FREE DISCONTINUITIES IN OPTIMAL TRANSPORT

JUN KITAGAWA AND ROBERT MCCANN

Abstract. We prove a nonsmooth implicit function theorem applicable to the zero set of the difference of convex functions. This theorem is explicit and global: it gives a formula representing this zero set as a difference of convex functions which holds throughout the entire domain of the original functions. As applications, we prove results on the stability of singularities of envelopes of semi-convex functions, and solutions to optimal transport problems under appropriate perturbations, along with global structure theorems on certain discontinuities arising in optimal transport maps for Ma-Trudinger-Wang costs. For targets whose components satisfy additional convexity, separation, multiplicity and affine independence assumptions we show these discontinuities occur on submanifolds of the appropriate codimension which are parameterized locally as differences of convex functions (DC, hence C^2 rectifiable), and — depending on the precise assumptions — $C^{1,\alpha}$ smooth. In this case the highest codimension submanifolds consists of isolated points, each uniquely identified by the (affinely independent) components of the target to which it is transported.

Contents

2010 *Mathematics Subject Classification.* 35J96.

*JK's research was supported in part by a Simons Foundation Travel Grant and National Science Foundation grant DMS-1700094. RM's research was supported in part by NSERC grants 217006-08 and -15 and by a Simons Foundation Fellowship. Parts of this project were carried out while both authors were in residence at the Mathematical Sciences Reseach Institute in Berkeley CA during the Fall 2013 program supported by National Science Foundation Grant No. 0932078 000, and later at the Fields Insititute for the Mathematical Sciences in Toronto during Fall 2014.

1. Introduction

The question of regularity for maps solving the optimal transportation problem of Monge and Kantorovich is a celebrated problem [32] [36]. Under strong hypotheses relating the target's convexity to curvature properties of the transportation cost, optimal maps are known to be smooth, following work of Caffarelli on quadratic costs [5] and Ma, Trudinger, and Wang more generally [27]. In the absence of such convexity and curvature properties, much less is true. Partial regularity results — which quantify the size of the singular set — are available in at least three flavors. The set of discontinuities of an optimal map is known to be contained in the non-differentiabilities of a (semi-)convex function, hence to have Hausdorff dimension at most $n-1$ in $\mathbf{R}^n$. In fact, Zajíček [40] has shown such discontinuities lie in a countable union of submanifolds parameterized as graphs of differences of convex functions — referred to as DC submanifolds hereafter. The *closure* of this set of discontinuities was shown to have zero volume by Figalli with Kim (for quadratic costs [14]) or with DePhilippis (for non-degenerate costs [12]), and is conjectured to have dimension at most $n-1$. However, this conjecture has only been verified in the special case of a quadratic transportation cost on $\mathbf{R}^2$ [13]. See related work of Chodosh et al [9] and Goldman and Otto [20]. The present manuscript is largely devoted to providing evidence for this conjecture in higher dimensions by providing concrete geometries in which it can be confirmed. Typically these consist of transportation to a collection of disjoint target components, which we allow to be convex or non-convex. This forces discontinuities along which the optimal map tears the source measure into separate components, one corresponding to each component of the target. We study the regularity of such tears. We show that when the target components can be separated by a hyperplane, the corresponding tear is a DC hypersurface. For quadratic costs, when several tears meet, their intersection is a DC submanifold of the appropriate codimension provided the corresponding target components are affinely independent. When the corresponding target components are strictly convex, we show the tears are $C^{1,\alpha}$ smooth, and that the optimal maps are smooth on their complement. We show stability of such tears when the data are subject to perturbations which are small in a sense made precise below.

A core result of this paper is a nonsmooth version of the classical implicit function theorem for convex functions. More specifically, we wish to write the set where two convex functions coincide as the graph of a DC function, where DC stands for difference of convex, alternately denoted $c-c$ [19] or Δ-convex [34] in some references. The idea of inverse and implicit function theorems have been explored in various nonsmooth settings, e.g. by Clarke [10] and Vesely and Zajíček [34, Proposition 5.9]; see also [39] [28, Appendix] [37, Theorem 10.50]. Two major aspects set apart the version we present here from previous theorems. The first is the explicit nature of the theorem: we are able to explicitly write down the function whose graph gives the coincidence set in terms of partial Legendre transforms of the original convex functions, thus we term this an "explicit function theorem" in contrast to the traditional implicit version. Second, our result is of a global, rather than a local nature: existing implicit function theorems generally state the existence of a neighborhood on which a surface can be written as the graph of a function, in our theorem we obtain that the domain of this function is actually the projection of the entire original domain on some hyperplane. Our method of proof relies on the construction of Alberti from [1, Lemma 2.7], foreshadowed in Zajíček's work [40].

Our interest in this theorem is motivated by its application to the *optimal transport* problem of Monge and Kantorovich mentioned above. Let Ω and $\overline{\Omega}$ be compact subsets of n-dimensional Riemannian manifolds (M, g) and $(\bar{M}, \bar{g})$ respectively, and a real valued *cost function* $c \in C^4(\Omega \times \overline{\Omega})$. The optimal transport problem is: given any two probability measures μ and ν on Ω and $\overline{\Omega}$ respectively, find a measurable mapping $T : \operatorname{spt}\mu \to \operatorname{spt}\nu$ pushing μ forward to ν (denoted $T_{\#}\mu = \nu$), such that

$$\int_{\Omega} c(x, T(x))\mu(dx) = \inf_{S_{\#}\mu=\nu} \int_{\Omega} c(x, S(x))\mu(dx). \tag{OT}$$

The applications we present here concern the global structure of discontinuities in T, stability results for such tears, and the regularity of T on their complement. For the first application, we ask if there is some structure for these discontinuities when the support of the target measure is separated into two compact sets — by a hyperplane (in appropriate coordinates). One would expect the source domain to be partitioned into two sets, which are then transported to each of the pieces in the target. Under suitable hypotheses we show this is the case, and the interface between these two pieces is actually a DC hypersurface (thus C^2 rectifiable) which can be parameterized as a globally Lipschitz graph. In the second application, we consider a target measure consisting of several connected components. This should result in a transport map that must split mass amongst the pieces, and we investigate the structure and stability of this splitting. It turns out a stability result can be obtained when considering perturbations of the target measure under the Kantorovich-Rubinstein-Wasserstein L^∞ metric ($\mathcal{W}_\infty$ in Definition 8.1 below), along with an appropriate notion of affine independence for the pieces (Definition 4.11 below). We also provide an example to illustrate this independence condition plays the role of an implicit function hypothesis and is crucial for stability.

The outline of the paper is as follows. In Section 2 we set up and prove the "explicit function theorem" for convex differences. We then apply the explicit function theorem in Section 3, to show stability for singular points of envelopes of semi-convex functions under certain perturbations. In Section 4, we recall some necessary background material concerning the optimal transport problem and begin to explore consequences of known regularity results in our setting. For the quadratic cost $c(x, \bar{x}) = -\langle x, \bar{x}\rangle$ on Euclidean space, Section 5 proves DC rectifiability of the (codimension k) tears along which the source is split into $k+1$ components whose images have affinely independent convex hulls. For $k = n$, Proposition 5.5 shows the corresponding tear consists of a single point. Section 6 shows these tears are $C^{1,\alpha}$ provided the corresponding target components are strictly convex; in the simplest case $k = 1$, a similar result was found by Chen [7] simultaneously and independently of the present manuscript: the main thrust of his work is to improve regularity of the tear to $C^{2,\alpha}$ when the pair of strictly convex target components are sufficiently far apart. Smoothness of the map away from such tears is shown for Ma-Trudinger-Wang costs — known as MTW costs [27] [33] — in Corollary 4.10. Section 7 extends our DC rectifiability result for tears to MTW costs in the prototypical case $k = 1$. Section 8 shows such tears are stable. Lastly, we include an appendix presenting an example to show the affine independence of target measures components is necessary for stability.

Throughout this paper, for $1 \leq i \leq n$ we will use the notation $\pi_i : \mathbf{R}^n \to \mathbf{R}^i$ to denote orthogonal projection onto the first ith coordinates, and e_i for the ith unit coordinate vector. We also reserve the notation A^{cl}, A^{int}, and A^{∂} for the closure, interior, and boundary of a set A respectively. Also, given any point $x \in \mathbf{R}^n$, we will write x^i for the ith coordinate of x. $\mathcal{H}^i$ will refer to the i-dimensional Hausdorff measure of a set in Euclidean space and $\mathcal{H}^i_g$ will be the i-dimensional Hausdorff measure of a set defined using the distance derived from a Riemannian metric g. Finally, $\mathrm{conv}(A)$ denotes the *closed* convex hull of a set A while $\mathcal{N}_\varepsilon(A) = \{x \mid dist(x, A) \leq \varepsilon\}$.

2. An "explicit function theorem" for convex differences

For the remainder of the paper, by *convex function* with no other qualifiers we will tacitly mean a *closed, proper, convex function on* $\mathbf{R}^n$ i.e., a function defined on $\mathbf{R}^n$ taking values in $\mathbf{R} \cup \{\infty\}$, whose epigraph is a non-empty, closed, convex set. If we refer to *a convex function on* Λ for some set $\Lambda \subset \mathbf{R}^n$, this will mean a function satisfying the above definition when it is extended lower semicontinuously to Λ^{cl} and (re)defined to be ∞ on $(\mathbf{R}^n \setminus \Lambda)^{\mathrm{int}}$. Also, we will use the notations $x' := \pi_{n-1}(x)$ and $A' := \pi_{n-1}(A)$ for any point $x \in \mathbf{R}^n$ and set $A \subset \mathbf{R}^n$. By the classical implicit function theorem, if f, $g : \mathbf{R}^n \to \mathbf{R}$ are smooth, the set $\{f = g\}$ is the graph of a smooth function of $n-1$ variables, near any point on the set where $\nabla f \neq \nabla g$. We aim to prove an analogue of this theorem, but for two convex functions without any assumptions of differentiability. In order to do so, we need an appropriate replacement for the inequality of gradients, which will be formulated in terms of the *subdifferential*: recall for a convex function u and x_0 in its domain,

$$\partial u\,(x_0) := \{\bar{x} \in \mathbf{R}^n \mid \langle x - x_0, \bar{x}\rangle + u(x_0) \leq u(x),\ \forall x\}, \tag{2.1}$$

while for a subset A of its domain,

$$\partial u\,(A) := \bigcup_{x \in A} \partial u\,(x).$$

We also recall here the *Legendre transform* of a (proper) convex function u with effective domain $\mathrm{Dom}(u) := \{x \in \mathbf{R}^n \mid u(x) < \infty\}$ as the (closed, proper, convex) function $u^* : \mathbf{R}^n \to \mathbf{R} \cup \{\infty\}$ defined by

$$u^*(\bar{x}) := \sup_{x \in \mathbf{R}^n} [\langle x, \bar{x}\rangle - u(x)] = \sup_{x \in \mathrm{Dom}(u)} [\langle x, \bar{x}\rangle - u(x)]. \tag{2.2}$$

Definition 2.1 (Separating hyperplane)**.** If Λ_+ and Λ_- are any two sets in $\mathbf{R}^n$ and v is a fixed unit vector, recall that a hyperplane $\{x \in \mathbf{R}^n \mid \langle x, v\rangle = a\}$ is said to *strongly separate* Λ_+ *and* Λ_- *(with spacing d)* if there exists a $d > 0$ such that

$$\langle x_1, v\rangle < a - d < a + d < \langle x_2, v\rangle$$

for any $x_1 \in \Lambda_+$ and $x_2 \in \Lambda_-$.

Let us also recall some terminology on DC (difference of convex) functions here.

Definition 2.2 (DC functions, mappings [2,30])**.** A function $h : \Lambda \to \mathbf{R}$ on a convex domain $\Lambda \subset \mathbf{R}^n$ is said to be a *DC function* if it can be written as the difference of two convex functions that are finite on Λ. A mapping from Λ to a Euclidean space $\mathbf{R}^m$ is said to be a *DC mapping* if each of its coordinate components is a DC function.

The key hypothesis of our theorem is the strong separation of the subdifferentials of two convex functions. One feature that differentiates our theorem from the usual implicit function theorem is that we can actually write down the function whose graph gives the equality set between the two convex functions we consider, and explicitly state the domain of this function. Thus we term this an "explicit function theorem." We first state the following Theorem 2.3 in terms of the subdifferential of the envelope of two convex functions, and formulate the actual explicit function theorem as Corollary 2.6 below.

Theorem 2.3 (DC tears)**.** *Let u_+ and u_- be convex functions, $\Lambda \subset \mathrm{Dom}(u) \subset \mathbf{R}^n$ a convex (but not necessarily bounded) set, and $\overline{\Lambda}_+$, $\overline{\Lambda}_-$ compact subsets of $\mathbf{R}^n$ with $\partial u_+(\Lambda) \subset \overline{\Lambda}_+$ and $\partial u_-(\Lambda) \subset \overline{\Lambda}_-$. We define*

$$
\begin{aligned}
u &:= \max\{u_+, u_-\},\\
\Sigma &:= \left\{x \in \Lambda^{\mathrm{cl}} \mid \partial u(x) \cap \overline{\Lambda}_+ \neq \emptyset \text{ and } \partial u(x) \cap \overline{\Lambda}_- \neq \emptyset\right\},\\
C_+ &:= \left\{x \in \Lambda^{\mathrm{cl}} \mid \partial u(x) \cap \overline{\Lambda}_- = \emptyset\right\},\\
C_- &:= \left\{x \in \Lambda^{\mathrm{cl}} \mid \partial u(x) \cap \overline{\Lambda}_+ = \emptyset\right\}.
\end{aligned}
$$

Also, suppose that (after a rotation of coordinates) for some $a_0 \in \mathbf{R}$ the hyperplane $\Pi := \{x^n = a_0\}$ strongly separates $\overline{\Lambda}_+$ and $\overline{\Lambda}_-$ with spacing $d_0 > 0$.

Writing $\Lambda' := \pi_{n-1}(\Lambda)$, define the functions $h^\pm : \mathbf{R}^{n-1} \to \mathbf{R}$, $h : (\Lambda')^{\mathrm{cl}} \to \mathbf{R}$ by

$$
h^\pm(x') := \begin{cases} -\frac{u^*_{x'}(a_0 \mp d_0)}{2d_0}, & x' \in (\Lambda')^{\mathrm{cl}},\\ \infty, & x' \in \mathbf{R}^{n-1} \setminus (\Lambda')^{\mathrm{cl}} \end{cases} \tag{2.3}
$$

$$
h(x') := h^+(x') - h^-(x'), \tag{2.4}
$$

where $u^*_{x'}$ is the Legendre transform of the function $u_{x'}(t) := u(x', t)$ of one variable. Then $h^\pm$ are both convex on $\mathbf{R}^{n-1}$ and finite on Λ' (so in particular, h is a DC function), with

$$
\begin{aligned}
\Sigma &= \{(x', h(x')) \mid x' \in \Lambda'\} \cap \Lambda^{\mathrm{cl}},\\
C_+ &= \{(x', x^n) \mid x' \in \Lambda',\ h(x') < x^n\} \cap \Lambda^{\mathrm{cl}},\\
C_- &= \{(x', x^n) \mid x' \in \Lambda',\ h(x') > x^n\} \cap \Lambda^{\mathrm{cl}}.
\end{aligned}
$$

Moreover,

$$
\|h\|_{\mathrm{Lip}\,((\Lambda')^{\mathrm{cl}})} \le \tan\Theta \le \frac{\mathrm{diam}[\pi_{n-1}(\overline{\Lambda}_+ \cup \overline{\Lambda}_-)]}{2d_0} \tag{2.5}
$$

where

$$
\cos\Theta := \inf_{\bar{x}_+ \in \overline{\Lambda}_+, \bar{x}_- \in \overline{\Lambda}_-} \langle \frac{\bar{x}_+ - \bar{x}_-}{|\bar{x}_+ - \bar{x}_-|}, e_n \rangle.
$$

Remark 2.4. Both functions $u_\pm$ can be extended in a continuous way to all of Λ^{cl}. Indeed, since $\partial u_\pm(\Lambda)$ is bounded, we can exhaust Λ by compact sets and apply [31, Theorem 24.7] to find that $u_\pm$ are uniformly Lipschitz on Λ; in particular they can be extended continuously to Λ^{cl} with finite values. Moreover, by compactness of $\overline{\Lambda}_\pm$ we see that $\partial u_\pm(x) \neq \emptyset$ for any $x \in \Lambda^\partial$ as well.

We will need the following classical result on subdifferentials of envelopes of convex functions (which can be obtained for example, by [11, Proposition 2.3.12] applied to convex functions).

Lemma 2.5. *If* $u = \max_i u_i$ *for some finite collection of convex functions* u_i*, then*

$$\partial u\,(x_0) = \operatorname{conv}\left(\bigcup_{i\in I} \partial u_i\,(x_0)\right)$$

where $I := \{i \mid u(x_0) = u_i(x_0)\}$.

Using this result, we find the following reformulation of Theorem 2.3.

Corollary 2.6 (Explicit function theorem)**.** *Under the same notation and hypotheses as Theorem 2.3,*

$$\begin{aligned}
\left\{x \in \Lambda^{\mathrm{cl}} \mid u_+(x) = u_-(x)\right\} &= \left\{(x', h(x')) \mid x' \in (\Lambda')^{\mathrm{cl}}\right\} \cap \Lambda^{\mathrm{cl}},\\
\left\{x \in \Lambda^{\mathrm{cl}} \mid u_+(x) > u_-(x)\right\} &= \left\{(x', x^n) \mid x' \in (\Lambda')^{\mathrm{cl}},\ h(x') < x^n\right\} \cap \Lambda^{\mathrm{cl}},\\
\left\{x \in \Lambda^{\mathrm{cl}} \mid u_+(x) < u_-(x)\right\} &= \left\{(x', x^n) \mid x' \in (\Lambda')^{\mathrm{cl}},\ h(x') > x^n\right\} \cap \Lambda^{\mathrm{cl}}.
\end{aligned}$$

Proof. Lemma 2.5 combined with Remark 2.4 immediately yields the corollary from Theorem 2.3. □

Proof of Theorem 2.3. Fix any such strongly separating hyperplane, by our assumptions we have $\overline{\Lambda}_+ \subset \{x^n > a_0 + d_0\}$ and $\overline{\Lambda}_- \subset \{x^n < a_0 - d_0\}$. Also, if $x' \in \Lambda'$, let us write $\Lambda^{x'} := \{t \in \mathbf{R} \mid (x', t) \in \Lambda\}$. By Remark 2.4, we can assume $u_\pm$ are both continuous up to Λ^{cl} which is also convex, thus we will tacitly assume Λ is a closed set for the remainder of the proof.

We first claim that given $x' \in \Lambda'$, there is at most one $x^n \in \Lambda^{x'}$ such that $(x', x^n) \in \Sigma$, and it must be that $x^n = h(x')$. Indeed, fix an $x' \in \Lambda'$ and suppose there exists such an x^n. First by [1, Proposition 2.4], for any $(x', t) \in \Lambda$ we have

$$\partial u_{x'}\,(t) = \pi^n\left(\partial u\,(x', t)\right). \tag{2.6}$$

As $\partial u\,(x', x^n)$ is convex and intersects both $\overline{\Lambda}_+$ and $\overline{\Lambda}_-$, we must have $[a_0 - d_0, a_0 + d_0] \subset \partial u_{x'}\,(x^n)$, which implies $x^n \in \partial u^*_{x'}\,([a_0 - d_0, a_0 + d_0])$ by [31, Theorem 23.5]. We also immediately see that the values $u^*_{x'}(a_0 \pm d_0)$ are both finite. By the definition of subdifferential, we have the inequalities

$$\begin{aligned}
u^*_{x'}(a_0 + d_0) &\geq u^*_{x'}(a_0 - d_0) + x^n(a_0 + d_0 - (a_0 - d_0)),\\
u^*_{x'}(a_0 - d_0) &\geq u^*_{x'}(a_0 + d_0) + x^n(a_0 - d_0 - (a_0 + d_0)),
\end{aligned}$$

which combined implies $x^n = h(x')$ defined by (2.3), and in particular there can only be at most one such x^n for each x'.

