-2016-035-8 Published online by Cambridge University Press

the equality following from (4.5).

By combining (4.12) with (4.13), we obtain

$$f_1(z_2^*) - f_1(y_2^*) \geqslant c_1 + \langle w_1, \uparrow_{y_1^*}^{y_2^*} \rangle d(y_1^*, y_2^*) - \frac{\lambda_D}{2} d^2(y_1^*, y_2^*). \tag{4.14}$$

On the other hand, since f_1 is λ_D-concave,

$$f_1(z_2^*) - f_1(y_2^*) \leqslant -d(y_1^*, y_2^*)|\nabla f_1(y_2^*)| + \frac{\lambda_D}{2} d^2(y_1^*, y_2^*). \tag{4.15}$$

By combining (4.14) with (4.15), we obtain

$$|\nabla f_1(y_2^*)| \leqslant \lambda_D d(y_1^*, y_2^*) - \langle w_1, \uparrow_{y_1^*}^{y_2^*} \rangle + \frac{c_1}{d(y_1^*, y_2^*)}.$$

This inequality, together with (4.6) imply that

$$|\overrightarrow{uv}| \leqslant \lambda_D d(y_1^*, y_2^*) + |w_1| + |w_2| + \frac{c_1}{d(y_1^*, y_2^*)}.$$

By combining (4.11) with (4.7), we have

$$|\overrightarrow{uv}| \leqslant \lambda_D d(y_1^*, y_2^*) + 6r + \frac{r}{d(y_1^*, y_2^*)}.$$

By the following lemma, we have

$$\begin{aligned} |uv| \leqslant L_D |\overrightarrow{uv}| &\leqslant L_D \Big(\lambda_D d(y_1^*, y_2^*) + 6r + \frac{r}{d(y_1^*, y_2^*)} \Big) \\ &\leqslant L_D \Big[\lambda_D d(y_1, y_2) + \Big(2\lambda_D + 6 + 2\frac{1}{d(y_1, y_2)} \Big) r \Big], \end{aligned}$$

where L_D is a constant depending only on D.

By the arbitrariness of r, we have $|uv| \leqslant L_D \lambda_D d(y_1, y_2)$.

Let $C_D = L_D \lambda_D$. If $y_2 = x_i$ for some $2 \leqslant i \leqslant m$, we just remove the term $\frac{d^2}{2}(x_i, z)$ from the function $f_1(z), f_2(z)$. If $y = u$ ($y = v$), we remove the term $\frac{d^2}{2}(u, z)$ ($\frac{d^2}{2}(v, z)$) from $f_1(z)$ ($f_2(z)$). In these cases, f_1 is $(m-1)D\frac{\cosh D}{\sinh D}$-concave; thus, in particular λ_D-concave. Then we can repeat the same argument as above and get the same result. ∎

The following lemma is well known; see [5, proposition 10.6.10] for a proof.

Lemma 4.5 *Let M be a compact, n-dimensional Alexandrov space with curvature at least -1. For any three points $p, q, r \in M$ such that there exists a unique geodesic connecting p to $q(r)$. If curvature $\geqslant k > 0$, we add assumption $d(p,q) + d(p,r) + d(q,r) < \frac{2\pi}{\sqrt{k}}$. Let $P = o_p$, $Q = |pq|\uparrow_p^q$, $R = |pr|\uparrow_p^r$ in T_pM. Then there exists a constant L_D only depending on the diameter D such that $d(q,r) \leqslant L_D|QR|$.*

Remark 4.6 If M has curvature $\geqslant 0$, then L_D can be 1; *i.e.*, the exponential map in non-expanding. In this case, for any $p \in M$, the function $d^2/2(p, \cdot)$ is 1-concave, then $\lambda_D = m$, thus $C_D = L_D \lambda_D = m$. This constant is sharp, since in $\mathbb{R}^n$ with the standard metric, $d(u,v) = md(y_1, y_2)$.

Now we prove Theorem 1.3.

Proof of Theorem 1.3 Since we have proved that G is Lipschitz, we can follow the same lines of the proofs of [13, Lemmas 5.3 & 5.4 and Theorem 5.1, p. 26], and finally prove Theorem 1.3. ∎

Next we give a proposition that highlights the relationship between the barycenter and the optimal maps in Theorem 1.2.