Now suppose $x' \in \Lambda'$ is such that $\Lambda^{x'} \neq \emptyset$ but there is no $t \in \Lambda^{x'}$ with $(x', t) \in \Lambda$ where $\partial u\,(x', t)$ intersects both of the sets $\overline{\Lambda}_\pm$. Note since Λ is convex the fiber $\Lambda^{x'}$ is connected. As the choice of cost function $c(x, \bar{x}) := -\langle x, \bar{x}\rangle$ satisfies conditions (B1) and (MTW) (see Section 4 below), we can apply Lemma 4.8 to see that $\partial u_{x'}\left(\Lambda^{x'}\right)$ is connected. We comment here, Lemma 4.8 does not directly apply if $\Lambda^{x'}$ is unbounded, but we can exhaust $\Lambda^{x'}$ with an increasing collection of bounded subintervals then take the union of their images under the subdifferential of $u_{x'}$ to obtain the claim. In particular by Lemma 2.5 (recalling (2.6)), either $\partial u_{x'}\left(\Lambda^{x'}\right) \subset [a_0 + d_0, \infty)$ or $\partial u_{x'}\left(\Lambda^{x'}\right) \subset (-\infty, a_0 - d_0]$, suppose it is the former; this is equivalent to having on the set $\Lambda^{x'}$,

$$u_{x'}(\cdot) \equiv u_+(x', \cdot). \tag{2.7}$$

Now we claim there exists a finite $t_0 \in \mathbf{R}$ such that

$$u_{x'}(t_0) = u_+(x', t_0) = u_-(x', t_0).$$

By (2.7), it is sufficient to show there is some t for which $u_-(x', t) > u_+(x', t)$, then then intermediate value theorem will finish the claim. Fix some $\tilde{t} \in \Lambda^{x'}$ and suppose the claim fails, then (2.7) would hold on all of $(-\infty, \tilde{t}]$. In turn, this means $u_+(x', t)$ is finite for all $t \leq \tilde{t}$, as if it was infinite anywhere the subdifferential of $u_+(x', \cdot)$ would contain an interval of the form $(-\infty, \bar{t})$ for some $\bar{t}$, contradicting (2.6) and the assumption $\partial u_+(\Lambda) \subset \overline{\Lambda}_+$. Now take a sequence $t_k \searrow -\infty$ where $u_-(x', t_k) \leq u_+(x', t_k)$, with $t_k < \tilde{t}$ for all k. By the above remark we can find $\bar{t}_k \in \pi^n(\partial u_+(x', t_k)) \subset (a_0 + d_0, \infty)$. Using [1, Proposition 2.4] we then have

$$u_+(x', t_k) \leq u_+(x', \tilde{t}) - \bar{t}_k(\tilde{t} - t_k) \leq u_+(x', \tilde{t}) - (a_0 + d_0)(\tilde{t} - t_k).$$

At the same time u_- is finite on Λ, hence there exists $\bar{t}_- \in \pi^n(\partial u_-\left(x', \tilde{t}\right)) \subset (-\infty, a_0 - d_0)$, again by [1, Proposition 2.4] we have

$$u_-(x', t_k) \geq u_-(x', \tilde{t}) + \bar{t}_-(t_k - \tilde{t}) \geq u_-(x', \tilde{t}) + (a_0 - d_0)(t_k - \tilde{t}),$$

thus

$$u_-(x', t_k) - u_+(x', t_k) \geq u_-(x', \tilde{t}) - u_+(x', \tilde{t}) + 2d_0(\tilde{t} - t_k) > 0$$

for large enough k, a contradiction, hence the claim is proven.

By Lemma 2.5 and (2.6) we can see that $a_0 + d_0 \in \partial u_{x'}(t_0)$, hence by [31, Theorem 23.5] we have

$$u^*_{x'}(a_0 + d_0) = t_0(a_0 + d_0) - u(x', t_0). \tag{2.8}$$

Since by definition

$$-u^*_{x'}(a_0 - d_0) = \inf_{t \in \mathbf{R}}(u(x', t) - t(a_0 - d_0)) \leq u(x', t_0) - t_0(a_0 - d_0),$$

we find that

$$h(x') \leq \frac{u(x', t_0) - t_0(a_0 - d_0) + t_0(a_0 + d_0) - u(x', t_0)}{2d_0} = t_0 \leq \inf \Lambda^{x'},$$

the last inequality from the fact that (2.7) holds on $\Lambda^{x'}$. The argument leading to (2.8) can also be applied to $u^*_{x'}(a_0 - d_0)$, since $u^*_{x'}$ is a proper convex function, an upper bound implies finiteness, hence $h^\pm$ are both finite valued for such x'. The case $\partial u_{x'}\left(\Lambda^{x'}\right) \subset (-\infty, a_0 - d_0]$ can be handled by a symmetric argument yielding that $h(x') \geq \sup \Lambda^{x'}$, and we find $h^\pm$ are both finite valued on all of Λ'. To show closedness of $h^\pm$, fix any $(x'_0, t_0) \in \mathbf{R}^n$. By [31, Theorem 7.1], u is lower semicontinuous on $\mathbf{R}^n$, thus for any $\varepsilon > 0$, there exists $\delta > 0$ such that $u(x'_0, t_0) \leq \varepsilon + \inf_{x' \in B_\delta(x'_0) \setminus \{x'_0\}} u(x', t_0)$, hence we have

$$\begin{aligned} -u^*_{x'_0}(a_0 \pm d_0) &\leq u(x'_0, t_0) - t_0(a_0 \pm d_0) \\ &\leq \inf_{x' \in B_\delta(x'_0) \setminus \{x'_0\}} u(x', t_0) - t_0(a_0 \pm d_0) + \varepsilon, \end{aligned}$$

taking an infimum over $t_0 \in \mathbf{R}$ shows that $h^\pm$ is lower semicontinuous, hence closed by [31, Theorem 7.1] again.

Next suppose $x \in \Lambda$ is such that $\partial u(x) \cap \overline{\Lambda}_- = \emptyset$, and there exists an $(x', t) \in \Lambda$ where $\partial u(x', t)$ intersects both of the sets $\overline{\Lambda}_\pm$. By the argument above, we must have $t = h(x')$. Take $\bar{x} \in \partial u(x)$ and $(\bar{y}', a_0) \in \partial u(x', h(x'))$. By monotonicity of the subdifferential we find that

$$\begin{aligned} 0 &\leq \langle x - (x', h(x')), \bar{x} - (\bar{y}', a_0) \rangle \\ &= (x^n - h(x'))(\bar{x}^n - a_0). \end{aligned}$$

However, by Lemma 2.5 and since $\partial u(x)$ does not intersect $\overline{\Lambda}_-$, we have must have $\bar{x}^n - a_0 \geq 0$, thus $x^n \geq h(x')$. A symmetric argument yields that if $\partial u(x) \cap \overline{\Lambda}_+ = \emptyset$, then $x^n \leq h(x')$. Since $\partial u(x', h(x'))$ intersects both sets $\overline{\Lambda}_\pm$, the above inequalities must be strict. Combined with the arguments above, this proves the characterizations of Σ, C_+, and C_- as the graph, epigraph, and subgraph of h intersected with Λ.

We will next show $h_\pm$ are both convex (essentially, this is just the fact that a supremum of a family of jointly convex functions gives a concave function). To this end, fix x_0', $x_1' \in \Lambda'$ and t_0, $t_1 \in \mathbf{R}$, and define $(x_\lambda', t_\lambda) := ((1-\lambda)x_0' + \lambda x_1', (1-\lambda)t_0 + \lambda t_1)$. Then $x_\lambda' \in \Lambda'$, hence $u^*_{x_\lambda'}(a_0 + d_0)$ is finite, in particular $h^\pm$ cannot take the value $-\infty$ anywhere and they must be proper. By the convexity of u, we can calculate

$$\begin{aligned} u^*_{x_\lambda'}(a_0 + d_0) &\geq t_\lambda(a_0 + d_0) - u(x_\lambda', t_\lambda) \\ &\geq (1-\lambda)t_0(a_0 + d_0) - (1-\lambda)u(x_0', t_0) + \lambda t_1(a_0 + d_0) - \lambda u(x_1', t_1), \end{aligned}$$

where the right hand sides of the second and third lines above may take the value $-\infty$. By taking a supremum on the right hand side, first over t_0, then over t_1, we obtain

$$u^*_{x_\lambda'}(a_0 + d_0) \geq (1-\lambda)u^*_{x_0'}(a_0 + d_0) + \lambda u^*_{x_1'}(a_0 + d_0),$$

then since Λ' is convex, the epigraph of h^+ will be a convex set. A similar argument for $u^*_{x_\lambda'}(a_0 - d_0)$ proves the epigraph of h^- is convex as well.

Lastly we prove the Lipschitz bound (2.5). To do so, we will show that any circular cone of slope $\tan\Theta$ opening in the positive or negative e_n direction, with vertex on the set $\Sigma \cap \Lambda$ remains on one side of Σ. Specifically, fix a point in $\Sigma \cap \Lambda$ and after a temporary translation, assume it is the origin. We claim that if $x^n \geq |x'| \tan\Theta$ with $x' \in \Lambda'$, then

$$h(x') \leq x^n. \tag{2.9}$$

Let us assume $h(x') \geq 0$, otherwise the above claim is immediate. First note that

$$\exists\, \bar{x}_\pm \in \overline{\Lambda}_\pm \text{ s.t. } \langle (x', h(x')), \bar{x}_+ - \bar{x}_- \rangle \leq 0 \implies \text{(2.9) holds.} \tag{2.10}$$

Indeed by the definition of Θ, this would imply that

$$\begin{aligned} 0 &\geq \langle x', \frac{\bar{x}_+' - \bar{x}_-'}{|\bar{x}_+ - \bar{x}_-|} \rangle + h(x') \left(\frac{\bar{x}_+^n - \bar{x}_-^n}{|\bar{x}_+ - \bar{x}_-|} \right) \\ &\geq \langle x', \frac{\bar{x}_+' - \bar{x}_-'}{|\bar{x}_+ - \bar{x}_-|} \rangle + h(x') \cos\Theta \end{aligned}$$

and rearranging terms,

$$\begin{aligned} h(x') &\leq \frac{1}{\cos\Theta}\langle -x', \frac{\bar{x}'_+ - \bar{x}'_-}{|\bar{x}_+ - \bar{x}_-|}\rangle \\ &\leq \frac{|x'|}{\cos\Theta}\frac{|\bar{x}'_+ - \bar{x}'_-|}{|\bar{x}_+ - \bar{x}_-|} \\ &\leq |x'|\tan\Theta \leq x^n, \end{aligned}$$

giving (2.9). Now let $\bar{x}_{0,\pm} \in \partial u_\pm(0)$ and $\tilde{\bar{x}}_\pm \in \partial u_\pm(x', h(x'))$; by Lemma 2.5 we have that $\bar{x}_{0,\pm} \in \partial u(0)$ and $\tilde{\bar{x}}_\pm \in \partial u(x', h(x'))$. In particular,

$$\begin{aligned} u(y) &\geq u(0) + \max\{\langle y, \bar{x}_{0,+}\rangle, \langle y, \bar{x}_{0,-}\rangle\}, \\ u(y) &\geq u(x', h(x')) + \max\{\langle y - (x', h(x')), \tilde{\bar{x}}_+\rangle, \langle y - (x', h(x')), \tilde{\bar{x}}_-\rangle\} \end{aligned}$$

for any y. Taking $y = (x', h(x'))$ in the first and $y = 0$ in the second inequality, plugging the second into the first and rearranging terms we obtain

$$\begin{aligned} \langle (x', h(x')), \tilde{\bar{x}}_-\rangle &\geq \min\{\langle (x', h(x')), \tilde{\bar{x}}_+\rangle, \langle (x', h(x')), \tilde{\bar{x}}_-\rangle\} \\ &\geq \max\{\langle (x', h(x')), \bar{x}_{0,+}\rangle, \langle (x', h(x')), \bar{x}_{0,-}\rangle\} \\ &\geq \langle (x', h(x')), \bar{x}_{0,+}\rangle. \end{aligned}$$

Thus we have (2.10), hence (2.9).

A symmetric argument can be used to show $x^n \leq h(x')$ whenever $x^n \leq -|x'|\tan\Theta$, as a result we obtain the Lipschitz bound (2.5). □

3. Stability of singularities

In this section, we will use the explicit function theorem from the previous section to show a stability result for singularities, we will extend our discussion from convex functions to semi-convex functions. First a few definitions.

Definition 3.1 (Semi-convexity)**.** Recall that a real valued function u defined on some $\Lambda \subset \mathbf{R}^n$ is said to be *semi-convex* if for any $x_0 \in \Lambda$, there exists a neighborhood of x_0 and some $C > 0$ for which the function $x \mapsto u(x) + C|x - x_0|^2$ is convex on that neighborhood. We will say that a family $\{u_j\}$ of semi-convex functions has *uniformly bounded constant of semi-convexity near* x_0 if there is some neighborhood of x_0 on which the same constant $C > 0$ can be chosen to make all of the functions $u_j + C|\cdot - x_0|^2$ convex on that neighborhood.

A function u defined on an open set in a smooth manifold is said to be *semi-convex* if the above definition holds near any point in a local coordinate chart.

Definition 3.2 (Subdifferential of a semi-convex function)**.** The *subdifferential* of a semi-convex function u defined on a subset of a Riemannian manifold (M, g) is defined by

$$\partial u(x_0) := \left\{p \in T^*_{x_0}M \mid u(\exp_{x_0}(v)) \geq u(x_0) + p(v) + o(|v|_g),\ \forall T_{x_0}M \ni v \to 0\right\}$$

where $\exp_{x_0}$ is the Riemannian exponential map.

If u is a convex function on a subdomain of $\mathbf{R}^n$, this definition is equivalent to (2.1).

Definition 3.3 (Legendre transform)**.** If u is a real-valued function defined on some subdomain $\mathrm{Dom}(u)$ of $\mathbf{R}^n$, its *Legendre transform* is the convex function defined by the equation (2.2) with the convention $u := \infty$ outside $\mathrm{Dom}(u)$.

It is well known that for a semi-convex function u, if $\partial u\,(x)$ is a singleton for some x, then u is actually differentiable at x. We will be interested in the behavior of u at points of *nondifferentiability*, namely we will be concerned with the *dimension* of $\partial u\,(x)$ (whenever we refer to the dimension of a convex set, we will always mean the dimension of its affine hull). In some sense, this dimension is a measure of how severe the singularity of u is at x: for example the function $|x|$ on $\mathbf{R}^n$ has an n dimensional subdifferential at the origin which corresponds to a conical singularity, while $|x^1|$ has a 1 dimensional subdifferential at the origin, and the function remains differentiable in the $\{x^1 = 0\}$ subspace.

In particular, we are interested in the stability of the dimension of the subdifferential of a sequence of semi-convex functions, as detailed in the following theorem, whose proof is deferred to the end of this section.

Theorem 3.4 (Stability of singularities)**.** *Suppose that u is a real valued function, finite on an open neighborhood $\mathcal{N}_{x_0}$ of some point $x_0 \in \mathbf{R}^n$, of the form*

$$u = \max_{1\leq i\leq K} u_i, \tag{3.1}$$

for some $K < \infty$ where all u_i are semi-convex. Also fix some $1 \leq k \leq \min\{K-1, n\}$ and assume that for any $1 \leq i \leq k+1$:

$$\begin{aligned} u_i &\in C^1(\mathcal{N}_{x_0}),\\ u(x_0) = u_i(x_0) &> u_{i'}(x_0),\ \forall\, k+2 \leq i' \leq K, \end{aligned}$$

and $\dim \partial u\,(x_0) = k$. Finally, let $\left\{u_i^j\right\}_{j=1}^{\infty}$ be a sequence for which each u_i^j is semi-convex with uniformly bounded constant of semi-convexity near x_0, $u_i^j \xrightarrow[j\to\infty]{} u_i$ uniformly in compact subsets of $\mathcal{N}_{x_0}$ for each $1 \leq i \leq K$, and write $u^j := \max_{1\leq i\leq K} u_i^j$. Then for any $\varepsilon > 0$, there exists an index J_ε such that for any $j > J_\varepsilon$, there exists a set $\Sigma^j_{n-k} \subset B_\varepsilon(x_0)$ with $\mathcal{H}^{n-k}\left(\Sigma^j_{n-k}\right) > 0$ on which

$$u^j(x) = u_i^j(x) > u_{i'}^j(x), \qquad \forall\, x \in \Sigma^j_{n-k},\ 1 \leq i \leq k+1,\ k+2 \leq i' \leq K. \tag{3.2}$$

Moreover, Σ^j_{n-k} is the graph of a DC mapping over an open set in $\mathbf{R}^{n-k}$ and

$$\dim \partial u^j\,(x) \geq k \qquad \forall x \in \Sigma^j_{n-k}, \tag{3.3}$$

with equality on a set of full $\mathcal{H}^{n-k}$ measure in Σ^j_{n-k}.

In preparation, we shall need a result on stability of the subdifferentials of a sequence of convergent convex functions. By a straightforward modification of the proof of [31, Theorem 25.7], we obtain the following lemma.

Lemma 3.5. *Suppose that u and $\{u_j\}_{j=1}^{\infty}$ are convex functions, finite and with $u_j \to u$ pointwise on some open convex domain Λ, and also assume that u is differentiable on Λ. Then for any compact $\Lambda_0 \subset \Lambda$ and $\varepsilon > 0$ there exists j_0 such that*

$$\partial u_j\,(x) \subset B_\varepsilon(\nabla u(x))$$

for all $j \geq j_0$ and $x \in \Lambda_0$.

Proof. Suppose that the proposition fails, then for some compact $\Lambda_0 \subset \Lambda$ and $\varepsilon > 0$, there exists a sequence $\{x_j\}_{j=1}^{\infty} \subset \Lambda_0$ and $p_j \in \partial u_j(x_j)$ for which $|p_j - \nabla u(x_j)| > \varepsilon$. By passing to subsequences, we may assume that $x_j \to x_0 \in \Lambda_0$, and for some fixed index $1 \leq i \leq n$ that $\langle p_j - \nabla u(x_j), e_i \rangle > \sqrt{\frac{\varepsilon}{n}}$ for all j (the case of $\langle p_j - \nabla u(x_j), e_i \rangle < -\sqrt{\frac{\varepsilon}{n}}$ is treated by a similar argument). Then, for any $\lambda > 0$, since $p_j \in \partial u_j(x_j)$ we find that

$$\frac{u_j(x_j + \lambda e_i) - u_j(x_j)}{\lambda} \geq \langle p_j, e_i \rangle > \sqrt{\frac{\varepsilon}{n}} + \langle \nabla u(x_j), e_i \rangle.$$

Recalling that u_j converges uniformly on compact subsets of Λ and ∇u is continuous on Λ ([31, Theorem 10.8 and Theorem 25.5]), by first taking the limit $j \to \infty$ (for all small enough $\lambda > 0$ so that $x_j + \lambda e_i \in \Lambda$) and then $\lambda \searrow 0$, we obtain the contradiction $\langle \nabla u(x_0), e_i \rangle \geq \sqrt{\varepsilon/n} + \langle \nabla u(x_0), e_i \rangle$, finishing the proof. ☐

Remark 3.6. We remark that if the limiting function u is not differentiable, then Lemma 3.5 above fails, even upon replacing $B_\varepsilon(\nabla u(x))$ by $\mathcal{N}_\varepsilon(\partial u(x))$, as seen by the following example. On $\Lambda = \mathbf{R}$ let $u_j := |x - 1/j|$ converging to $u := |x|$, and take the compact subdomain $\Lambda_0 := [-1, 1]$. Then if $\varepsilon = 1/2$, for any $j_0 \in \mathbf{N}$ we see that

$$\partial u_{j_0}\left(\frac{1}{j_0}\right) = [-1, 1] \not\subset [\frac{1}{2}, \frac{3}{2}] = \mathcal{N}_{1/2}\left(\partial u\left(\frac{1}{j_0}\right)\right),$$

hence there is no choice of j_0 for which the proposition holds uniformly over $[-1, 1]$.

Next we recall the *generalized (Clarke) Jacobian* of a mapping G (at a point x_0, in the last k variables).

Definition 3.7 (Clarke Jacobian)**.** If $G : B_\varepsilon(x_0) \subset \mathbf{R}^n \to \mathbf{R}^k$ is a Lipschitz function on a neighbourhood of x_0, we define $J^C G(x_0)$ to be the closed convex hull of all $k \times n$ matrices which can be written as limits of the form

$$\lim_{n\to\infty} DG(x_n)$$

where $x_n \to x_0$ and G is differentiable at each x_n.

Moreover if $1 \leq k \leq n$, using the notation $x = (x', x'') \in \mathbf{R}^{n-k} \times \mathbf{R}^k$ we write $J^C_{x''} G(x_0)$ for the set of $k \times k$ matrices consisting of the last k columns of elements in $J^C G(x_0)$.

A combination of Clarke's inverse function theorem [10, Theorem 1] and results of Vesely and Zajíček [34] on DC mappings yields the following DC implicit function theorem.

Theorem 3.8 (DC implicit mapping theorem [34, Proposition 5.9])**.** *Suppose $U \subset \mathbf{R}^{n-k} \times \mathbf{R}^k$ is open, $G : U \to \mathbf{R}^k$ is a DC mapping, and $G(x_0) = 0$ for some $x_0 = (x_0', x_0'') \in U$. Then if every element of $J^C_{x''} G(x_0)$ is invertible, there exists $\delta > 0$ and a bi-Lipschitz, DC mapping ϕ from $B_\delta(x_0') \subset \mathbf{R}^{n-k}$ into $\mathbf{R}^k$ such that for all $(x', x'') \in B_\delta(x_0') \times B_\delta(x_0'') \subset \mathbf{R}^{n-k} \times \mathbf{R}^k$:*

$$G(x', x'') = 0 \quad \text{if and only if} \quad x'' = \phi(x').$$

Additionally, a careful inspection of the proof of [4, Theorem 3.1] combined with [34, Theorem 5.1] yields the following DC constant rank theorem.

Theorem 3.9 (DC constant rank theorem)**.** *Suppose $U \subset \mathbf{R}^n$ is open, $G : U \to \mathbf{R}^k$ is a DC mapping, and $G(x_0) = 0$ for some $x_0 \in U$. Then if every element of $J^C G(x_0)$ has rank k, after a possible re-ordering and rotation of coordinates, the same conclusion as Theorem 3.8 above holds.*

We shall also need:

Lemma 3.10 (Coincident roots)**.** *Suppose $\phi_1^{\pm}, \ldots, \phi_k^{\pm}$ are real valued convex functions on $[-1, 1]^n$, such that $\phi_i^{\pm} > \phi_i^{\mp}$ on the set $\{x \in [-1, 1]^n \mid x^i = \pm 1\}$, and $\partial\phi_i^+ ([-1, 1]^n)$ is strongly separated from $\partial\phi_i^- ([-1, 1]^n)$ by a hyperplane normal to e_i for each $1 \leq i \leq k$. Then, there exists a point in $]-1, 1[^n$ where all $2k$ functions $\phi_1^{\pm} = \ldots = \phi_k^{\pm}$ agree.*

Proof. For any $x \in \mathbf{R}^n$, let us write $\hat{x}^i := (x^1, \ldots, x^{i-1}, x^{i+1}, \ldots, x^n)$. Fix $1 \leq i \leq k$, by Corollary 2.6, there is a DC function h_i defined on all of $\hat{I}_i := \{\hat{x}^i \mid x \in [-1, 1]^n\}$ such that the graph of h_i over this set is exactly

$$\{x \in [-1, 1]^n \mid \phi_i^+(x) = \phi_i^-(x)\};$$

by the intermediate value theorem we see for any $\hat{x} \in \hat{I}_i$ there exists $x \in [-1, 1]^n$ where $\phi_i^+(x) = \phi_i^-(x)$ and $\hat{x}^i = \hat{x}$, and in particular the range of h_i is contained in $[-1, 1]$. Now define the mapping $F : [-1, 1]^n \to [-1, 1]^n$ by

$$F(x) := (h_1(\hat{x}^1), \ldots, h_k(\hat{x}^k), x^{k+1}, \ldots, x^n),$$

this mapping is continuous by Theorem 2.3, thus by Brouwer's fixed point theorem it has a fixed point in $[-1, 1]^n$. However, we see that at this fixed point we must have $\phi_1^{\pm} = \ldots = \phi_k^{\pm}$, by the assumptions on the $\phi_i^{\pm}$ this point clearly must be in the interior $]-1, 1[^n$. □

With these preparations, we are ready to prove the main stability result.

Proof of Theorem 3.4. By [1, Theorem 1], the set of points x where $\dim \partial u(x) \geq k + 1$ has zero $\mathcal{H}^{n-k}$ measure, hence the final claim will follow immediately from (3.3).

Suppose we are given u, x_0, and a sequence $\{u^j\}_{j=1}^{\infty}$ as in the hypotheses of Theorem 3.4. Now by Lemma 2.5 we have

$$\partial u(x_0) = \operatorname{conv}\left(\bigcup_{1 \leq i \leq k+1} \{\nabla u_i(x_0)\}\right), \tag{3.4}$$

and since $\dim(\partial u(x_0)) = k$, the collection $\{\nabla u_i(x_0) - \nabla u_{k+1}(x_0)\}_{i=1}^k$ must be linearly independent, subtraction of a fixed linear function followed by a linear change of coordinates allows us to assume $\nabla u_i(x_0) = e_{n-k+i}$ for $1 \leq i \leq k$ and $\nabla u_{k+1}(x_0) = 0$. Next fix $\varepsilon > 0$, without loss of generality assume that $B_\varepsilon(x_0) \subset \mathcal{N}_{x_0}$. By our assumptions, we may add a fixed quadratic function centered at x_0 to assume all u_i^j and u_i are convex on $B_\varepsilon(x_0)$, for $1 \leq i \leq k + 1$ (possibly shrinking ε as well). By taking j large enough and possibly shrinking ε further, by the uniform convergence of each u_i^j we may assume

$$\min_{1 \leq i \leq k+1} u_i^j > \max_{k+2 \leq i \leq K} u_i^j \tag{3.5}$$

on $B_\varepsilon(x_0)$.

Define the mapping $F^j : B_\varepsilon(x_0) \to \mathbf{R}^k$ by

$$F^j(x) := (u_1^j(x) - u_{k+1}^j(x), \dots, u_k^j(x) - u_{k+1}^j(x))$$

then we see that if $x \in B_\varepsilon(x_0)$, the set $J^C_{x''}F^j(x)$ is contained in the collection of $k \times k$ matrices for which the ith row is contained in the convex hull of vectors of the form

$$\lim_{m\to\infty} D_{x''}(u_i^j - u_{k+1}^j)(x_m)$$

where $x_m \to x$ and u_i^j, u_{k+1}^j are differentiable at each x_m. Here $D_{x''}$ indicates the projection of the gradient of a function onto the last k variables. Since each function u_i is C^1, after shrinking ε if necessary and taking j large enough, by applying Lemma 3.5 we can assume that for any $x \in B_\varepsilon(x_0)$ and $p_i^j \in \partial u_i^j(x)$ we have

$$\begin{cases} p_i^j \in B_{\frac{1}{4}}(e_i), & 1 \le i \le k, \\ p_{k+1}^j \in B_{\frac{1}{4}}(0). \end{cases} \tag{3.6}$$

In particular, this implies that every matrix in $J^C_{x''}F^j(x)$ will be invertible, thus we can apply the DC implicit mapping theorem above to F^j, provided there exists at least one point $x_j \in B_\varepsilon(x_0)$ where F^j vanishes.

To this end, we translate so $x_0 = 0$, then we can apply the C^1 implicit function theorem to $u_i - u_{k+1}$ for each $1 \le i \le k$. For $\eta > 0$ small enough we then get $u_i - u_{k+1} > 0$ on $\{x \in [-\eta, \eta]^n \mid x^i = \eta\}$ while $u_i - u_{k+1} < 0$ on $\{x \in [-\eta, \eta]^n \mid x^i = -\eta\}$ for all $i \le k$. For any j large enough $u_i^j - u_{k+1}^j$ satisfies the same inequalities. Thus recalling (3.6), a dilation by $1/\eta$ allows us to apply Lemma 3.10 above to conclude the existence of a sequence $x_j \in\,]-\eta, \eta[^n \subset B_\varepsilon(x_0)$ such that $F^j(x_j) = 0$. In particular, we may now apply the DC implicit mapping theorem to find a ball $B^j \subset \pi_{n-k}(B_\varepsilon(x_0))$ and a DC mapping $\Phi^j : B^j \to B_\varepsilon(x_0)$ whose graph passes through x_j for which $u_1^j(\Phi^j(x')) = \ldots = u_{k+1}^j(\Phi^j(x'))$ for all $x' \in B^j$. Let

$$\Sigma^j_{n-k} := \{(x', \Phi^j(x')) \mid x' \in B^j\} \cap B_\varepsilon(x_0).$$

As a Lipschitz graph over $B^j \subset \mathbf{R}^{n-k}$ we see Σ^j_{n-k} has strictly positive $\mathcal{H}^{n-k}$ measure. Thus by Lemma 2.5, this implies (3.3), while (3.5) yields (3.2) to finish the proof. □

4. Applications to optimal transport

In this sequel, we apply the explicit function theorem and stability theorems from the previous two sections to the optimal transport problem. Throughout, Ω and $\overline{\Omega}$ are compact subsets of n-dimensional Riemannian manifolds (M, g) and $(\bar{M}, \bar{g})$ respectively, Ω^∂ is assumed to have dimension less than or equal to $n-1$, and $c \in C^4(\Omega \times \overline{\Omega})$. Also the notation $\mathcal{H}^i_g$ will refer to the i-dimensional Hausdorff measure of a set defined using the distance derived from the Riemannian metric g.

We begin by recalling a number of notions and conditions from the theory of the optimal transportation problem (OT) which adapt concepts from convex analysis such as the Legendre transform (2.2) to choices of cost other than $c(x, \bar{x}) = -\langle x, \bar{x}\rangle$.

Definition 4.1 (c-convex functions)**.** For a proper lower semicontinuous function $u : \Omega \to \mathbf{R} \cup \{\infty\}$, its *c-transform* $u^c : \overline{\Omega} \to \mathbf{R} \cup \{\infty\}$ is defined by

$$u^c(\bar{x}) := \sup_{x \in \Omega}(-c(x, \bar{x}) - u(x)).$$

Also its *double c-transform* $u^{cc^*} : \Omega \to \mathbf{R} \cup \{\infty\}$ is defined by

$$u^{cc^*}(x) := \sup_{\bar{x} \in \overline{\Omega}}(-c(x, \bar{x}) - u^c(\bar{x}));$$

u is said to be *c-convex* if $u = u^{cc^*}$ on Ω. Its *c-subdifferential* at a point $x_0 \in \Omega$ is the set

$$\partial_c u(x_0) := \left\{\bar{x} \in \overline{\Omega} \mid -c(x, \bar{x}) + c(x_0, \bar{x}) + u(x_0) \leq u(x), \quad \forall\, x \in \Omega\right\}.$$

Likewise, the *c^*-subdifferential* of u^c at $\bar{x}_0 \in \overline{\Omega}$ is defined as

$$\partial_{c^*} u^c(\bar{x}_0) := \left\{x \in \Omega \mid -c(x, \bar{x}) + c(x, \bar{x}_0) + u^c(\bar{x}_0) \leq u^c(\bar{x}), \quad \forall\, \bar{x} \in \overline{\Omega}\right\}.$$

Finally, for any subset $A \subset \Omega$, we write

$$\partial_c u(A) := \bigcup_{x \in A} \partial_c u(x),$$

and analogously for $\bar{A} \subset \overline{\Omega}$ and $\partial_{c^*} u^c(\bar{A})$.

Remark 4.2 (Strict c-convexity)**.** It can be shown u is a c-convex function if and only if for every $x_0 \in \Omega$, the set $\partial_c u(x_0) \neq \emptyset$. If in addition, for every $x_0 \in \Omega$ and $\bar{x}_0 \in \partial_c u(x_0)$, we have

$$\{x \in \Omega \mid -c(x, \bar{x}_0) + c(x_0, \bar{x}_0) + u(x_0) = u(x)\} = \{x_0\}$$

we say that u is *strictly c-convex.*

We also say c satisfies (B1) if for any $x_0 \in \Omega$ and $\bar{x}_0 \in \overline{\Omega}$, the mappings

$$\begin{aligned} \bar{x} &\mapsto -D_x c(x_0, \bar{x}) \in T^*_{x_0} M, \\ x &\mapsto -D_{\bar{x}} c(x, \bar{x}_0) \in T^*_{\bar{x}_0} \bar{M}, \end{aligned} \tag{B1}$$

are diffeomorphisms on Ω and $\overline{\Omega}$ respectively (these are classical conditions in optimal transport, corresponding for example to the twist and non-degeneracy conditions (A1) and (A2) in [23]). We will also write $exp^c_{x_0}(\cdot)$ for the inverse of the map in the first line above. Also for any sets $A \subset \Omega$ and $\bar{A} \subset \overline{\Omega}$, we will use the shorthand notation

$$\begin{aligned} [A]_{\bar{x}} &: = -D_{\bar{x}} c(A, \bar{x}), \\ \left[\bar{A}\right]_x &: = -D_x c(x, \bar{A}). \end{aligned} \tag{4.1}$$

At this point we recall a classical result about existence of solutions to (OT), originally due to Brenier for the case of the cost function $c(x, \bar{x}) = -\langle x, \bar{x}\rangle$.

Theorem 4.3 (Optimal transport maps [3, 16, 17, 24, 27, 29])**.** *If c satisfies* (B1) *and μ is absolutely continuous with respect to the volume measure on M, then there exists a c-convex function $u : \Omega \to \mathbf{R}$ which is differentiable almost everywhere, and the map $T(x) := exp^c_x(Du(x))$ is a solution to* (OT) *with $T(\mathrm{Dom}\, Du) \subset \mathrm{spt}\, \mu$. We call such a u an* optimal potential *transporting μ to ν, with cost c.*

In this first lemma, we show that if the support of the target measure consists of a (finite) union of disjoint, compact pieces, we can write the optimal potential as a maximum (of a finite number) of corresponding c-convex functions. For any function u, we will write $\mathrm{Dom}(Du)$ for the set of points where u is differentiable, which in the case of a semi-convex function (thus in particular, for any c-convex function under our assumptions) is a set of full Lebesgue measure in $\mathrm{Dom}(u)$.

Lemma 4.4 (Optimal maps to separated targets)**.** *Suppose a cost function c satisfies* (B1)*, μ is absolutely continuous, and ν is such that* $\operatorname{spt}\nu$ *is a disjoint union of an arbitrary (i.e. finite, coutable, or uncountable) collection $\{\overline{\Omega}_i\}_{i\in I}$ of compact subsets of the compact set $\overline{\Omega}$. The c-convex functions $u_i : \Omega \to \mathbf{R}$, $i \in I$ defined by*

$$u_i(x) := \sup_{\bar{x}\in\overline{\Omega}_i} (-c(x,\bar{x}) - u^c(\bar{x})) \tag{4.2}$$

satisfy

$$exp^c_x(Du_i(x)) \in \overline{\Omega}_i, \quad \forall\, x \in \mathrm{Dom}(Du),\ \forall\, i \in I, \tag{4.3}$$

$$u(x) = \sup_{i\in I} u_i(x), \quad \forall\, x \in \Omega. \tag{4.4}$$

Proof. First observe u_i is finite valued on all of Ω. Clearly u_i is c-convex, hence differentiable a.e.. Fix i and let x be such a point of differentiability, by compactness of $\overline{\Omega}_i$ there exists an $\bar{x} \in \overline{\Omega}_i$ achieving the supremum in the definition of $u_i(x)$. The inclusion (4.3) then follows immediately by differentiation of u_i at x and (B1).

Now as u is c-convex by we see that for $x \in \Omega$,

$$\begin{aligned} u(x) &= \sup_{\bar{x}\in\overline{\Omega}} [-c(x,\bar{x}) - u^c(\bar{x})] \\ &= \sup_{\bar{x}\in\operatorname{spt}\nu} [-c(x,\bar{x}) - u^c(\bar{x})] \\ &= \sup_{i\in I} u_i(x), \end{aligned} \tag{4.5}$$

proving (4.4). The reason why we may change the supremum above from being over $\overline{\Omega}$ to just over $\operatorname{spt}\nu$ is as follows. As mentioned previously, u is differentiable almost everywhere on Ω, so there exists a sequence $x_j \to x$ where u is differentiable at x_j and $\exists\ \bar{x}_j \in \partial_c u(x_j) = \left\{exp^c_{x_j}(Du(x_j))\right\}$ for each j. By [37, Theorem 10.28] (the assumption **(H∞)** of the reference is automatically satisfied by our assumption that $\overline{\Omega}$ is bounded) we must have $\bar{x}_j \in \operatorname{spt}\nu$, then by compactness, we may pass to a subsequence and assume $\bar{x}_j \to \bar{x}_0$ for some $\bar{x}_0 \in \operatorname{spt}\nu$, necessarily $\bar{x}_0 \in \partial_c u(x)$. However, this implies

$$\begin{aligned} \sup_{\bar{x}\in\overline{\Omega}} [-c(x,\bar{x}) - u^c(\bar{x})] &= \sup_{\bar{x}\in\overline{\Omega}} \inf_{y\in\Omega} [-c(x,\bar{x}) + c(y,\bar{x}) + u(y)] \\ &\le u(x) \le -c(x,\bar{x}_0) + c(y,\bar{x}_0) + u(y) \end{aligned}$$

for any $y \in \Omega$, thus we may take the supremum merely over $\operatorname{spt}\nu$. □

Remark 4.5. We pause to remark here that the above lemma will hold true even if Ω is not necessarily bounded. This is relevant for the cost $c(x,\bar{x}) = -\langle x, \bar{x}\rangle$ with $\Omega = \mathbf{R}^n$, for which it is established as in [18]; see also [17] for more general costs.

In order to discuss regularity, we will require some more geometric structure.

Definition 4.6 (c-convex sets)**.** $A \subset \Omega$ ($\bar{A} \subset \overline{\Omega}$) is *$c$-convex ($c^*$-convex) with respect to $\bar{x}_0$ (x_0)* if the set $[A]_{\bar{x}_0}$ ($\left[\bar{A}\right]_{x_0}$) from (4.1) is convex. We say *A is c-convex with respect to $\bar{A}$* if A is c-convex with respect to every $\bar{x} \in \bar{A}$, and *$\bar{A}$ is c^*-convex with respect to A* analogously. Finally, the phrase *A and $\bar{A}$ are c-convex with respect to each other* if both hold. We also refer to *strictly c-convex* and *strongly c-convex* by adding the corresponding modifiers to the convexity of $[A]_{\bar{x}_0}$ or $\left[\bar{A}\right]_{x_0}$, *strict* convexity meaning the midpoint of any nontrivial segment in $[A]_{\bar{x}_0}$ lies in the interior of $[A]_{\bar{x}_0}$, and *strong* convexity meaning $[A]_{\bar{x}_0}$ can be expressed as the intersection of a family of balls with fixed radii.

Lastly, $c \in C^4(\Omega \times \bar{\Omega})$ satisfies the (MTW) or (Ma-Trudinger-Wang) condition if

Definition 4.7 (MTW costs [27,33])**.** For some constant $a_3 \geq 0$, all $(x, \bar{x}) \in \Omega \times \overline{\Omega}$, $V \in T_xM$ and $\eta \in T^*_xM$ with $\eta(V) = 0$ satisfy

$$-(c_{ij,\bar{r}\bar{s}} - c_{ij,\bar{t}}c^{\bar{t},s}c_{s,\bar{r}\bar{s}})c^{\bar{r},k}c^{\bar{s},l}(x, \bar{x})V^iV^j\eta_k\eta_l \geq a_3|V|^2_g|\eta|^2_g \geq 0. \qquad \text{(MTW)}$$

Here, local coordinate systems are fixed near x and $\bar{x}$, and subscripts before a comma indicate differentiation with respect to the x variable, those after a comma are differentiation with respect to the $\bar{x}$ variable, and two raised indices indicate the matrix inverse. This last condition was crucial for regularity in the pioneering works [27, 33], it was later shown to have geometric implications by Loeper in [26] and by Kim and McCann in [23].

A particular geometric consequence of (MTW) that we will need is the following lemma, which follows from [26] [23].

Lemma 4.8 (Connected c-subdifferential images)**.** *Let c satisfy* (B1) *and* (MTW)*, and $\overline{\Omega}$ and Ω be c-convex with respect to each other. Then if $\mathcal{C} \subset \Omega$ is connected and u is a c-convex function on Ω, then $\partial_c u(\mathcal{C})$ is connected.*

Proof. Suppose not, then there exist disjoint, closed sets $\bar{C}_1$ and $\bar{C}_2 \subset \overline{\Omega}$ such that $\partial_c u(\mathcal{C}) \subset \bar{C}_1 \cup \bar{C}_2$, and $\partial_c u(\mathcal{C}) \cap \bar{C}_i \neq \emptyset$ for $i = 1, 2$. Define $C_i := \partial_{c^*} u^c(\bar{C}_i)$ for $i = 1, 2$. Since $\bar{x} \in \partial_c u(x)$ if and only if $x \in \partial_{c^*} u^c(\bar{x})$ we immediately have $\mathcal{C} \cap C_i \neq \emptyset$ for each i while $\mathcal{C} \subset C_1 \cup C_2$. On the other hand, suppose there exists $x \in C_1 \cap C_2 \cap \mathcal{C}$. Then there exist $\bar{x}_1 \in \bar{C}_1$ and $\bar{x}_2 \in \bar{C}_2$ such that both are contained in $\partial_c u(x)$. However [26, Theorem 3.1] implies the set $L_x := \{exp^c_x((1 - \lambda)(-D_x c(x, \bar{x}_1)) + \lambda(-D_x c(x, \bar{x}_2)) \mid \lambda \in [0, 1]\}$ is contained in $\partial_c u(x) \subset \bar{C}_1 \cup \bar{C}_2$. This is a contradiction as $\bar{C}_1$ and $\bar{C}_2$ would disconnect L_x, thus $C_1 \cap C_2 \cap \mathcal{C} = \emptyset$. Last, since each $\bar{C}_i$ is compact, we immediately see that C_i is also compact, hence closed. Thus we have a contradiction with the connectedness of $\mathcal{C}$. □

We are particularly interested in optimal transport problems between measures μ and ν satisfying the following properties, which are related to regularity results proved for $c(x, \bar{x}) = -\langle x, \bar{x}\rangle$ by Caffarelli [5] and extended to other costs in [15] [21]; see also Chen and Wang [8], and Vétois [35]:

(I) Both μ and ν are absolutely continuous with respect to the respective volume measures on M and $\bar{M}$, and with densities bounded a.e. away from 0 and ∞ on their supports.