Proposition 4.7 *Under the assumption of Theorem 1.3, the optimal maps F_i are of the form $G_i \circ G_1^{-1}$, where G_1^{-1} is the optimal map pushing μ_1 forward to ν for the quadratic cost $d^2(x_i, y)$, and G_i is the optimal map pushing ν forward to μ_i.*

The proof is quite similar to the proof of [14, Proposition 5.1]. Just note that since we have proved the absolutely continuity of Wasserstein barycenter, G_i are optimal maps pushing ν to μ_i.

A Existence and Uniqueness: the Non-compact Case

In this appendix, we prove Theorem 1.2 for non-compact Alexandrov spaces. The proof is similar to that of [4, theorem 4.2]. We refer the reader to [3, thm 6.2.4] for a detailed proof in the Euclidean case. Recall that a point $x \in M$ is a Lebesgue point of a function f if

$$\lim_{r\to 0} \frac{1}{\mathcal{H}^n(B_r(x))} \int_{B_r(x)} f(y) d\mathcal{H}^n(y) = \lim_{r\to 0} ⨍_{B_r(x)} f(y) d\mathcal{H}^n(y)$$

exists and coincides with $f(x)$. We also have a Lebesgue differentiation theorem for Alexandrov spaces. If $f: M \to \mathbb{R}$ is a locally integrable function, then H^n-a.e. $x \in M$ are Lebesgue points. Since for any measurable subset $U \subset M$, the characteristic function χ_U is locally integrable,

$$\lim_{r\to 0} \frac{\mathcal{H}^n(U \cap B_r(x))}{\mathcal{H}^n(B_r(x))} = 1 \text{ for } \mathcal{H}^n a.e. x \in U; \tag{A.1}$$

that is, $\mathcal{H}^n$-a.e. $x \in U$ have density 1 in U.

Since the potential function u_i may not be locally Lipschitz, we use the approximate differential of a map. We recall the definition in this setting.

Definition A.1 ([3, 4]) We say that $f: M \to \mathbb{R}$ has an approximate differential at $x \in \mathrm{Reg(M)}$ (denoted by $\widetilde{\nabla} f(x)$) if there exists a function $g: M \to \mathbb{R}$ which is differentiable at x such that the set $\{f \neq g\} := \{x \in M : f(x) \neq g(x)\}$ has density 0 at x.

Next, we prove Theorem 1.2 for the non-compact case.

Proof By Theorem 3.1, we have the existence of an optimal measure γ and a c-conjugate solution $(u_1, \dots, u_m)$ to (D) with $u_i \in L^1(\mu_i)$ for each $1 \leqslant i \leqslant m$ that satisfies (3.2). Set $S := \pi_1(\mathrm{spt}(\gamma))$; then $\mu_1(S) = 1$ and for any $x_1 \in S$, there exists

https://doi.org/10.4153/CJM-2016-035-8 Published online by Cambridge University Press

$(x_2,\ldots,x_m)\in M^{m-1}$ such that

$$\sum_{i=1}^{m} u_i(x_i)=c(x_1,\ldots,x_m). \tag{A.2}$$

Now fix a point $p\in M$, for any $k\in\mathbb{N}^+$, define functions

$$u_1^k(x_1):=\inf_{x_2,\ldots,x_m\in B_k(p)}\Big[c(x_1,\ldots,x_m)-\sum_{i=2}^{m}u_i(x_i)\Big].$$

By an argument similar to Lemma 3.2, we can prove that $c(x_1,\ldots,x_m)$ is locally Lipschitz on M^m and u_1^k is locally Lipschitz on M. Denote by

$$U_k:=\{x_1\in M: u_1^k \text{ is differentiable at } x_1\};$$

then we have $\mu_1(U_k)=1$ for each k.