(II) Ω and $\overline{\Omega}$ are c-convex with respect to each other and either

$$\operatorname{spt}\mu \text{ and } \overline{\Omega} \text{ are strongly } c\text{-convex with respect to each other} \tag{4.6}$$

or

$$\operatorname{spt}\mu \subset \Omega^{\text{int}},\ \operatorname{spt}\nu \subset \overline{\Omega}^{\text{int}},$$
$$\operatorname{spt}\mu \text{ is } c\text{-convex with respect to } \overline{\Omega}. \tag{4.7}$$

Under the above conditions and (MTW), we can make the following improvement of Lemma 4.4. The idea is based on one used by Caffarelli and McCann [6, Theorem 6.3] for the cost function $c(x, \bar{x}) = -\langle x, \bar{x}\rangle$.

Proposition 4.9 (Continuous optimal maps onto closed c-convex target pieces)**.** *In addition to the hypotheses of Lemma 4.4 and* (MTW)*, assume that μ and ν on M and $\bar{M}$ respectively satisfy conditions (I) and (II) above, and for some $i \in I$ the compact set $\overline{\Omega}_i$ is strictly c-convex with respect to the compact set Ω. Then the c-convex function u_i from Lemma 4.4 belongs to $C^1(\Omega)$,*

$$\partial_c u_i(\Omega) \subset \overline{\Omega}_i, \tag{4.8}$$

and for any $x \in \operatorname{spt}\mu$ the intersection $\partial_c u(x) \cap \overline{\Omega}_i$ contains at most one point.

Proof. Since $\overline{\Omega}_i$ is c-convex with respect to Ω, combining [27, Lemma 5.1] with (4.3) yields $\partial_c u_i(\Omega) \subset \overline{\Omega}_i$ to establish (4.8).

Next we show that each u_i is C^1 on Ω. Indeed, note that u^c is an optimal potential transporting ν to μ with cost function c^* defined on $\overline{\Omega} \times \Omega$ by $c^*(\bar{x}, x) := c(x, \bar{x})$, then by [15, Theorem 2.1] under (4.6), or [21, Lemma 2.19, Theorem 1.2] under (4.7), we have that $u^c \in C^{1,\bar{\alpha}}_{loc}$ for $\bar{\alpha}$ as described and strictly c^*-convex when restricted to each $\overline{\Omega}_i^{\text{int}}$. If there was a point x where u_i fails to be differentiable, by [26, Theorem 3.1] this implies the existence of some nontrivial line segment $\ell \subset \partial u_i(x) = [\partial_c u_i(x)]_x \subset [\overline{\Omega}_i]_x$. However, by the strict convexity of $[\overline{\Omega}_i]_x$, this would imply that $\ell \cap [\overline{\Omega}_i]_x^{\text{int}}$ contains more than one point. It can be seen that this contradicts the strict c^*-convexity of u^c on $\overline{\Omega}_i^{\text{int}}$, thus u_i must be differentiable on Ω. The fact that the c-subdifferential of a c-convex function has a closed graph then implies $u_i \in C^1(\Omega)$.

Now if $x \in \operatorname{spt}\mu$ and $\partial_c u(x) \cap \overline{\Omega}_i$ contains more than one point, the same argument as the previous paragraph combined with the representation (4.5) again yields a contradiction. □

As a corollary to its proof we obtain the following interior homeomorphism result, which can be upgraded to a diffeomorphism using results from the literature.

Corollary 4.10 (Optimal homeomorphisms onto open c-convex target pieces)**.** *Assume the same hypotheses as Proposition 4.9, but if condition* (4.6) *is assumed from (II), additionally suppose that $\overline{\Omega}_i$ is strongly c-convex with respect to Ω. Then the map the map $T_i(x) := exp^c_x(Du_i(x))$ is a homeomorphism from the interior of $\{x \in \operatorname{spt}\mu \mid u(x) = u_i(x)\}$ to $\overline{\Omega}_i^{\text{int}}$; its inverse is $C^{\bar{\alpha}}_{loc}$ for some $\bar{\alpha} > 0$ depending only on n and the bounds (I). If the densities of μ and ν are locally Dini continuous (respectively $C^{k+\alpha}_{loc}$ for any $0 < k + \alpha \notin \mathbf{N}$) on the interiors of these two sets, then T_i defines a diffeomorphism whose derivatives are locally Dini continuous (or $C^{k+\alpha}_{loc}$ respectively), at least if $a_3 > 0$ or $c(x, \bar{x}) = -\langle x, \bar{x}\rangle$; see [25] re $a_3 = 0$.*

Proof. The strict c^*-convexity of $u^c \in C^{1,\bar\alpha}_{loc}$ from the preceding proof shows the map $S(\bar x) := exp^{c^*}_{\bar x}(Du^c(\bar x))$ restricted to $\overline{\Omega}_i^{\text{int}}$ is a homeomorphism (and $C^{\bar\alpha}_{loc}$). We assert this restriction has range R^{int} where $R := \{x \in \operatorname{spt}\mu \mid u(x) = u_i(x)\}$, and its inverse is T_i.

First note that $u^c(\bar x) = (u_i)^c(\bar x)$ for $\bar x \in \overline{\Omega}_i$. Indeed, $u_i \le u$ implies $(u_i)^c \ge u^c$ everywhere, while for $\bar x \in \overline{\Omega}_i$ the opposite inequality is obtained by taking $\bar y = \bar x$ in

$$(u_i)^c(\bar x) = \sup_{x\in\Omega}[-c(x,\bar x) + \inf_{\bar y\in\bar\Omega_i}(c(x,\bar y) + u^c(\bar y))].$$

Then, recall

$$u(x) + u^c(\bar x) + c(x,\bar x) \ge 0 \qquad \text{for all } (x,\bar x) \in \Omega\times\bar\Omega, \tag{4.9}$$

and equality holds if and only if $\bar x \in \partial_c u(x)$ (or equivalently $x \in \partial_{c^*}u^c(\bar x)$). For $\bar x \in \overline{\Omega}_i^{\text{int}}$, we have $\partial_{c^*}u^c(\bar x) = \{S(\bar x)\}$ thus

$$\begin{aligned} u(S(\bar x)) &= -c(S(\bar x),\bar x) - u^c(\bar x) = -c(S(\bar x),\bar x) - (u_i)^c(\bar x) \\ &= -c(S(\bar x),\bar x) + \inf_{y\in\Omega}(c(y,\bar x) + u_i(y)) \le u_i(S(\bar x)). \end{aligned}$$

Since the reverse inequality always holds, we have $u(S(\bar x)) = u_i(S(\bar x))$. Then as S is injective and continuous, the set $S(\overline{\Omega}_i^{\text{int}})$ is open, hence it must be contained in R^{int}.

We now claim that T_i pushes the restriction of μ to R^{int} forward to the restriction of ν to $\overline{\Omega}_i$. Let us write $T(x) := exp^c_x(Du(x))$, defined for $x \in \operatorname{Dom}(Du)$ so $T_\#\mu = \nu$. By Lemma 2.5 and (4.3), we see that $x \in \operatorname{Dom}(Du)$ with $T(x) \in \overline{\Omega}_i$ only if $u(x) = u_i(x)$ and $u(x) > u_j(x)$ for all $j \ne i$, in particular, $T^{-1}(\overline{\Omega}_i) \subset R^{\text{int}}$. On the other hand, if $x \in R^{\text{int}}$, then $u = u_i$ on a neighborhood of x and in particular, u is differentiable at x. Hence by (B1) we must have $\partial_c u(x) = \{T_i(x)\} = \{T(x)\}$ for all $x \in R^{\text{int}}$. Thus if $\bar E \subset \overline{\Omega}_i$ is measurable, we have

$$\mu(R^{\text{int}} \cap T_i^{-1}(\bar E)) = \mu(R^{\text{int}} \cap T^{-1}(\bar E)) = \mu(T^{-1}(\bar E)) = \nu(\bar E)$$

and the claim is proven.

Thus again using [15, Theorem 2.1] under (4.6) (along with the assumption of strong c-convexity of $\overline{\Omega}_i$ with respect to Ω in this case), and [21, Lemma 2.19, Theorem 1.2] under (4.7) gives that T_i is continuous and injective on R^{int}, hence $T_i(R^{\text{int}}) \subset \overline{\Omega}_i^{\text{int}}$.

We complete the proof of the claim by showing $S\circ T_i = id_{R^{\text{int}}}$. Since for each $x \in R^{\text{int}}$, we have $\partial_c u(x) = \{T_i(x)\} \subset \overline{\Omega}_i^{\text{int}}$, as argued above this yields $\partial_{c^*}u^c(T_i(x))) = \{S(T_i(x)))\}$. The equality conditions in (4.9) then force $x = S(T_i(x))$ as required.

When (MTW) holds with $a_3 > 0$, the local Dini or Hölder continuity asserted then follows from [25], where it is also claimed that the results extend to $a_3 = 0$, although details of this extension are deferred to a forthcoming publication. For $c(x,\bar x) = -\langle x,\bar x\rangle$, the details can be found in [5], [38], and [22]. □

Next we wish to make some finer observations on the structure of the boundaries of the sets above, and in particular the sets where more than two of the functions u_i coincide. For this we need some notion of "independence" for subcollections of $\{\overline{\Omega}_i\}_{i\in I}$, which we call *affine independence*. Its role is to guarantee the natural implicit function theorem hypothesis is satisfied in the applications which follow.

Definition 4.11 (Affine independence)**.** A finite collection $\{\overline{\Lambda}_i\}_{i=1}^k$ of $k \leq n+1$ subsets of an n dimensional vector space is said to be *affinely independent* if no $k-2$ dimensional affine subspace intersects all of the sets in the collection. (Equivalently, any collection of k points, each from a different set $\overline{\Lambda}_i$, is affinely independent in the usual sense.)

We also define an alternate notion measuring the "size" of a singular point that we call the *multiplicity*. Essentially the multiplicity of a singular point counts "how many pieces of the target domain does a singular point get transported to?"

Definition 4.12 (Multiplicity along tears)**.** Let c be a cost function satisfying (B1) and μ, ν probability measures with μ absolutely continuous with respect to volume measure. Also suppose $\operatorname{spt}\nu \subset \overline{\Omega}$ is a disjoint union of some collection of sets $\{\overline{\Omega}_i\}_{i\in I}$ for some index set I and u is an optimal potential of (OT) transporting μ to ν, with $x_0 \in \operatorname{spt}\mu$. Then we define the *multiplicity of u at x_0 relative to $\{\overline{\Omega}_i\}_{i\in I}$* by

$$\#\left\{i \in I \mid \overline{\Omega}_i \cap \partial_c u(x_0) \neq \emptyset\right\}.$$

When the collection $\{\overline{\Omega}_i\}_{i\in I}$ is clear, we will simply refer to the multiplicity of u at x_0.

Finally, in order to simplify the statements and proofs of our results, we define notation for coincidence sets and multiplicity sets of the functions u_i and u.

Definition 4.13 (Tearing and coincidence sets)**.** Let c be a cost function satisfying (B1). Also take compactly supported probability measures μ and ν with μ absolutely continuous, and $\operatorname{spt}\nu = \cup_{i\in I}\overline{\Omega}_i$ a *finite disjoint* union of *compact* sets $\overline{\Omega}_i$. Then Lemma 4.4 asserts

$$u = \sup_{i\in I} u_i \quad \text{with} \quad exp_x^c(Du_i(x)) \in \overline{\Omega}_i,\ \forall\, x \in \operatorname{Dom}(Du).$$

For any subset $I' \subset I$ of indices, we then define the set

$$\Sigma_{I'} := \{x \in \Omega \mid u_i(x) = u_j(x),\ \forall\, i,\ j \in I'\}, \tag{4.10}$$

$$\Sigma_{I'}^{\uparrow} := \{x \in \Omega \mid u(x) = u_i(x),\ \forall\, i \in I'\}. \tag{4.11}$$

Also for any $k \in \mathbf{Z}_{\geq 0}$ we define

$$M_k := \{x \in \mathbf{R}^n \mid u \text{ has multiplicity exactly } k \text{ at } x\}, \tag{4.12}$$

$$M_{\geq k} := \{x \in \mathbf{R}^n \mid u \text{ has multiplicity at least } k \text{ at } x\}, \tag{4.13}$$

where u is the optimal potential as in (5.5) and multiplicity here taken relative to the collection $\{\overline{\Omega}_i\}_{i\in I}$ in Definition 4.12.

Under a suitable assumption of affine independence, a quick application of the usual implicit function theorem yields the following corollary from Proposition 4.9.

Corollary 4.14 (Affine independence of convex targets yields C^1 smooth tears of each expected codimension)**.** *Assume c is a cost function satisfying* (B1) *and* (MTW)*, μ and ν are probability measures on M and $\bar{M}$ respectively, and conditions (I) and (II) (before* (4.6)*) hold. Let $\operatorname{spt}\nu = \cup_{i\in I}\overline{\Omega}_i$ be a finite disjoint union of compact sets, and $u = \max u_i$ be from Lemma 4.4. Finally suppose there is a collection of indices $i_1, \ldots, i_k \in I$ for which $\{[\overline{\Omega}_{i_1}]_x, \ldots, [\overline{\Omega}_{i_k}]_x\}$ forms an affinely independent collection of strictly convex sets for every $x \in \Sigma_{i_1,\ldots,i_k}$. Then $\Sigma_{i_1,\ldots,i_k}$ is a C^1 submanifold of M having codimension $k-1$.*

Proof. Reordering if necessary, we may assume $i_j = j$ for each $j \leq k$. The set $\Sigma_{1,\dots,k}$ then consists of the zero set of the system of $k-1$ equations

$$u_1(x) = u_2(x) = \cdots = u_k(x), \tag{4.14}$$

which are all contained in $C^1(\Omega)$ by Proposition 4.9. The implicit function theorem condition for the zero set of this system to be a C^1 submanifold of the appropriate dimension is that the vectors $\{Du_j(x) - Du_k(x)\}_{j=1}^{k-1}$ be linearly independent when (4.14) holds, which is equivalent to affine independence of $\{Du_j(x)\}_{j=1}^{k}$. But since $Du_i(x) \in \left[\overline{\Omega}_i\right]_x$ by (4.3) and (B1), this follows from the affine independence of $\{\left[\overline{\Omega}_i\right]_x\}_{i=1}^{k}$. □

Next, we establish two elementary relationships between the sets $\Sigma^{\uparrow}$ and M. Specifically, we show that the closure M_k^{cl} of all points with multiplicity lie in a union of tears; we later prove that when the disjoint components of $\operatorname{spt}\nu = \cup_{i\in I}\overline{\Omega}_i$ can be separated by hyperplanes pairwise (5.6), these tears lie in DC submanifolds.

Lemma 4.15 (Covering multiplicity sets with tears)**.** *Suppose that c is a cost function satisfying* (B1)*, μ and ν are probability measures with μ absolutely continuous with respect to the volume measure, and* $\operatorname{spt}\nu = \bigcup_{i\in I}\overline{\Omega}_i$ *is a disjoint union of compact sets. Then multiplicity is upper semicontinuous:*

$$M_k^{\mathrm{cl}} \subset M_{\geq k}. \tag{4.15}$$

Additionally, fix a positive integer k and suppose that for any collection of indices $I' \subset I$ with $\#(I') = k$ and $x \in \Omega$

$$\left\{\left[\overline{\Omega}_i\right]_x\right\}_{i\in I'} \tag{4.16}$$

is affinely independent. Then

$$M_{\geq k} \subset \bigcup_{\{I'\subset I \mid \#(I')=k\}} \Sigma^{\uparrow}_{I'}. \tag{4.17}$$

Proof. Suppose $x_0 \in M_k^{\mathrm{cl}}$, so there is a sequence $\{x_m\}_{m=1}^{\infty} \subset M_k$ converging to x_0. We may pass to a subsequence and assume, without loss of generality, that each $\partial_c u(x_m)$ only intersects $\overline{\Omega}_1, \dots, \overline{\Omega}_k$ out of the collection $\{\overline{\Omega}_i\}_{i\in I}$, and take $\bar{x}_{i,m} \in \partial_c u(x_m) \cap \overline{\Omega}_i$ for $i \in \{1, \dots, k\}$. Since each $\overline{\Omega}_i$ is compact, we may pass to further subsequences to assume each $\bar{x}_{i,m}$ converges as $m \to \infty$ to some $\bar{x}_i \in \overline{\Omega}_i$, and by upper semicontinuity of the c-subdifferential we see that $\bar{x}_i \in \partial_c u(x_0)$, meaning $x_0 \in M_{\geq k}$.

Now assume (4.16) holds and take $x_0 \in \Omega \setminus \bigcup_{\{I'\subset I \mid \#(I')=k\}} \Sigma^{\uparrow}_{I'}$. If $\#(I) < k$, then clearly $x_0 \notin M_{\geq k}$, thus assume $\#(I) \geq k$. From (5.5) it is clear that $u(x_0) = u_i(x_0)$ for at least one index i, and this can only hold for at most $k' \leq k-1$ distinct indices; suppose we have $u(x_0) = u_{i_j}(x_0)$ for $1 \leq j \leq k'$ and strict inequality for all other indices. Then by Lemma 2.5 and (4.3)

$$[\partial_c u(x_0)]_{x_0} \subset \partial u\,(x_0) \subset \operatorname{conv}\left(\bigcup_{1\leq j\leq k'} \operatorname{conv}\left(\left[\overline{\Omega}_{i_j}\right]_{x_0}\right)\right) = \operatorname{conv}\left(\bigcup_{1\leq j\leq k'} \left[\overline{\Omega}_{i_j}\right]_{x_0}\right).$$

Thus if the multiplicity of u at x_0 is k or greater, there exists an index $i' \notin \{i_1, \dots, i_{k'}\}$ for which $\partial_c u(x_0) \cap \overline{\Omega}_{i'} \neq \emptyset$, by the above inclusion this implies there is a point in $\left[\overline{\Omega}_{i'}\right]_{x_0}$ which can be written as the convex combination of k' points, one from each of the sets $\{\left[\overline{\Omega}_{i_1}\right]_{x_0},\ldots,\left[\overline{\Omega}_{i'_k}\right]_{x_0}\}$. Since $k'\leq k-1$ and $\#(I)\geq k$, we can complete $\{i_1,\ldots,i'_k,\ i'\}$ to a subset of I with cardinality k to obtain a contradiction with (4.16), hence $x_0\notin M_{\geq k}$.

□

5. Global structure of optimal map discontinuities: quadratic cost

We state the results of this section in the model case $c(x,\bar{x}) = -\langle x,\bar{x}\rangle$ on $\mathbf{R}^n\times\mathbf{R}^n$, where the proofs are much simpler and the geometric picture easier to understand. It is easily verified that this cost function satisfies (B1) and (MTW), both $exp^c_x(\cdot)$ and $exp^{c^*}_{\bar{x}}(\cdot)$ are the identity mapping for any x and $\bar{x}$, and c- and c^*-convexity of sets reduces to the usual convexity of a set.

Our first result is the following proposition which — apart from its final sentence — follows rapidly from our explicit function theorem. It will be extended to MTW costs in a subsequent section.

Proposition 5.1 (Hyperplane separated components induce DC tears)**.** *Let $c(x,\bar{x}) = -\langle x,\bar{x}\rangle$. Also suppose μ and ν are absolutely continuous probability measures with bounded supports, and $\operatorname{spt}\nu=\overline{\Omega}_1\cup\overline{\Omega}_2$ is such that $\overline{\Omega}_1$ and $\overline{\Omega}_2$ are strongly separated by some hyperplane Π.*

Then an optimal potential u transporting μ to ν can be written $u=\max\{u_1,u_2\}$, where u_1 and u_2 are convex functions, finite on $\mathbf{R}^n$ such that

$$\nabla u_i(x)\in\overline{\Omega}_i,\quad \forall\, x\in\operatorname{Dom}(\nabla u). \tag{5.1}$$

Moreover, the sets

$$\begin{aligned}
\Sigma :&= \left\{x\in\mathbf{R}^n \mid \partial u\,(x)\cap\overline{\Omega}_i\neq\emptyset,\ i=1,2\right\} = \{x\in\mathbf{R}^n\mid u_1(x)=u_2(x)\},\\
C_1 :&= \left\{x\in\mathbf{R}^n \mid \partial u\,(x)\cap\overline{\Omega}_2=\emptyset\right\} = \{x\in\mathbf{R}^n\mid u_1(x)>u_2(x)\},\\
C_2 :&= \left\{x\in\mathbf{R}^n \mid \partial u\,(x)\cap\overline{\Omega}_1=\emptyset\right\} = \{x\in\mathbf{R}^n\mid u_1(x)<u_2(x)\}.
\end{aligned}$$

are connected and given by the graph, open epigraph, and open subgraph respectively of a globally Lipschitz DC function h defined on the hyperplane Π.

If $\operatorname{spt}\mu$ is convex and $\overline{\Omega}_i$ is connected for either $i=1$ or 2, then $\operatorname{spt}\mu\cap(C_i\cup\Sigma)$ is also connected.

Proof. Let us assume $\Pi=\{x\in\mathbf{R}^n\mid x^n=0\}=\mathbf{R}^{n-1}$. By Lemma 4.4 (note we do not necessarily need boundedness of Ω, see Remark 4.5) we find that $u=\max\{u_1,u_2\}$, both u_i are convex and finite on $\mathbf{R}^n$, and we have (5.1). Since $\overline{\Omega}_1$ and $\overline{\Omega}_2$ are strongly separated by $\mathbf{R}^{n-1}$, so are their convex hulls, and (5.1) implies $\partial u_i\left(\mathbf{R}^n\right)\subset\operatorname{conv}(\overline{\Omega}_i)$. Thus we can apply Corollary 2.3 to obtain the function h defined on $\mathbf{R}^{n-1}$ along with all claimed properties above; the connectedness from continuity of h.

Now assume $\operatorname{spt}\mu$ is convex and $\overline{\Omega}_1$ is connected. Let $d(x):=d(x,\operatorname{spt}\mu)^2$ which is finite and convex on $\mathbf{R}^n$, and define $\tilde{u}:=u+d$. An easy calculation gives

$$\partial d\,(x)=\begin{cases}\{0\}, & x\in\operatorname{spt}\mu,\\ d(x,\operatorname{spt}\mu)\frac{x-\pi_{\operatorname{spt}\mu}(x)}{|x-\pi_{\operatorname{spt}\mu}(x)|}, & x\notin\operatorname{spt}\mu,\end{cases}$$

where $\pi_{\operatorname{spt}\mu}(x)$ is the (unique) closest point projection of x onto $\operatorname{spt}\mu$. Thus we see by [31, Theorem 23.8] that

$$\partial\tilde{u}\,(x)=\partial u\,(x)\,,\quad \forall\, x\in\operatorname{spt}\mu. \tag{5.2}$$

Next we will show that $\partial \tilde{u}^* (\bar{x}) \subset \operatorname{spt} \mu$ for every $\bar{x} \in \overline{\Omega}_1$ (this is a nontrivial claim for $\bar{x} \in \overline{\Omega}_1^{\partial}$). By [31, Theorem 16.4] we have

$$\tilde{u}^*(\bar{x}) = \inf_{\bar{y} \in \mathbf{R}^n} (u^*(\bar{x} - \bar{y}) + d^*(\bar{y})), \tag{5.3}$$

we will calculate $d^*(\bar{y})$. Let us write $h(\bar{y}) := \sup_{x \in \operatorname{spt} \mu} \langle x, \bar{y} \rangle$ for the *support function* of $\operatorname{spt} \mu$, since $\operatorname{spt} \mu$ is compact, for each $\bar{y} \in \mathbf{R}^n$ there exists $z(\bar{y}) \in \operatorname{spt} \mu$ such that $h(\bar{y}) = \langle z(\bar{y}), \bar{y} \rangle$. Clearly $d^*(0) = 0$, so assume $\bar{y} \neq 0$. Then by definition,

$$d^*(\bar{y}) = \sup_{x \in \mathbf{R}^n} (\langle x, \bar{y} \rangle - d(x, \operatorname{spt} \mu)^2) = \sup_{\{x \in \mathbf{R}^n \mid \langle x, \bar{y} \rangle > \langle z(\bar{y}), \bar{y} \rangle\}} (\langle x, \bar{y} \rangle - d(x, \operatorname{spt} \mu)^2).$$

Fix any x such that $\langle x, \bar{y} \rangle > \langle z(\bar{y}), \bar{y} \rangle$, and an arbitrary $y \in \operatorname{spt} \mu$, then for some $\lambda \in [0, 1)$ we have $x_\lambda := (1-\lambda) y + \lambda x$ satisfies $\langle x_\lambda, \bar{y} \rangle = \langle z(\bar{y}), \bar{y} \rangle$. Then we calculate

$$|x - y| \geq |x - x_\lambda| \geq \langle x - x_\lambda, \frac{\bar{y}}{|\bar{y}|} \rangle = \langle x - z(\bar{y}), \frac{\bar{y}}{|\bar{y}|} \rangle,$$

hence taking an infimum over $y \in \operatorname{spt} \mu$,

$$\langle x, \bar{y} \rangle - d(x, \operatorname{spt} \mu)^2 \geq h(\bar{y}) + \langle x - z(\bar{y}), \bar{y} \rangle - \frac{\langle x - z(\bar{y}), \bar{y} \rangle^2}{|\bar{y}|^2}.$$

This last quantity can be seen to be maximized over $\langle x, \bar{y} \rangle > \langle z(\bar{y}), \bar{y} \rangle$ when $\langle x - z(\bar{y}), \bar{y} \rangle = \frac{|\bar{y}|^2}{2}$, yielding

$$d^*(\bar{y}) = h(\bar{y}) + \frac{|\bar{y}|^2}{2} - \frac{|\bar{y}|^2}{4} = h(\bar{y}) + \frac{|\bar{y}|^2}{4}.$$

By choosing $\bar{y} = 0$ in (5.3), for any $\bar{x} \in \mathbf{R}^n$ we clearly have

$$\tilde{u}^*(\bar{x}) \leq u^*(\bar{x}).$$

On the other hand, suppose $\bar{x}_0 \in \overline{\Omega}_1^{\mathrm{int}}$. By [36, Theorem 2.12] u^* is an optimal potential transporting ν to μ, then by [37, Theorem 10.28] and convexity of $\operatorname{spt} \mu$, we have that $\partial u^* (\bar{x}_0) \in \operatorname{spt} \mu$, let $x_0 \in \partial u^* (\bar{x}_0)$. Then for any $\bar{y} \in \mathbf{R}^n$,

$$u^*(\bar{x}_0 - \bar{y}) + h(\bar{y}) + \frac{|\bar{y}|^2}{4} \geq u^*(\bar{x}_0) + \langle \bar{x}_0 - \bar{y} - \bar{x}_0, x_0 \rangle + \langle \bar{y}, x_0 \rangle = u^*(x_0),$$

thus taking an infimum over $\bar{y} \in \mathbf{R}^n$ and recalling (5.3) gives $\tilde{u}^* \geq u^*$ on $\overline{\Omega}_1^{\mathrm{int}}$. Since the Legendre transform of a convex function is always closed, we then have $\tilde{u}^* \equiv u^*$ on all of $\overline{\Omega}_1 = \overline{\Omega}_1^{\mathrm{cl}}$. Now let $\bar{x}_0 \in \overline{\Omega}_1$ and suppose $x_0 \in \partial \tilde{u} (\bar{x}_0)$. Then for any $\bar{x}$, $\bar{y} \in \mathbf{R}^n$, again using (5.3),

$$\begin{aligned} u^*(\bar{x} - \bar{y}) + h(\bar{y}) + \frac{|\bar{y}|^2}{4} &\geq \tilde{u}^*(\bar{x}) \geq \tilde{u}^*(\bar{x}_0) + \langle \bar{x} - \bar{x}_0, x_0 \rangle \\ &= u^*(\bar{x}_0) + \langle \bar{x} - \bar{x}_0, x_0 \rangle. \end{aligned}$$

We can let $\bar{y}$ vary over $\mathbf{R}^n \setminus \{0\}$ while setting $\bar{x} = \bar{y} + \bar{x}_0$ in the equation above, then dividing through by $|\bar{y}|$ we find

$$\sup_{x \in \operatorname{spt} \mu} \langle x, \frac{\bar{y}}{|\bar{y}|} \rangle + \frac{|\bar{y}|}{4} \geq \langle x_0, \frac{\bar{y}}{|\bar{y}|} \rangle,$$

taking $\bar{y} \to 0$ radially gives

$$\sup_{x \in \operatorname{spt} \mu} \langle x, \omega \rangle \geq \langle x_0, \omega \rangle, \quad \forall \omega \in \mathbf{S}^{n-1},$$

hence we must have $x_0 \in \operatorname{spt} \mu$ as claimed.

We now claim that

$$\partial\tilde{u}^*\left(\overline{\Omega}_1\right) = \operatorname{spt}\mu \cap (C_1 \cup \Sigma), \tag{5.4}$$

then the proof will be complete by applying Lemma 4.8. Suppose $x_0 \in \operatorname{spt}\mu \cap (C_1 \cup \Sigma)$. Recall by (5.2), $\partial u\,(x_0) = \partial\tilde{u}\,(x_0)$. There are two possibilities, either $u_1(x_0) > u_2(x_0)$, or $u_1(x_0) = u_2(x_0)$. In the first case, $\partial u\,(x_0) = \partial u_1\,(x_0)$, while in the second case, by Lemma 2.5 we have $\partial u\,(x_0) = \operatorname{conv}(\partial u_1\,(x_0) \cup \partial u_2\,(x_0))$. In either case, since $\partial u_1\,(x_0) \cap \overline{\Omega}_1 \neq \emptyset$ by (5.1), there exists $y_0 \in \overline{\Omega}_1$ such that $y_0 \in \partial\tilde{u}\,(x_0)$. Hence $x_0 \in \partial\tilde{u}^*\,(y_0) \subset \partial\tilde{u}^*\left(\overline{\Omega}_1\right)$.

Now suppose $x_0 \in \partial\tilde{u}^*\left(\overline{\Omega}_1\right)$ but $u_2(x_0) > u_1(x_0)$. As we have shown above, $x_0 \in \operatorname{spt}\mu$. Then by (5.2) combined with Lemma 2.5, $\partial\tilde{u}\,(x_0) = \partial u\,(x_0) = \partial u_2\,(x_0) \subset \operatorname{conv}(\overline{\Omega}_2)$. However this is a contradiction, as this gives $\partial\tilde{u}\,(x_0) \cap \overline{\Omega}_1 = \emptyset$. This concludes the proof of (5.4). □

We can also obtain some structure in the case where $\operatorname{spt}\nu$ consists of more than two regions separated by hyperplanes. Before we state the results, some setup.

Again, we restrict the discussion to the bilinear cost $c(x,\bar{x}) = -\langle x,\bar{x}\rangle$ on $\mathbf{R}^n \times \mathbf{R}^n$, while μ and ν are absolutely continuous probability measures with bounded supports. We'll assume $\operatorname{spt}\nu = \cup_{i\in I}\overline{\Omega}_i$ is a decomposition into finitely many compact disjoint sets; i.e. henceforth we assume that I is *finite*. Then if u is an optimal potential transporting μ to ν, by Lemma 4.4 there exist convex functions u_i, $i \in I$ on $\mathbf{R}^n$ such that

$$u = \sup_{i\in I} u_i \quad \text{with} \quad \nabla u_i(x) \in \overline{\Omega}_i,\ \forall\, x \in \operatorname{Dom}(\nabla u). \tag{5.5}$$

If some $\overline{\Omega}_i$ is strictly convex, $\operatorname{spt}\mu$ is convex, and the densities of μ and ν are bounded away from zero and infinity on their supports, by Proposition 4.9 we have $u_i \in C^1(\mathbf{R}^n)$. We'll often require that each $\overline{\Omega}_i$ can be strongly separated from each $\overline{\Omega}_j$ by a hyperplane, so that their convex hulls are disjoint: on is that the sets $\operatorname{conv}(\overline{\Omega}_i)$ are mutually disjoint, hence

$$\partial u_i\,(\mathbf{R}^n) \subset \operatorname{conv}(\overline{\Omega}_i) \text{ are mutually disjoint.} \tag{5.6}$$

We begin with two corollaries of Theorem 2.3 (the sets $\Sigma_{I'}$ and $\Sigma^{\uparrow}_{I'}$ below for a collection of indices I' are defined by (4.10) and (4.11) respectively):

Corollary 5.2 (DC rectifiability of Σ_{ij})**.** *If $\overline{\Omega}_i$ and $\overline{\Omega}_j$ can be strongly separated by a hyperplane Π for some $i \neq j$ in Definition 4.13, then $\Sigma_{ij} := \Sigma_{\{i,j\}}$ is a globally Lipschitz DC graph over Π.*

Proof. The convex hull of $\overline{\Omega}_i$ contains $\partial u_i(\mathbf{R}^n)$ and is strongly separated from $\partial u_j(\mathbf{R}^n) \subset \operatorname{conv}(\overline{\Omega}_j)$ by Π. The claim therefore follows from Theorem 2.3. □

For the quadratic cost, this result allows us to deduce a variant of Proposition 4.9 which requires neither convexity of $\operatorname{spt}\mu$ nor *strict* convexity of $\overline{\Omega}_1$:

Corollary 5.3 (Continuous optimal maps to convex target pieces)**.** *Fix $c(x,\bar{x}) = -\langle x,\bar{x}\rangle$ and absolutely continuous probability measures μ and ν on $\mathbf{R}^n$ whose densities are bounded away from zero and infinity on their (compact) supports. Let $u = \max u_i$ be from Lemma 4.4. Assume $\overline{\Omega}_1$ is convex, and disjoint from $\operatorname{conv}(\overline{\Omega}_i)$ for each $i > 1$ such that $\Sigma^{\uparrow}_i$ intersects $\Omega_1 := (\operatorname{spt}\mu)\cap\Sigma^{\uparrow}_1$. If, in addition $(\operatorname{spt}\mu)^{\partial}\cap\Sigma^{\uparrow}_1$ has zero volume, then $Du_1 \in C^{\alpha}_{loc}(\Omega_1^{\mathrm{int}})$ and is injective on Ω_1^{int}.*

Proof. The boundary of Ω_1 is contained in the union of those $\Sigma^{\uparrow}_{1,i}$ intersecting $\operatorname{spt}\mu$ and $(\operatorname{spt}\mu)^{\partial} \cap \Sigma^{\uparrow}_1$. Corollary 5.2 shows the former are DC hypersurfaces, hence contain zero volume, like the latter. Caffarelli's results [5] now assert $u_1 \in C^{1,\alpha}_{loc}(\Omega_1^{\rm int})$ and is strictly convex there. □

In the above corollary, Du_1 gives a homeomorphism between the interior of $\Omega_1 := (\operatorname{spt}\mu) \cap \Sigma^{\uparrow}_1$ and some open subset $V_1 := Du_1(\Omega_1^{\rm int})$ of full volume in $\overline{\Omega}_1$; however, the price we pay for the lack of convexity of $\operatorname{spt}\mu$ is that we can no longer conclude differentiability of u_1 up to the boundary of Ω_1 because we cannot preclude the possibility that u^* fails to be strictly convex along a segment in $\overline{\Omega}_1 \setminus V_1$.

The next theorem shows that $\Sigma^{\uparrow}_{i_1,\dots,i_k}$ is a disjoint union of $\Sigma^{\uparrow}_{i_1,\dots,i_k} \cap M_k$ and $\bigcup_{j\in I\setminus\{i_1,\dots,i_k\}} \Sigma^{\uparrow}_{i_1,\dots,i_k,j}$: the first being a DC submanifold of codimension $k-1$, the second a finite union of closed sets with Hausdorff dimension at most $n-k$. For implications of affine independence in a simpler setting, see the C^1 description of higher codimension tears coming from strictly convex target components in Corollary 4.14.

Theorem 5.4 (DC rectifiability of higher multiplicity tears). *Fix $c(x,\bar{x}) = -\langle x, \bar{x}\rangle$ and probability measures μ and ν on $\mathbf{R}^n$ with μ absolutely continuous and $\operatorname{spt}\nu = \cup_{i\in I}\overline{\Omega}_i$ a finite disjoint union of compact sets. Let $\nu = (Du)_{\#}\mu$ where u from (5.5) is convex. Fix a collection of indices $i_1, \dots, i_k \in I$. If $\{\operatorname{conv}(\overline{\Omega}_{i_1}), \dots, \operatorname{conv}(\overline{\Omega}_{i_k})\}$ is an affinely independent collection, for any $x_0 \in \Sigma_{i_1,\dots,i_k}$ there exists $r_0 > 0$ such that $B_{r_0}(x_0) \cap \Sigma_{i_1,\dots,i_k}$ is contained in the image of an open subset of $\mathbf{R}^{n+1-k}$ under a bi-Lipschitz DC mapping.*

Suppose in addition that the existence of a point x such that $\partial u(x) \cap \overline{\Omega}_i \neq \emptyset$ for all of $i = i_1, \dots, i_k$, and j implies

$$\{\operatorname{conv}(\overline{\Omega}_{i_1}), \dots, \operatorname{conv}(\overline{\Omega}_{i_k}),\ \operatorname{conv}(\overline{\Omega}_j)\} \text{ is an affinely independent collection.} \tag{5.7}$$

Then

$$\Sigma^{\uparrow}_{i_1,\dots,i_k} \cap M_k = \{x \in \mathbf{R}^n \mid u(x) = u_{i_1}(x) = \dots = u_{i_k}(x) > \max_{j\in I\setminus\{i_1,\dots,i_k\}} u_j(x)\}, \tag{5.8}$$

$$(\Sigma^{\uparrow}_{i_1,\dots,i_k} \cap M_k) \cap \bigcup_{j\in I\setminus\{i_1,\dots,i_k\}} \Sigma^{\uparrow}_{i_1,\dots,i_k,j} = \emptyset, \tag{5.9}$$

$$(\Sigma^{\uparrow}_{i_1,\dots,i_k} \cap M_k) \cup \bigcup_{j\in I\setminus\{i_1,\dots,i_k\}} \Sigma^{\uparrow}_{i_1,\dots,i_k,j} = \Sigma^{\uparrow}_{i_1,\dots,i_k}. \tag{5.10}$$

Moreover $\Sigma^{\uparrow}_{i_1,\dots,i_k} \cap M_k$ is a relatively open subset of $\Sigma^{\uparrow}_{i_1,\dots,i_k}$.

Proof. Without loss of generality, we may assume $i_1 = 1, \dots, i_k = k$.

First assume $\{\operatorname{conv}(\overline{\Omega}_i)\}_{i=1}^k$ is an affinely independent collection and $x_0 \in \Sigma_{1,\dots,k}$. Defining $F : \mathbf{R}^{n+1-k} \times \mathbf{R}^{k-1} \to \mathbf{R}^{k-1}$ by

$$F(x) := (u_1(x) - u_k(x), \dots, u_{k-1}(x) - u_k(x)),$$

by assumption $F(x_0) = 0$, we will now show that every element of $J^C F(x_0)$ has rank $k-1$. Let $M \in J^C F(x_0)$, and suppose the ith row is given by a vector of the form

$$v_i := \lim_{m\to\infty} \nabla(u_i - u_k)(x_m)$$

with $x_m \to x_0$ and $x_m \in \mathrm{Dom}(\nabla u_i) \cap \mathrm{Dom}(\nabla u_k)$. Then there must exist points $\bar{x}_i \in \overline{\Omega}_i$ for $i \in \{1,\dots,k\}$ such that $v_i = \bar{x}_i - \bar{x}_k$, and the assumption of affine independence implies M has rank $k-1$. By Carathéodory's theorem ([31, Theorem 17.1] any other $M \in J^C F(x_0)$ can be written as the convex combination of $n+1$ matrices as above, meaning that we have $v_i = \bar{x}_i - \bar{x}_k$ this time with $\bar{x}_i \in \mathrm{conv}(\overline{\Omega}_i)$ for $i \in \{1,\dots,k\}$, again the hypothesis yields that M has rank $k-1$. Thus we can apply the DC constant rank theorem (Theorem 3.9) to obtain the first claim.