For any $x_1\in S$, choose an $m-1$ tuple $(x_2,\ldots,x_m)$ satisfying (A.2) and a k_0 sufficiently large such that $x_2,\ldots,x_m\in B_{k_0}(p)$. Then for $k\geqslant k_0$, we have

$$c(x_1,\ldots,x_m)-\sum_{i=2}^{m}u_i(x_i)\geqslant u_1^{k_0}(x_1)\geqslant u_1^k(x_1)\geqslant u_1(x_1)=c(x_1,\ldots,x_m)-\sum_{i=2}^{m}u_i(x_i).$$

It follows that $u_1^k(x_1)=u_1(x_1)$ for any $k\geqslant k_0$. Set $V_k=\{x_1\in S: u_1^k(x_1)=u_1(x_1)\}$; then we have $V_k\subset V_{k+1}$ and $\bigcup_{k=1}^\infty V_k=S$. It follows that

$$\lim_{k\to\infty}\mu_1(V_k)=\mu_1\Big(\bigcup_{k=1}^{\infty}V_k\Big)=\mu_1(S)=1. \tag{A.3}$$

Set

$$W_k:=\{x_1\in V_k: x_1 \text{ has density 1 in } V_k\}.$$

If $V_k\neq\varnothing$, then by (A.1), $\mu_1(W_k)=\mu_1(V_k)$. By (A.3), we have

$$1\geqslant\mu_1\Big(\bigcup_{k=1}^{\infty}W_k\Big)\geqslant\lim_{k\to\infty}\mu_1(W_k)=\lim_{k\to\infty}\mu_1(V_k)=1.$$

It follows that $\mu_1(\bigcup_{k=1}^\infty W_k)=1$.

Set

$$A=\Big(\bigcap_{k=1}^{\infty}U_k\Big)\cap\Big(\bigcup_{k=1}^{\infty}W_k\Big)\cap S;$$

then $\mu_1(A)=1$.

For any $x_1\in A$, there exists $k\in\mathbb{N}^+$ such that $x_1\in W_k$. By Definition A.1, u_1 has approximate differential at x_1, and $\widetilde{\nabla u_1}(x_1)=\nabla u_1^k(x_1)$. By a similar argument as the proof of lemma 3.3, we can get that the mean y is uniquely determined by x_1 and

$$y=\begin{cases}x_1, & \text{if } \widetilde{\nabla u_1}(x_1)=0,\\ \exp_{x_1}\Big(-(f_1')^{-1}(|\widetilde{\nabla u_1}(x_1)|)\frac{\widetilde{\nabla u_1}(x_1)}{|\widetilde{\nabla u_1}(x_1)|}\Big) & \text{if } \widetilde{\nabla u_1}(x_1)\neq 0.\end{cases}$$

Note that Theorem 3.10 also holds on non-compact Alexandrov spaces; it follows that the m-1 tuple $(x_2,\ldots,x_m)$ with $(x_1,\ldots,x_m)\in\operatorname{spt}(\gamma)$ is uniquely determined by y, hence by $x_1\in A$. Following the last part of the proof of Theorem 1.2 for the compact case, we can get that Theorem 1.2 also holds on non-compact Alexandrov spaces. ∎