Now assume condition (5.7) holds. For brevity, let us notate the set on the right hand side of (5.8) by S_k. Suppose $u(x_0) = u_i(x_0)$ for any fixed index $i \in I$, then by Lemma 2.5 we have $\partial u_i(x_0) \subset \partial u(x_0)$. Any extremal point of $\partial u_i(x_0)$ is a limit of points of the form $\nabla u_i(x_m)$ where $x_m \in \mathrm{Dom}(\nabla u_i)$ and $x_m \to x_0$, then since $\nabla u_i(\mathrm{Dom}(\nabla u)) \subset \overline{\Omega}_i$ which is a closed set, we see $\partial u(x_0) \cap \overline{\Omega}_i \neq \emptyset$. Thus, we immediately see $\Sigma^{\uparrow}_{1,\dots,k} \cap M_k \subset S_k$. On the other hand suppose $x_0 \in S_k$, then by definition $x_0 \in \Sigma^{\uparrow}_{1,\dots,k}$. Suppose by contradiction $x_0 \notin M_k$, then there must exist $j \in I \setminus \{1,\dots,k\}$ such that $\exists \bar{x}_0 \in \partial u(x_0) \cap \overline{\Omega}_j$. Since $\partial u(x_0) \cap \overline{\Omega}_i \neq \emptyset$ for $i \in \{1,\dots,k\}$ by Lemma 2.5, (5.7) implies the collection

$$\{\mathrm{conv}(\overline{\Omega}_1), \dots, \mathrm{conv}(\overline{\Omega}_k),\ \mathrm{conv}(\overline{\Omega}_j)\}$$

is affinely independent. However, by Lemma 2.5 and the definition of S_k, we must have that $\bar{x}_0$ is contained in the convex hull of k points, one from each of $\{\mathrm{conv}(\overline{\Omega}_1), \dots, \mathrm{conv}(\overline{\Omega}_k)\}$ contradicting this affine independence, proving (5.8). The claim (5.9) then follows immediately.

Next, by continuity of the u_i and u we immediately see

$$\Sigma^{\uparrow}_{1,\dots,k} \subset (\Sigma^{\uparrow}_{1,\dots,k} \cap M_k) \cup \bigcup_{j\in I\setminus\{1,\dots,k\}} \Sigma^{\uparrow}_{1,\dots,k,j},$$

while by (5.8) the opposite inclusion holds proving (5.10).

Finally, suppose $x \in \Sigma^{\uparrow}_{1,\dots,k} \cap M_k$. By (5.8), there is some open ball $B_r(x)$ on which $\min_{1\le i\le k} u_i > \max_{k+1\le j\le K} u_j$. Then clearly $B_r(x) \cap \Sigma^{\uparrow}_{1,\dots,k} \subset \Sigma^{\uparrow}_{1,\dots,k} \cap M_k$, hence $\Sigma^{\uparrow}_{1,\dots,k} \cap M_k$ is relatively open in $\Sigma^{\uparrow}_{1,\dots,k}$. □

We also mention that under affine independence, there can be at most one tear of multiplicity $n+1$.

Proposition 5.5 (Uniqueness of maximal multiplicity tears). *Let $c(x,\bar{x}) = -\langle x, \bar{x}\rangle$, and assume μ, ν are absolutely continuous probabilities measures on $\mathbf{R}^n$ with bounded supports. Also suppose $\{\overline{\Omega}_i\}_{i=1}^{n+1}$ is* any *affinely independent collection of path connected subsets of $\mathbf{R}^n$ (which may or may not decompose* spt ν*). Then if u is an optimal potential transporting μ to ν, it can have at most one point of multiplicity $n+1$ relative to $\{\overline{\Omega}_i\}_{i=1}^{n+1}$.*

Proof. Suppose by contradiction there exist two points $x_0 \neq y_0$ where u has multiplicity $n+1$, then $\partial u(x_0)$ and $\partial u(y_0)$ each must intersect all of the sets $\overline{\Omega}_i$. First note that $\partial u(x_0)$, $\partial u(y_0)$ must have affine dimension n (hence nonempty interior), otherwise there would be an $n-1$ dimensional affine plane intersecting all $\overline{\Omega}_i$. Now the convex function u^* is seen to be nondifferentiable on $\partial u\,(x_0) \cap \partial u\,(y_0)$, hence this intersection must have zero Lebesgue measure. In particular, the interiors of $\partial u\,(x_0)$ and $\partial u\,(y_0)$ are disjoint, and by [31, Theorem 11.3], $\mathbf{R}^n$ is divided into two closed, opposing halfspaces H_+ and H_- with $\partial u\,(x_0) \subset H_+$, $\partial u\,(y_0) \subset H_-$.

Let us take $\bar{x}_i \in \partial u\,(x_0) \cap \overline{\Omega}_i$ and $\bar{y}_i \in \partial u\,(y_0) \cap \overline{\Omega}_i$; we see that $\bar{x}_i \in H_+$ while $\bar{y}_i \in H_-$ for each $1 \leq i \leq n+1$. Now each $\overline{\Omega}_i$ is path connected, thus for each i there exists some continuous path $\gamma_i(t)$ with $\gamma_i(0) = \bar{x}_i$ and $\gamma_i(1) = \bar{y}_i$, which remains inside $\overline{\Omega}_i$. Clearly there must exist some time $t_i \in [0, 1]$ at which γ_i intersects the hyperplane $H_+ \cap H_-$ for each $1 \leq i \leq n+1$. However, this would imply that $H_+ \cap H_-$ is an $n-1$ dimensional affine plane intersecting all of the sets $\overline{\Omega}_i$, a contradiction. □

6. $C^{1,\alpha}$ SMOOTHNESS OF OPTIMAL MAP DISCONTINUITIES: QUADRATIC COST

In a previous section, affine independence of the target pieces was identified as the geometric manifestation of the implicit function theorem hypothesis which guarantees DC smoothness of the corresponding tears. This section is devoted to improving this smoothness to $C^{1,\alpha}_{loc}$ on $(\operatorname{spt}\mu)^{\text{int}}$. In order to establish this goal, we begin by recalling the required machinery from [6]. Again, we will be working in the setting of $c(x, \bar{x}) = -\langle x, \bar{x}\rangle$ on $\mathbf{R}^n \times \mathbf{R}^n$.

Definition 6.1 (Affine doubling)**.** Suppose μ is a Borel measure on $\mathbf{R}^n$ and $x \in X \subset \mathbf{R}^n$. An open neighborhood $\mathcal{N}_x$ of x is said to be a *doubling neighborhood of μ with respect to X* if there exists a constant $\delta > 0$ (called the *doubling constant of μ on $\mathcal{N}_x$*) such that for any convex set $Z \subset \mathcal{N}_x$ whose (Lebesgue) barycenter is in X,

$$\mu(\frac{1}{2}Z) \geq \delta^2 \mu(Z),$$

here the dilation of Z is with respect to its barycenter.

Definition 6.2 (Centered sections)**.** If $\phi : \mathbf{R}^n \to \mathbf{R} \cup \{\infty\}$ is a convex function with $\partial\phi\,(\mathbf{R}^n)^{\text{int}} \neq \emptyset$, $\varepsilon > 0$, and $x_0 \in \mathbf{R}^n$, the *centered section* of ϕ at x_0 of height ε is defined by

$$Z^{\phi}_{\varepsilon}(x_0) := \{x \in \mathbf{R}^n \mid \phi(x) < \varepsilon + \phi(x_0) + \langle v_\varepsilon, x - x_0\rangle\}$$

where v_ε is chosen so that x_0 is the barycenter of $Z^{\phi}_{\varepsilon}(x_0)$, which is bounded.

It is known that such a v_ε exists, and is unique (see [6, Theorems A.7 and A.8]).

Definition 6.3 (p-uniform convexity)**.** If $p \geq 2$, a convex function u is said to be *p-uniform convex* on a set Ω if there is a finite constant $k > 0$ such that for any x_1, $x_2 \in \Omega$ and $\bar{x}_1 \in \partial u\,(x_1)$, $\bar{x}_2 \in \partial u\,(x_2)$,

$$\langle \bar{x}_2 - \bar{x}_1, x_2 - x_1\rangle \geq k^{1-p}|x_2 - x_1|^p.$$

Remark 6.4. This definition differs from the *a priori* weaker [6, Definition 7.9], but is equivalent. Indeed, the above inequality still holds if $\bar{x}_1$, $\bar{x}_2$ are points that can be written as limits of the form $\lim_{k\to\infty} \nabla u(x_{k,i})$ where $x_{k,i} \in \operatorname{Dom}(\nabla u) \cap \Omega$ and $x_{k,i} \to x_i$ as $k \to \infty$. Then since any $\bar{x}_i \in \partial u\,(x_i)$ can be written as a convex combination of such points, the formulation in Definition 6.3 holds.

With these definitions in hand, we can state and prove the following refinement in the case when one of the pieces $\overline{\Omega}_i$ is strictly convex. Corollary 6.6 below will give conditions under which the exceptional sets E_i of the theorem below lie in the boundary of $\operatorname{spt}\mu$.

Theorem 6.5 (Hölder continuity of optimal maps to closed convex target pieces)**.** *Fix $c(x,\bar{x}) = -\langle x, \bar{x}\rangle$ and probability measures μ, ν with densities bounded away from zero and infinity on their supports in $\mathbf{R}^n$. Let $\operatorname{spt}\mu$ be convex and $\operatorname{spt}\nu = \cup_{i\in I}\bar{\Omega}_i$ a finite disjoint union of closed sets strongly separated by hyperplanes pairwise* (5.6)*. If $\overline{\Omega}_i$ is strictly convex, then*

$$u_i \in C^{1,\alpha}_{loc}((\Sigma_i^{\uparrow} \cap \operatorname{spt}\mu) \setminus E_i)$$

for some $\alpha \in (0,1)$ (which depends only n, and the bounds of the densities of μ and ν away from zero and infinity on their supports) where

$$E_i := \{x \in (\Sigma_i^{\uparrow} \cap \operatorname{spt}\mu)^{\partial} \mid \nabla u_i(x) \in N_{\operatorname{spt}\mu}(x) + \operatorname{conv}\left(\partial u\left(x\right) \cap (\operatorname{spt}\nu \setminus \bar{\Omega}_i)\right)\} \quad (6.1)$$

and $N_{\operatorname{spt}\mu}(x) := \{v \in \mathbf{R}^n \mid \langle v, y - x\rangle \leq 0 \text{ for all } y \in \operatorname{spt}\mu\}$ denotes the outer normal cone to the convex set $\operatorname{spt}\mu$ at x.

Proof. We may assume $i = 1$. Proposition 4.9 asserts that $u_1 \in C^1(\mathbf{R}^n)$, and Corollary 4.10 implies ∇u gives a homeomorphism between $(\operatorname{spt}\nu \cap \Sigma_1^{\uparrow})^{\mathrm{int}}$ and $\bar{\Omega}_i^{\mathrm{int}}$ which extends continuously to the boundary. The purpose of this theorem is to establish a Hölder estimate away from the exceptional set E_1.

Let us write for any Borel $A \subset \mathbf{R}^n$, $M_1(A) := |\nabla u_1(A)|_{\mathcal{L}}$, the *Monge-Ampère measure* of u_1 (here $|\cdot|_{\mathcal{L}}$ denotes the Lebesgue measure). Since $\nabla u_1(\mathbf{R}^n) \subset \overline{\Omega}_1$ which is convex, by [5, Lemma 2] we have for some constant $C > 0$ depending only the bounds of the densities of μ and ν away from zero and infinity on their supports, for any Borel $A \subset \mathbf{R}^n$

$$C^{-1}|A \cap \Sigma_1^{\uparrow} \cap \operatorname{spt}\mu|_{\mathcal{L}} \leq M_1(A) \leq C|A \cap \Sigma_1^{\uparrow} \cap \operatorname{spt}\mu|_{\mathcal{L}}. \quad (6.2)$$

Suppose $x_0 \in (\operatorname{spt}\mu)^{\partial} \cap (\Sigma_1^{\uparrow})^{\mathrm{int}}$. Then for some $r_0 > 0$ small, the intersection $B_{r_0}(x_0) \cap \operatorname{spt}\mu \cap \Sigma_1^{\uparrow}$ is convex and any convex $Z \subset B_{r_0}(x_0) \cap \operatorname{spt}\mu \cap \Sigma_1^{\uparrow}$ satisfies (6.2). Thus the proof of [6, Lemma 7.5] applies and we see $B_{r_0}(x_0)$ is a doubling neighborhood of M_1 with respect to $\operatorname{spt}\mu \cap \Sigma_1^{\uparrow}$, with doubling constant δ_0 depending only on μ, ν, and n.

Next define the convex function $\tilde{u}$ by

$$\tilde{u}(x) = \begin{cases} u(x), & x \in \operatorname{spt}\mu, \\ \infty, & \text{else,} \end{cases}$$

then its Legendre transform $\tilde{u}^*$ is an optimal potential transporting ν to μ which is finite on all of $\mathbf{R}^n$ with $\partial\tilde{u}^*(\mathbf{R}^n) \subset \operatorname{spt}\mu$ by convexity of $\operatorname{spt}\mu$. Since the restriction of $\tilde{u}^*$ will be an optimal potential transporting the restriction of ν to $\overline{\Omega}_1$ to the restriction of μ to $\Sigma_1^{\uparrow} \cap \operatorname{spt}\mu$ and $\overline{\Omega}_1$ is connected, by subtracting a constant we can assume $\tilde{u}^* = u^*$ on $\overline{\Omega}_1$. Writing for any Borel $A \subset \mathbf{R}^n$, $\tilde{M}(A) := |\partial\tilde{u}^*(A)|_{\mathcal{L}}$ (the *Monge-Ampère measure* of $\tilde{u}^*$), by [5, Lemma 2] we then have for some constant $C > 0$ depending only the bounds of the densities of μ and ν away from zero and infinity on their supports, for any Borel $A \subset \mathbf{R}^n$

$$C^{-1}|A \cap \operatorname{spt}\nu|_{\mathcal{L}} \leq \tilde{M}(A) \leq C|A \cap \operatorname{spt}\nu|_{\mathcal{L}}.$$

In turn, since $\overline{\Omega}_1$ is convex we find the proof of [6, Lemma 7.5] applies hence for any $x \in \overline{\Omega}_1$ and $r > 0$ such that $B_r(x) \cap \bigcup_{i \in I \setminus \{1\}} \overline{\Omega}_i = \emptyset$, the open ball $B_r(x)$ is a doubling neighborhood of $\tilde{M}$ with respect to $\overline{\Omega}_1$, with doubling constant δ_0 depending only on μ, ν, and n.

Next, we will show that for $r > 0$ fixed, there is some $\varepsilon_0 > 0$ such that whenever $x \in \Sigma_1^{\uparrow} \cap \operatorname{spt}\mu$ and $\bar{x} = \nabla u_1(x)$ are such that

$$(\nabla u_1)^{-1}(B_r(\bar{x})) \cap E_1 = \emptyset, \tag{6.3}$$

and $\varepsilon < \varepsilon_0$, then the centered section $Z_\varepsilon^{\tilde{u}^*}(\bar{x}) \subset B_r(\bar{x})$. The proof will closely follow that of [6, Lemma 7.11]. Suppose the claim fails: for some fixed $r > 0$ there exist sequences $\bar{x}_j = \nabla u_1(x_j)$ with $x_j \in \Sigma_1^{\uparrow} \cap \operatorname{spt}\mu$ satisfying (6.3), $\varepsilon_j \searrow 0$ with $Z_{\varepsilon_j}^{\tilde{u}^*}(\bar{x}_j) \not\subset B_r(\bar{x}_j)$. Extracting subsequences yields $\bar{x}_j \to \bar{x}_\infty$ and $x_j \to x_\infty$ with $\nabla u_1(x_\infty) = \bar{x}_\infty \in \overline{\Omega}_1$, still satisfying (6.3); let us also define

$$Z_{\min} := \{\bar{x} \in \mathbf{R}^n \mid \tilde{u}^*(\bar{x}) = \tilde{u}^*(\bar{x}_\infty) + \langle \bar{x} - \bar{x}_\infty, x_\infty \rangle\} = \partial \tilde{u}\,(x_\infty)\,.$$

We can see that Claim #1 in the proof of [6, Lemma 7.11] still holds, so in particular there is a nontrivial line segment contained in $Z_{\min}$, centered at $\bar{x}_\infty$ but otherwise disjoint from the set $\overline{\Omega}_1$ on which $\tilde{u}$ is strictly convex. Thus $\bar{x}_\infty \in (\overline{\Omega}_1)^\partial$ and Corollary 4.10 implies $x_\infty \in (\Sigma_1^{\uparrow} \cap \operatorname{spt}\mu)^\partial$. Reordering if necessary, we may assume $u_i(x_\infty)$ depends monotonically on i, with $u_1(x_\infty) = u_2(x_\infty) = \cdots = u_k(x_\infty) > u_{k+1}(x_\infty)$ for some $k \geq 1$. Then

$$\begin{aligned} \partial \tilde{u}\,(x_\infty) &= \partial u\,(x_\infty) + N_{\operatorname{spt}\mu}(x_\infty) \\ &= \operatorname{conv}(\{\bar{x}_\infty\} \cup \partial u_2\,(x_\infty) \cup \cdots \cup \partial u_k\,(x_\infty)) + N_{\operatorname{spt}\mu}(x_\infty) \end{aligned} \tag{6.4}$$

decomposes as the sum of a bounded component and a convex cone, in view of Lemma 2.5. Since (6.3) for $\bar{x}_k$ implies $(\nabla u_1)^{-1}(B_r(\bar{x}_\infty)) \cap E_1 = \emptyset$, we see $\bar{x}_\infty$ is not contained in the closed convex set

$$\operatorname{conv}(N_{\operatorname{spt}\mu}(x_\infty) + \cup_{i=2}^k \partial u_i\,(x_\infty)) = \operatorname{conv}(\partial u\,(x_\infty) \cap (\operatorname{spt}\nu \setminus \bar{\Omega}_1)),$$

hence can be strongly separated from it by a hyperplane ([31, Corollary 11.4.2]). Any segment in (6.4) centered at x_∞ must be parallel to this hyperplane. But this can only occur if the closed cone $N_{\operatorname{spt}\mu}(x_\infty)$ contains a complete line parallel to this segment, contradicting the fact that $\operatorname{spt}\mu$ has non-empty interior.

Thus [6, Theorem 7.13 and Corollary 7.14] will apply (note that differentiability of $\tilde{u}^*$ is not actually necessary to do this), proving that u is $C^{1,\alpha}_{loc}$ on $(\Sigma_1^{\uparrow} \cap \operatorname{spt}\mu) \setminus E_1$. □

In addition to giving conditions under which the exceptional sets E_i of the theorem above lie in the boundary of $\operatorname{spt}\mu$, the following corollary shows the codimension k submanifolds of Corollary 4.14 enjoy Hölder differentiability, except possibly where they intersect the boundary $(\operatorname{spt}\mu)^\partial$ tangentially.

Corollary 6.6 (Hölder regularity away from tangential tear-boundary intersections)**.** *Fix $x \in E_1$ in Theorem 6.5. Assume $u_i(x) \geq u_{i+1}(x)$ for all $i \in I$, and $u_1(x) = u_k(x) > u_{k+1}(x)$. Also suppose the collection $\{\operatorname{conv}\left(\partial u\,(x) \cap \overline{\Omega}_i\right)\}_{i=1}^k$ is affinely independent. If $\overline{\Omega}_1$ is strictly convex then $x \in (\operatorname{spt}\mu)^\partial$. If additionally, $\overline{\Omega}_i$ is strictly convex for all $i \leq k$ and $(\operatorname{spt}\mu)^\partial$ is differentiable at x, then $\Sigma_{\{1,2,\dots,k\}}$ intersects $(\operatorname{spt}\mu)^\partial$ tangentially, meaning that the outer unit normal to $\operatorname{spt}\mu$ at x is also normal to the C^1 submanifold $\Sigma_{\{1,2,\dots,k\}}$. In this case, $\Sigma^{\uparrow}_{\{1,2,\dots,k\}} \cap \operatorname{spt}\mu$ is*

$C^{1,\alpha}_{loc}$ *smooth, away from any such tangential intersections (and any possible non-differentiabilities of* $(\operatorname{spt}\mu)^\partial$*).*

Proof. Suppose $x \in E_1 \subset \Sigma_1^\uparrow \cap \operatorname{spt}\mu$. By our assumptions and Lemma 2.5, we have $\partial u(x) \subset \operatorname{conv}\left(\bigcup_{i=1}^k \overline{\Omega}_i\right)$, hence

$$\operatorname{conv}\left(\partial u(x) \cap (\operatorname{spt}\nu \setminus \bar{\Omega}_1)\right) \subset \operatorname{conv}\left(\bigcup_{i=2}^k (\partial u(x) \cap \overline{\Omega}_i)\right) = \bigcup_{i=2}^k \operatorname{conv}\left(\partial u(x) \cap \overline{\Omega}_i\right)$$

Thus there exist $\bar{x}_i \in \operatorname{conv}\left(\partial u(x) \cap \overline{\Omega}_i\right)$ and $t_i \geq 0$ with $1 = \sum_{i=2}^k t_i$ such that

$$\sum_{i=2}^k t_i(\nabla u_1(x) - \bar{x}_i) \in N_{\operatorname{spt}\mu}(x) \tag{6.5}$$

according to (6.1) of Theorem 6.5. Setting $\bar{x}_1 = \nabla u_1(x)$, the affine independence of $\{\bar{x}_i\}_{i \leq k}$ makes $\{\bar{x}_1 - \bar{x}_i\}_{2 \leq i \leq k}$ linearly independent. Thus $\sum_{i=2}^k t_i = 1$ forces the sum in (6.5) not to vanish.

Now $x \in (\operatorname{spt}\mu)^{\text{int}}$ would force $N_{\operatorname{spt}\mu}(x) = \{0\}$, contradicting the last sentence. Thus we conclude x is contained in the boundary of $\operatorname{spt}\mu$. If, in addition, $\overline{\Omega}_i$ is strictly convex for all $i \leq k$ then $\nabla u_1(x) - \bar{x}_i = \nabla u_1(x) - \nabla u_i(x)$ is a (non-zero) normal to the hypersurface $\Sigma_{\{1,i\}} = \{u_1 = u_i\}$, which is C^1 smooth by Corollary 4.14, noting that a collection of two sets is affinely independent if they are disjoint. Thus the sum in (6.5) is normal to the codimension $k-1$ submanifold $\Sigma_{\{1,\ldots,k\}} = \cap_{i=2}^k \Sigma_{\{1,i\}}$ of the same corollary. Since (6.5) is non-vanishing, it is an outer normal to $\operatorname{spt}\mu$ when the latter is differentiable at x. Away from such points, the improvement in regularity from C^1 to $C^{1,\alpha}_{loc}$ comes from Theorem 6.5 and the implicit function theorem. □

When $k = 2$ and both target components are strictly convex, an analous result was shown simultaneously and independently from us by Chen [7], who went on to show $C^{2,\alpha}$ regularity of the tear provided the target components are sufficiently far apart.

7. Global structure of optimal map discontinuities: MTW costs

For quadratic transportation costs, we have already shown that when the support of the target measure consists of a number of affinely independent regions, the optimal transport map induces a partition of the source domain into sets corresponding to each of these regions. In this section, we extend one such result — Proposition 5.1 — to MTW costs. While we expect other results from Sections 5 and 6 also to have analogs for such costs, we do not pursue such extensions in the present manuscript.

Theorem 7.1 (Pairwise partitions of source). *Suppose the cost function c satisfies* (B1) *and* (MTW)*, and Ω and $\overline{\Omega}$ are c-convex with respect to each other. Also suppose μ and ν are absolutely continuous probability measures on Ω and $\overline{\Omega}$ respectively, where $\operatorname{spt}\nu = \overline{\Omega}_1 \cup \overline{\Omega}_2$ is such that there exists $\bar{x}_0 \in \overline{\Omega}$ for which the sets $\bigcup_{x\in\Omega}[-D_{\bar{x}}D_x c(x,\bar{x}_0)]^{-1}(\left[\overline{\Omega}_1\right]_x)$ and $\bigcup_{x\in\Omega}[-D_{\bar{x}}D_x c(x,\bar{x}_0)]^{-1}(\left[\overline{\Omega}_2\right]_x)$ are strongly separated by a hyperplane $\Pi \subset T_{\bar{x}_0}\bar{M}$.*

Then an optimal potential u transporting μ to ν can be written as $u = \max\{u_1, u_2\}$, where u_1 and u_2 are c-convex functions such that

$$exp_x^c(Du_i(x)) \in \overline{\Omega}_i, \quad a.e.\ x \in \Omega. \tag{7.1}$$

Moreover, under the global coordinates induced by $x \mapsto -D_{\bar{x}}c(x, \bar{x}_0)$ on Ω, the sets $\{u_1 = u_2\}$, $\{u_1 > u_2\}$, and $\{u_1 < u_2\}$ in Ω are given by the graph, open epigraph, and open subgraph respectively of a DC function h defined on the projection of $[\Omega]_{\bar{x}_0}$ onto the hyperplane Π.

Additionally, if

$$\left[\bigcup_{x\in\Omega} exp_x^c(\operatorname{conv}\left(\left[\overline{\Omega}_1\right]_x\right))\right] \cap \left[\bigcup_{x\in\Omega} exp_x^c(\operatorname{conv}\left(\left[\overline{\Omega}_2\right]_x\right))\right] = \emptyset, \tag{7.2}$$

then the sets $\{u_1 = u_2\}$, $\{u_1 \geq u_2\}$, and $\{u_1 \leq u_2\}$ are connected.

Proof of Theorem 7.1. Again Lemma 4.4 gives the representation $u = \max\{u_1, u_2\}$ and (7.1), note we have not exploited any convexity properties of the $\overline{\Omega}_i$ so far. Write

$$\begin{aligned}
\Omega_= &:= \{x \in \Omega \mid u_1(x) = u_2(x)\},\\
\Omega_< &:= \{x \in \Omega \mid u_1(x) < u_2(x)\}, \quad \Omega_> := \{x \in \Omega \mid u_1(x) > u_2(x)\},\\
\Omega_\leq &:= \{x \in \Omega \mid u_1(x) \leq u_2(x)\}, \quad \Omega_\geq := \{x \in \Omega \mid u_1(x) \geq u_2(x)\}.
\end{aligned}$$

Now make a change of variables under $exp_{\bar{x}_0}^{c^*}(\cdot)$ and define $\tilde{u}_i : [\Omega]_{\bar{x}_0} \to \mathbf{R}$ by

$$\tilde{u}_i(p) := u_i(exp_{\bar{x}_0}^{c^*}(p)), \quad \tilde{u} := \max\{\tilde{u}_1, \tilde{u}_2\}.$$

We will identify $T^*_{\bar{x}_0}\bar{M} \cong T_{\bar{x}_0}\bar{M} \cong \mathbf{R}^n$, and without loss of generality assume the separating hyperplane Π is $\{p^n = a_0\}$ for some $a_0 \in \mathbf{R}$, with width $d_0 > 0$.

Now take any point $p_0 \in \Omega^{\text{int}}$ with $\tilde{u}_1(p_0) = \tilde{u}_2(p_0)$. For a sufficiently small $r > 0$, there is some $C > 0$ for which $\tilde{\tilde{u}}_i := \tilde{u}_i + \frac{C}{2}|p - p_0|^2$ are both convex functions on $B_r(p_0) \subset \Omega$. Since c satisfies (MTW) and Ω and $\overline{\Omega}$ are c-convex with respect to each other, writing $x_0 := exp_{\bar{x}_0}^{c^*}(p_0)$ we have for $i = 1$, 2,

$$\begin{aligned}
\partial\tilde{\tilde{u}}_i\,(p_0) = \partial\tilde{u}_i\,(p_0) &= [-D_{\bar{x}}D_xc(x_0, \bar{x}_0)]^{-1}(\partial u_i\,(x_0))\\
&= [-D_{\bar{x}}D_xc(x_0, \bar{x}_0)]^{-1}\,[\partial_c u_i(x_0)]_{x_0}\\
&\subset \operatorname{conv}\left([-D_{\bar{x}}D_xc(x_0, \bar{x}_0)]^{-1}(\left[\overline{\Omega}_i\right]_{x_0})\right),
\end{aligned}$$

which are strongly separated from each other by $\{p^n = a_0\}$ with spacing d_0 by assumption. Then by [31, Corollary 24.5.1], there is some $r > 0$ for which $\partial\tilde{\tilde{u}}_1\,(B_r(p_0))$ and $\partial\tilde{\tilde{u}}_2\,(B_r(p_0))$ are still strongly separated with spacing d_0. Let us write $\tilde{\Omega}_=$, $\tilde{\tilde{\Omega}}_=$ for $\Omega_=$ with u_i replaced by $\tilde{u}_i$ or $\tilde{\tilde{u}}_i$ and Ω by $[\Omega]_{\bar{x}_0}$ or $[\Omega]_{\bar{x}_0} \cap B_r(p_0)$ respectively (and likewise for $<$ and $>$). Then we may apply Corollary 2.6 to find that the sets $\tilde{\tilde{\Omega}}_=$, $\tilde{\tilde{\Omega}}_<$, and $\tilde{\tilde{\Omega}}_>$ are the graph, open subgraph, and open epigraph respectively of the function

$$p' \mapsto \frac{-\tilde{\tilde{u}}^*_{p'}(a_0 - d_0) + \tilde{\tilde{u}}^*_{p'}(a_0 + d_0)}{2d_0} = h(p')$$

over $B_r(p_0) \subset \mathbf{R}^{n-1}$ where again, $\tilde{\tilde{u}}^*_{p'}$ is the Legendre transform in just the nth variable. Now by [31, Theorem 16.4] we see that

$$\tilde{\tilde{u}}^*_{p'}(a_0 - d_0) = \inf_{s\in\mathbf{R}}(\tilde{u}^*_{p'}(a_0 - d_0 - s) + \left(\frac{C(|p' - p_0'|^2 + (\cdot - p_0^n)^2)}{2}\right)^*(s)),$$

and a quick calculation yields

$$\begin{aligned}\left(\frac{C(|p' - p_0'|^2 + (\cdot - p_0^n)^2)}{2}\right)^*(s) &= \sup_{t\in\mathbf{R}}(ts - \frac{C}{2}(|p' - p_0'|^2 + (t - p_0^n)^2)) \\ &= \sup_{t\in\mathbf{R}}(ts - \frac{C}{2}(t - p_0^n)^2) - \frac{C}{2}|p' - p_0'|^2 \\ &= \frac{s^2}{2C} + sp_0^n - \frac{C(p_0^n)^2}{2} - \frac{C}{2}|p' - p_0'|^2. \qquad (7.3)\end{aligned}$$

At the same time, from the proof of Theorem 2.3 we see that $\tilde{u}^*_{p'}$ is a convex function with $p_0^n \in \partial\tilde{u}^*_{p'}(a_0 - d_0)$. Thus by (7.3) we find,

$$\begin{aligned}\tilde{\tilde{u}}^*_{p'}(a_0 - d_0) &= \inf_{s\in\mathbf{R}}(\tilde{u}^*_{p'}(a_0 - d_0 - s) + \frac{s^2}{2C} + sp_0^n - \frac{C(p_0^n)^2}{2} - \frac{C}{2}|p' - p_0'|^2) \\ &\geq \tilde{u}^*_{p'}(a_0 - d_0) - sp_0^n + \frac{s^2}{2C} + sp_0^n - \frac{C(p_0^n)^2}{2} - \frac{C}{2}|p' - p_0'|^2 \\ &\geq \tilde{u}^*_{p'}(a_0 - d_0) - \frac{C(p_0^n)^2}{2} - \frac{C}{2}|p' - p_0'|^2,\end{aligned}$$

with equality achieved for the choice $s = 0$. A similar calculation shows

$$\tilde{\tilde{u}}^*_{p'}(a_0 + d_0) = \tilde{u}^*_{p'}(a_0 + d_0) - \frac{C(p_0^n)^2}{2} - \frac{C}{2}|p' - p_0'|^2$$

hence

$$h(p') = \frac{-\tilde{u}^*_{p'}(a_0 - d_0) + \tilde{u}^*_{p'}(a_0 + d_0)}{2d_0},$$

the significance being that this function does not depend on the constant C in $\tilde{\tilde{u}}$, hence is independent of the point p_0. Since $\tilde{\tilde{u}}_1(p) = \tilde{\tilde{u}}_2(p)$ for $p \in B_r(p_0)$ if and only if $\tilde{u}_1(p) = \tilde{u}_2(p)$, by the continuity of u_i up to the boundary of Ω we obtain that $\tilde{\Omega}_=$, $\tilde{\Omega}_<$, $\tilde{\Omega}_>$ are equal to the graph, open subgraph, open epigraph respectively of h over the projection of $[\Omega]_{\bar{x}_0}$ on Π.

Now assume condition (7.2) holds. Since $\partial_c u_i(x) \subset exp^c_x(\text{conv}\left([\overline{\Omega}_i]_x\right))$, by (7.1) combined with [26, Theorem 3.1], we see that

$$\partial_c u_1(\Omega) \cap \partial_c u_2(\Omega) = \emptyset.$$

We now claim

$$\partial_{c^*} u^c(\partial_c u_1(\Omega)) = \Omega_\geq, \qquad (7.4)$$

which concludes the proof by Lemma 4.8 (and a symmetric argument switching the roles of u_1 and u_2).

Suppose $u_1(x) \geq u_2(x)$. Then since $[\partial_c u_i(x)]_x = \partial u_i(x)$, Lemma 2.5 yields that

$$\partial_c u(x) = \partial_c u_1(x) \text{ or } \partial_c u(x) = exp^c_x(\text{conv}(\partial u_1(x) \cup \partial u_2(x))).$$

In either case, there exists $\bar{x} \in \partial_c u_1(x) \cap \partial_c u(x)$ which implies $x \in \partial_{c^*} u(\bar{x})$, and in particular $x \in \partial_{c^*} u^c(\partial_c u_1(\Omega))$, thus $\partial_{c^*} u^c(\partial_c u_1(\Omega)) \supset \Omega_\geq$. On the other hand, suppose $x \in \partial_{c^*} u^c(\partial_c u_1(\Omega))$ but $u_1(x) < u_2(x)$. Then there exist $y \in \Omega$ and $\bar{x} \in \overline{\bar{\Omega}}$

with $\bar{x} \in \partial_c u_1(y)$ and $x \in \partial_{c^*} u^c(\bar{x})$, or equivalently $\bar{x} \in \partial_c u(x)$. However, since $u_1(x) < u_2(x)$, we can again use Lemma 2.5 to see that $\partial_c u(x) = \partial_c u_2(x)$. This contradicts the disjointness of $\partial_c u_1(\Omega)$ and $\partial_c u_2(\Omega)$, thus we must have (7.4). □

8. Stability of tears

Our main goal of this section is to establish a stability result for the multiplicity of singularities of an optimal potential, under certain perturbations of the target measure. To do so, we must first choose an appropriate notion of perturbation for the target measure. In this case, we would only expect stability under perturbations of the target measure that prohibit moving even small amounts of mass to a far away location. Thus a good candidate is the $\mathcal{W}_\infty$ metric defined below.

Definition 8.1 (∞-Kantorovich-Rubinstein-Wasserstein distance)**.** Given two probability measures ν_1 and ν_2 on $\bar{M}$, the $\mathcal{W}_\infty$ distance between them is defined by

$$\mathcal{W}_\infty(\nu_1, \nu_2) := \inf \left\{ \|d_{\bar{g}}\|_{L^\infty(\gamma)} \mid \gamma \in \Pi(\nu_1, \nu_2) \right\}.$$

Here, $d_{\bar{g}}$ is the geodesic distance on $\bar{M}$ induced by the associated Riemannian metric, and $\Pi(\nu_1, \nu_2)$ is the set of probability measures on $\bar{M} \times \bar{M}$ whose left and right marginals are ν_1 and ν_2, repectively

To obtain stability, we again require affine independence (Definition 4.11) of the pieces of the support of the target measure. See Example A.1 for a counterexample to stability when this independence is not present.

We are now ready to state the stability result.

Theorem 8.2 (Stability of tears)**.** *Suppose a cost function $c : \Omega \times \overline{\Omega} \to \mathbf{R}$ satisfies* (B1) *and* (MTW)*, and the measures μ and ν on $\Omega \subset M$ and $\overline{\Omega} \subset \bar{M}$ respectively satisfy conditions (I) and (II) above* (4.6)–(4.7)*. Also let u be an optimal potential transporting μ to ν with cost c and suppose u has multiplicity $k+1 \leq K$ at $x_0 \in (\operatorname{spt} \mu)^{\text{int}}$, relative to a finite collection $\left\{\overline{\Omega}_i\right\}_{i=1}^K$ of disjoint compact sets whose union is $\operatorname{spt} \nu$. Reorder if necessary, so that u also has multiplicity $k+1$ with respect to the subcollection $\left\{ \left[\overline{\Omega}_i\right]_{x_0} \right\}_{i=1}^{k+1}$ consisting of the first $k+1$ sets; assume this subcollection is affinely independent and consists of strictly convex sets.*

Then for any $\varepsilon > 0$, there exists a $\delta > 0$ depending only on $\varepsilon, c,$ $\operatorname{spt} \mu$, and $\{\overline{\Omega}_i\}_{i=1}^K$, such that for any ν^δ with $\mathcal{W}_\infty\left(\nu, \nu^\delta\right) < \delta$ and any optimal potential u^δ transporting μ to ν^δ, there is a DC submanifold of dimension $n-k$ in $B_\varepsilon(x_0) \subset \mathbf{R}^n$ on which u^δ has multiplicity $k+1$ relative to $\left\{\mathcal{N}_\delta\left(\overline{\Omega}_i\right)\right\}_{i=1}^K$ at every point.

The discrepancy of k versus $k+1$ between Theorem 3.4 and Theorem 8.2 arises because the affine hull of $k+1$ affinely independent points generates an affine subspace of dimension k.

We first show a lemma which uses the affine independence assumption to deduce $\dim \partial u(x_0) = k$, so that Theorem 3.4 can be applied. To do so requires some finer properties of the c-subdifferentials of each of the functions u_i which make up u in the decomposition constructed in Lemma 4.4.

Lemma 8.3. *Suppose $\{u_i\}_{i=1}^K$ is the collection of c-convex functions obtained by applying Lemma 4.4 to the optimal potential u under the conditions of Theorem 8.2.*

Ordering indices as in Theorem 8.2, $u_i \in C^1(\Omega)$ *for* $i \leq k+1$ *and*

$$\partial u\,(x_0) \cap \left[\overline{\Omega}_i\right]_{x_0} = \begin{cases} \{Du_i(x_0)\}, & 1 \leq i \leq k+1, \\ \emptyset, & k+1 < i \leq K, \end{cases} \tag{8.1}$$

$$\partial u\,(x_0) = \operatorname{conv}\left(\bigcup_{1 \leq i \leq k+1} \{Du_i(x_0)\} \right), \tag{8.2}$$

$$u(x_0) = u_i(x_0),\ 1 \leq i \leq k+1, \tag{8.3}$$

$$u(x_0) > u_i(x_0),\ k+1 < i \leq K. \tag{8.4}$$

Additionally, $\dim \partial u\,(x_0) = k$.

Proof. Apply Proposition 4.9 to obtain $\{u_i\}_{i=1}^K$. Recall that $[\partial_c u(x_0)]_{x_0} = \partial u\,(x_0)$ under (B1), (MTW), and conditions (I)–(II); thus Proposition 4.9 and the fact that the multiplicity of u at x_0 relative to $\{\overline{\Omega}_i\}_{i=1}^K$ is $k+1$ implies that $\partial u\,(x_0)$ intersects exactly $k+1$ of the sets $\left[\overline{\Omega}_i\right]_{x_0}$, each at exactly one point.

Re-number the indices $1 \leq i \leq K$ so that $\partial u\,(x_0)$ intersects $\left[\overline{\Omega}_i\right]_{x_0}$ only for $1 \leq i \leq k+1$. Since $Du_i(x_0) \in \left[\overline{\Omega}_i\right]_{x_0}$ for each i, Lemma 2.5 along with the mutual disjointness of the $\overline{\Omega}_i$ immediately gives (8.1), (8.2), (8.3), and (8.4).

Finally by (8.2), it is clear that $\dim(\partial u\,(x_0)) \leq k$. However, if $\dim(\partial u\,(x_0)) < k$, the collection $\left\{\left[\overline{\Omega}_i\right]_{x_0}\right\}_{i=1}^{k+1}$ would fail to be affinely independent, thus we must have equality. This finishes the proof. $\square$

We are now in a situation to appeal to Theorem 3.4 and finish the proof of the stability theorem.

Proof of Theorem 8.2. We first apply Lemma 8.3 and reorder indices if necessary to obtain c-convex functions u_i, $1 \leq i \leq K$ with properties (8.1) through (8.4).

Now fix an $\varepsilon > 0$ and suppose by contradiction that the theorem fails to hold: then there exist sequences $\delta_j \searrow 0$ and ν^j with $\mathcal{W}_\infty(\nu, \nu^j) < \delta_j$, and optimal potentials u^j transporting μ to ν^j with cost function c, but u^j does *not* have δ_j-multiplicity $k+1$ at each point of a codimension k, DC submanifold of $B_\varepsilon(x_0)$. Since $c \in C^4(\Omega \times \overline{\Omega})$ and each u^j is c-convex, the collection $\{u^j\}_{j=1}^\infty$ is uniformly Lipschitz. Then by Arzelà-Ascoli (after adding constants to each u^j, which does not change the δ_j-multiplicity of any points) we can extract a subsequence, still indexed by j, that converges uniformly. By stability of optimal transport maps (see for example, [37, Corollary 5.23]) and convexity of $\operatorname{spt}\mu$ this limit must be (again, up to adding a constant) equal to u.