References

[1] M. Agueh and G. Calier, *Barycenters in the Wasserstein space*. SIAM J. Math. Anal. **43**(2011), no. 2, 904–924. http://dx.doi.org/10.1137/100805741
[2] L. Ambrosio and N. Gigli, *A user's guide to optimal transport*. In: Modelling and optimisation of flows on networks, Lecture Notes in Math., 2062, Springer, Heidelberg, 2013, 1–155. http://dx.doi.org/10.1007/978-3-642-32160-3_1
[3] L. Ambrosio, N. Gigli, and G. Savaré, *Gradient flows: in metric spaces and in the space of probability measures*. Lectures in Mathematics ETH Źurich, Birkh́auser Verlag, Basel, 2005.
[4] J. Bertrand, *Existence and uniqueness of optimal maps on Alexandrov spaces*. Adv. Math. **219**(2008), no. 3, 838–851. http://dx.doi.org/10.1016/j.aim.2008.06.008
[5] D. Burago, Y. Burago, and S. Ivanov, *A course in metric geometry*. Graduate Studies in Mathematics, 33, American Mathematical Society, Providence, RI, 2001. http://dx.doi.org/10.1090/gsm/033
[6] Y. Burago, M. Gromov, and G. Perel'man, *A. D. Alexandrov spaces with curvature bounded below*. Russian Math. Surveys **47**(1992), no. 2, 1–58. http://dx.doi.org/10.1070/RM1992v047n02ABEH000877
[7] G. Carlier and I. Ekeland, *Matching for teams*. Econom. Theory **42**(2010), no. 2, 397–418. http://dx.doi.org/10.1007/s00199-008-0415-z
[8] J. Cheeger, *Differentiability of Lipschitz functions on metric measure spaces*. Geom. Funct. Anal. **9**(1999), no. 3, 428–517. http://dx.doi.org/10.1007/s000390050094
[9] D. Cordero-Erausquin, R. McCann, and M. Schmuckenschläger, *A Riemannian interpolation inequality á la Borel, Brascamp and Lieb*. Invent. Math. **146**(2001), 219–257. http://dx.doi.org/10.1007/s002220100160
[10] A. Figalli and N. Juillet, *Absolute continuity of Wasserstein geodesics in the Heisenberg group*. J. Funct. Anal. **255**(2008), 133–141. http://dx.doi.org/10.1016/j.jfa.2008.03.006
[11] W. Gangbo and A. Święch, *Optimal maps for the multidimensional Monge-Kantorovich problem*. Comm. Pure Appl. Math. **51**(1998), 23–45. http://dx.doi.org/10.1002/(SICI)1097-0312(199801)51:1<23::AID-CPA2>3.0.CO;2-H
[12] H. G. Kellerer, *Duality theorems for marginal problems*. Z. Wahrsch. Verw. Gebiete **67**(1984), 399–432. http://dx.doi.org/10.1007/BF00532047
[13] Y.-H. Kim and B. Pass, *Wasserstein barycenters over Riemannian manifolds*. arxiv:1412.7726
[14] ———, *Multi-marginal optimal transport on Riemannian manifolds*. Amer. J. Math. **137**(2015), 1045–1060. http://dx.doi.org/10.1353/ajm.2015.0024
[15] R. McCann, *A convexity principle for interacting gases and equilibrium crystals.* Ph.D. thesis, Princeton University, 1994.
[16] ———, *Polar factorization of maps on Riemannian manifolds*. Geom. Funct. Anal. **11**(2001), 589–608. http://dx.doi.org/10.1007/PL00001679
[17] S. Ohta, *Barycenters in Alexandrov spaces of curvature bounded below*. Adv. Geom. **12**(2012), 571–587.
[18] Y. Otsu and T. Shioya, *The Riemannian structure of Alexandrov spaces*. J. Diferential Geom. **39**(1994), 629–658.
[19] B. Pass, *Uniqueness and Monge solutions in the multimarginal optimal transportation problem*. SIAM J. Math. Anal. **43**(2011), no. 6. 2758–2775. http://dx.doi.org/10.1137/100804917
[20] ———, *Multi-marginal optimal transport and multi-agent matching problems: uniqueness and structure of solutions*. Discrete Contin. Dyn. Syst. **34**(2014), no. 4, 1623–1639. http://dx.doi.org/10.3934/dcds.2014.34.1623
[21] ———, *Optimal transportation with infinitely many marginals*. J. Funct. Anal. **264**(2013), no. 4, 947–963. http://dx.doi.org/10.1016/j.jfa.2012.12.002
[22] G. Perelman, *Elements of Morse theory on Alexandrov spaces*. St. Petersburg Math. J. **5**(1994), 205–213.
[23] G. Perelman and A. Petrunin, *Quasigeodesics and gradient curves in Alexandrov spaces*. http://www.math.psu.edu/petrunin/
[24] A. Petrunin, *Subharmonic functions on Alexandrov space*. http://www.math.psu.edu/petrunin/
[25] ———, *Parallel transportation for Alexandrov spaces with curvature bounded below*. Geom. Funct. Anal. **8**(1998), 123–148. http://dx.doi.org/10.1007/s000390050050
[26] ———, *Semiconcave functions in Alexandrov's geometry*. Surv. Differ. Geom., 11, Int. Press, Somerville, MA, 2007, pp. 137–201. http://dx.doi.org/10.4310/SDG.2006.v11.n1.a6
[27] S. T. Rachev and L. Rüschendorf, *Mass transportation problems. Vol. I*. Probability and its Applications (New York), Springer-Verlag, New York, 1998, pp. 57–106.

[28] C. Villani, *Topics in optimal transportation*. Graduate Studies in Mathematics, 58, American Mathematical Society, Providence, RI, 2003. http://dx.doi.org/10.1007/b12016
[29] H. C. Zhang and X. P. Zhu, *Yau's gradient estimates on Alexandrov spaces*. J. Diferential Geom. **91**(2012), 445–522.

Department of Mathematics, Sun Yat-sen University, Guangzhou, 510275, China
e-mail: jiangy39@mail2.sysu.edu.cn