Now by taking j large enough we may ensure the sets $\mathcal{N}_{\delta_j}(\overline{\Omega}_i)$ are mutually disjoint for each j; note that by the definition of $\mathcal{W}_\infty$, the assumption $\mathcal{W}_\infty(\nu, \nu^j) < \delta_j$ implies $\operatorname{spt}\nu^j \subset \bigcup_{i=1}^K \mathcal{N}_{\delta_j}(\overline{\Omega}_i)$. Thus, as in Lemma 4.4 we obtain

$$u_i^j(x) := \sup_{\bar{x} \in \mathcal{N}_{\delta_j}(\overline{\Omega}_i)} (-c(x, \bar{x}) - (u^j)^c(\bar{x})),$$

$$u^j(x) = \max_{1 \leq i \leq K} u_i^j(x),$$

for $x \in \operatorname{spt}\mu$ as long as j is large enough. We also comment here that u_i^j converges uniformly to u_i for each $1 \leq i \leq K$, while the compactness of each set $\mathcal{N}_{\delta_j}(\overline{\Omega}_i)$ implies that

$$\partial_c u_i^j(x) \cap \mathcal{N}_{\delta_j}(\overline{\Omega}_i) \neq \emptyset, \quad \forall\, x \in \Omega. \tag{8.5}$$

We can now take a local coordinate system near x_0 to view all functions as defined in a subset of $\mathbf{R}^n$, since all u_i^j and u_i are c-convex they have uniformly bounded constant of semi-convexity near x_0 (see [17, Proposition C.2]). Thus u and $\{u^j\}_{j=1}^{\infty}$ satisfy the conditions of Theorem 3.4, and for j sufficiently large, we obtain existence of a DC submanifold $\Sigma_{n-k}^j \subset B_\varepsilon(x_0)$ of codimension k satisfying $\dim \partial u^j(x) \geq k$ for every $x \in \Sigma_{n-k}^j$.

At this point, fix any $x \in \Sigma_{n-k}^j$. By (3.2) and Lemma 2.5 we see that

$$\begin{aligned}\partial_c u^j(x) &= exp_x^c(\partial u^j(x))\\ &= exp_x^c(\operatorname{conv}\left(\bigcup_{1\leq i\leq k+1} \partial u_i^j(x)\right)),\end{aligned}$$

thus (8.5) implies that for j large enough u^j has δ_j-multiplicity at least $k+1$ at x. On the other hand by the mutual disjointness of $\{\overline{\Omega}_i\}_{i=1}^{K}$ and recalling $\partial u_i^j(x) = \left[\partial_c u_i^j(x)\right]_x$, Lemma 3.5 yields that for j large enough, $1 \leq i \leq k+1$, and $i \neq i' \leq K$, we have $\partial_c u_i^j(x) \cap \mathcal{N}_{\delta_j}(\overline{\Omega}_{i'}) = \emptyset$; in particular this implies u^j has δ_j-multiplicity no more than $k+1$ at x. Thus if j is large enough, u^j has δ_j-multiplicity exactly $k+1$ at every point in Σ_{n-k}^j, which finishes the proof by contradiction. □

Appendix A. Failure of stability without affine independence

In this appendix, we provide an example to illustrate the importance of the affine independence condition on the support of the target measure in Theorem 8.2. Note simply by definition, no collection of $n+2$ or more sets can be affinely independent in $\mathbf{R}^n$. The example we illustrate below has a target measure on $\mathbf{R}^2$ whose support consists of four strictly convex sets, and the associated optimal potential has a point of multiplicity 4 which is unstable under certain $\mathcal{W}_\infty$ perturbations. The source measure will have constant density, and the target measure will be absolutely continuous with density bounded from above. This density does not have a lower bound away from zero in its whole support, so it does not exactly satisfy all of the remaining (i.e. other than affine independence) hypotheses of Theorem 8.2, but we comment that the resulting optimal potential is an envelope of globally C^1 functions, which is the only way in which these other conditions are required in the proof of this theorem. In particular, this example strongly suggests that to obtain stability there must be some restriction on the multiplicity in relation to the ambient dimension.

Proposition A.1. *Let $c(x, \bar{x}) = -\langle x, \bar{x}\rangle$ on $\mathbf{R}^2 \times \mathbf{R}^2$. Denoting points $(x, y) \in \mathbf{R}^2$, let*

$$D := \left\{(x, y) \in \mathbf{R}^2 \mid x^2 - r_0^2 \leq y \leq r_0^2 - x^2\right\}$$

where $r_0 > 0$ is a small constant to be determined, and take μ to be the uniform probability measure on D (see Figure 1). Also define the function

$$u = \max_{1\leq i\leq 4} u_i$$

where

$$\begin{aligned} u_1(x,y) &= x^2+y^2-x^6+y, \\ u_2(x,y) &= 4x^2+y^2-y^6+x-3xy, \\ u_3(x,y) &= 4x^2+y^2-y^6-x+3xy, \\ u_4(x,y) &= 4y^4+y^2-|x|^3+y^2\max\{0,-\operatorname{sgn}(y)\}+3|x|^{\frac{3}{2}}, \end{aligned}$$

and take ν to be the pushforward of μ under Du. Then ν is absolutely continuous with density bounded away from infinity on its support, $\operatorname{spt}\nu$ *is the disjoint union of nonempty, compact, strictly convex sets $\{\overline{\Omega}_1,\dots,\overline{\Omega}_4\}$, each $u_i \in C^1(\mathbf{R}^n)$, and u has a singularity of multiplicity 4 at $(0,0)$ relative to this collection. Moreover, for any $\delta > 0$ there exists a sequence of measures ν^j converging to ν in $\mathcal{W}_\infty$ for which the associated optimal potentials mapping μ to ν^j do not have any singularities of δ-multiplicity 4 relative to $\{\overline{\Omega}_1,\dots,\overline{\Omega}_4\}$.*

Proof. First, we mention the choice of r_1 is taken so that the line $y=-r_1$ passes through the intersection of the curves $y = x^2 - r_0^2$ and $y = -|x|$. Thus it is easy to see that D is convex.

Second, we note that $u_1,\dots u_4$ are convex on D if r_0 is sufficiently small. Indeed, since we are in two dimensions, the characteristic polynomial of the Hessian matrix of a C^2 function f is $\lambda^2 - \Delta f\lambda + \det D^2 f$. Thus by the quadratic formula, if $\Delta f \geq 0$ and $\det D^2 f \geq 0$ both eigenvalues will be nonnegative, hence f will be convex. This immediately gives the convexity of u_1, u_2, u_3 by a quick calculation near the origin. For u_4, we can first see $3|x|^{\frac{3}{2}} - |x|^3$ is a convex function of one variable in $\mathbf{R}$ near zero (by calculating the subdifferential of the function), hence also as a function on $\mathbf{R}^2$. Then the remaining terms are also clearly convex, thus so is their sum u_4.

Next, if we let $U_i := \{u = u_i\} \cap D$, some tedious but routine calculations yield that

$$\begin{aligned} U_1 &= \{(x,y)\in D \mid y \geq |x|\}, \\ U_2 &= \left\{(x,y)\in D \mid x\geq 0,\ -\sqrt{|x|} \leq y \leq |x|\right\}, \\ U_3 &= \left\{(x,y)\in D \mid x\leq 0,\ -\sqrt{|x|} \leq y \leq |x|\right\}, \\ U_4 &= \left\{(x,y)\in D \mid y \leq -\sqrt{|x|}\right\}, \end{aligned}$$

for r small enough. We will show that $Du_i(U_i)$ is a strictly convex set for each i.

Before embarking on this verification, let us record

$$\begin{aligned} \nabla u_1(x,y) &= (2x-6x^5, 2y+1), \\ \nabla u_2(x,y) &= (8x+1-3y, 2y-6y^5-3x), \\ \nabla u_3(x,y) &= (8x-1-3y, 2y-6y^5+3x), \\ \nabla u_4(x,y) &= (\operatorname{sgn}(x)(\frac{9}{2}|x|^{\frac{1}{2}}-3x^2), 16y^3+2(1+max\{0,-\operatorname{sgn}(y)\})y). \end{aligned}$$

The idea will be to take a portion of U_i^∂ and write it parametrically as $\gamma(t)$. Then we can write

$$(f(t), g(t)) := \nabla u_i(\gamma(t)),$$

and consider one of either

$$y(x) := g(f^{-1}(x)), \quad x(y) := f(g^{-1}(y)),$$

(i.e., we write one of the coordinates as a function of the other, and consider the image of the boundary curve as the graph of this function). By determining the strict convexity or concavity of these functions (depending on which variable we have solved for, and which side the image of U_i lies), we can then conclude strict convexity of $\nabla u_i(U_i)$ (see Figure 1 below for a rough sketch of these regions, diagram is not to scale). We will either directly solve for a variable and verify convexity / concavity, or use the formulae

$$(f^{-1})'(x) = \frac{1}{f'(f^{-1}(x))}, \quad (f^{-1})''(x) = -\frac{f''(f^{-1}(x))}{(f'(f^{-1}(x)))^3}$$

to see

$$\begin{aligned} y''(x) &= g''(f^{-1}(x))((f^{-1})'(x))^2 + g'(f^{-1}(x))(f^{-1})''(x) \\ &= ((f^{-1})'(x))^2(g''(f^{-1}(x)) - g'(f^{-1}(x))\frac{f''(f^{-1}(x))}{f'(f^{-1}(x))}). \end{aligned} \tag{A.1}$$

(or their analogues if considering $x(y)$).

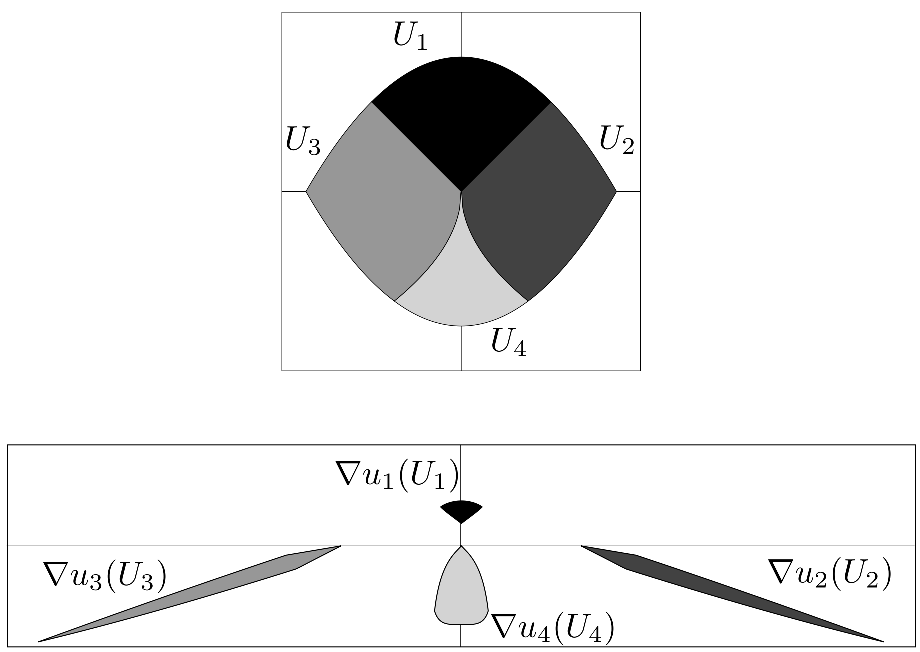

Figure 1.

The remainder is a series of calculations, below r_1 and r_2 are the x-coordinates of the intersection of the line $y = x$ with the curve $y = r_0^2 - x^2$, and the intersection of the curves $y = -|x|^{\frac{1}{2}}$ and $y = x^2 - r_0^2$ respectively, note that r_1, $r_2 < r_0$, hence we will always be in the situation $|t| \leq r_0$ in the calculations below.

$$\boldsymbol{\nabla u_1(U_1^{\partial} \cap \{y = r_0^2 - x^2\}),\ \gamma(t) := (t, r_0^2 - t^2),}$$
$$\boldsymbol{-r_1 \leq t \leq r_1,\ \text{region below curve}.}$$

Then

$$\begin{aligned}(f(t), g(t)) &= (2t - 6t^5, 2(r^2 - t^2) + 1),\\ (f'(t), g'(t)) &= (2 - 30t^4, -4t),\\ (f''(t), g''(t)) &= (-120t^3, -4)\end{aligned}$$

so for $t = f^{-1}(x)$,

$$\frac{y''(x)}{((f^{-1})'(x))^2} = -4 - \frac{480t^4}{2 - 30t^4} < 0,$$

if r_0 is small so y is strictly concave.

$$\boldsymbol{\nabla u_1(U_1^{\partial} \cap \{y = x\}),\ \gamma(t) := (t, t),}$$
$$\boldsymbol{0 \leq t \leq r_1,\ \text{region above curve}.}$$

Then

$$(f(t), g(t)) = (2t - 6t^5, 2t + 1),$$

so directly solving:

$$x(y) = f(g^{-1}(y)) = f(\frac{y-1}{2}) = y - 1 - \frac{6}{2^5}(y-1)^5$$

which is strictly concave. We make a similar calculation for $U_1^{\partial} \cap \{y = -x\}$, then since ∇u_1 maps vertical line segments to vertical line segments with the same orientation and the first coordinate is strictly increasing as long as r_0 is small, this shows that $\nabla u_1(U_1)$ is a strictly convex set.

$$\boldsymbol{\nabla u_2(U_2^{\partial} \cap \{y = x\}),\ \gamma(t) := (t, t),}$$
$$\boldsymbol{0 \leq t \leq r_1,\ \text{region below curve}.}$$

Then

$$(f(t), g(t)) = (8t + 1 - 3t, 2t - 6t^5 - 3t) = (5t + 1, -t - 6t^5),$$

directly solving,

$$y(x) := g(f^{-1}(x)) = g(\frac{x-1}{5}) = \frac{1-x}{5} - \frac{6(x-1)^5}{5^5}$$

which is strictly concave.

$$\boldsymbol{\nabla u_2(U_2^\partial \cap \{y = -\sqrt{|x|}\}),\ \gamma(t) := (t, -t^{\frac{1}{2}}),}$$
$$\boldsymbol{0 \le t \le r_2,\ \text{region above curve}.}$$

Then

$$\begin{aligned}
(f(t), g(t)) &= (8t + 1 + 3t^{\frac{1}{2}}, -2t^{\frac{1}{2}} + 6t^{\frac{5}{2}} - 3t),\\
(f'(t), g'(t)) &= (8 + \frac{3}{2}t^{-\frac{1}{2}}, -t^{-\frac{1}{2}} + 15t^{\frac{3}{2}} - 3),\\
(f''(t), g''(t)) &= (-\frac{3}{4}t^{-\frac{3}{2}}, \frac{t^{-\frac{3}{2}}}{2} + \frac{45}{2}t^{\frac{1}{2}}),
\end{aligned}$$

for $t = f^{-1}(x)$,

$$\begin{aligned}
\frac{y''(x)}{((f^{-1})'(x))^2} &= \frac{t^{-\frac{3}{2}}}{2} + \frac{45}{2}t^{\frac{1}{2}} + \frac{\frac{3}{4}t^{-\frac{3}{2}}(-t^{-\frac{1}{2}} + 15t^{\frac{3}{2}} - 3)}{8 + \frac{3}{2}t^{-\frac{1}{2}}}\\
&\ge t^{-\frac{3}{2}}\left(\frac{1}{2} - \frac{3(t^{-\frac{1}{2}} + 3)}{4(8 + \frac{3}{2}t^{-\frac{1}{2}})}\right)\\
&\ge t^{-\frac{3}{2}}\left(\frac{1}{2} - \frac{3(1 + 3\sqrt{t})}{4(8\sqrt{t} + \frac{3}{2})}\right) \ge t^{-\frac{3}{2}}\left(\frac{1}{2} - \frac{3(1 + 3\sqrt{r_2})}{6}\right) > 0
\end{aligned}$$

if r_0 is small enough, making y strictly convex.

$$\boldsymbol{\nabla u_2(U_2^\partial \cap \{y = r_0^2 - x^2\}),\ \gamma(t) := (t, r_0^2 - t^2),}$$
$$\boldsymbol{r_1 \le t \le r_0,\ \text{region below curve}.}$$

$$\begin{aligned}
(f(t), g(t)) &= (8t + 1 - 3(r_0^2 - t^2), 2(r_0^2 - t^2) - 6(r_0^2 - t^2)^5 - 3t),\\
(f'(t), g'(t)) &= (8 + 6t, -4t + 60t(r_0^2 - t^2)^4 - 3),\\
(f''(t), g''(t)) &= (6, -4 + 60(r_0^2 - t^2)^4 - 480t^2(r_0^2 - t^2)^3),
\end{aligned}$$

for $t = f^{-1}(x)$,

$$\begin{aligned}
\frac{y''(x)}{((f^{-1})'(x))^2} &= -4 + 60(r_0^2 - t^2)^4 - 480t^2(r_0^2 - t^2)^3 + \frac{6(4t - 60t(r_0^2 - t^2)^4 + 3)}{8 + 6t}\\
&\le -4 + 60r_0^8 + \frac{3(4t + 3)}{4} \le -4 + 60r_0^8 + \frac{9}{4} + 3r_0 < 0
\end{aligned}$$

when r_0 is small, so y is strictly concave.

$$\boldsymbol{\nabla u_2(U_2^{\partial} \cap \{y = x^2 - r_0^2\}),\ \gamma(t) := (t, t^2 - r_0^2),}$$
$$\boldsymbol{r_2 \le t \le r_0,\ \text{region above curve}.}$$

$$\begin{aligned}
(f(t), g(t)) &= (8t + 1 - 3(t^2 - r_0^2), 2(t^2 - r_0^2) - 6(t^2 - r_0^2)^5 - 3t),\\
(f'(t), g'(t)) &= (8 - 6t, 4t - 60t(t^2 - r_0^2)^4 - 3),\\
(f''(t), g''(t)) &= (-6, 4 - 60(t^2 - r_0^2)^4 - 480t^2(t^2 - r_0^2)^3),
\end{aligned}$$

for $t = f^{-1}(x)$,

$$\begin{aligned}
\frac{y''(x)}{((f^{-1})'(x))^2} &= 4 - 60(t^2 - r_0^2)^4 - 480t^2(t^2 - r_0^2)^3 + \frac{6(4t - 60t(t^2 - r_0^2)^4 - 3)}{8 - 6t}\\
&\ge 4 - 60r_0^8 - \frac{6(60t(r_0^2 - t^2)^4 + 3)}{8 - 6t} \ge 4 - 60r_0^8 - \frac{6(60r_0^9 + 3)}{8 - 6r_0} > 0
\end{aligned}$$

when r_0 is small, so y is strictly convex.

Since ∇u_2 maps all horizontal lines to lines with the same slope, the above verifications give strict convexity of $\nabla u_2(U_2)$, a symmetric argument shows the strict convexity of $\nabla u_3(U_3)$.

$$\boldsymbol{\nabla u_4(U_4^{\partial} \cap \{y = -\sqrt{|x|}\}),\ \gamma(t) := (t, -t^{\frac{1}{2}}),}$$
$$\boldsymbol{0 \le t \le r_2,\ \text{region below curve}.}$$

$$\begin{aligned}
(f(t), g(t)) &= (\frac{9}{2}t^{\frac{1}{2}} - 3t^2, -16t^{\frac{3}{2}} - 3t^{\frac{1}{2}}),\\
(f'(t), g'(t)) &= (\frac{9}{4}t^{-\frac{1}{2}} - 6t, -24t^{\frac{1}{2}} - \frac{3}{2}t^{-\frac{1}{2}}),\\
(f''(t), g''(t)) &= (-\frac{9}{8}t^{-\frac{3}{2}} - 6, -12t^{-\frac{1}{2}} + \frac{3}{4}t^{-\frac{3}{2}}),
\end{aligned}$$

for $t = f^{-1}(x)$,

$$\begin{aligned}
\frac{y''(x)}{((f^{-1})'(x))^2} &= -12t^{-\frac{1}{2}} + \frac{3}{4}t^{-\frac{3}{2}} - \frac{(\frac{9}{8}t^{-\frac{3}{2}} + 6)(24t^{\frac{1}{2}} + \frac{3}{2}t^{-\frac{1}{2}})}{\frac{9}{4}t^{-\frac{1}{2}} - 6t}\\
&< \frac{3}{4}t^{-\frac{3}{2}} - \frac{(\frac{9}{8}t^{-\frac{3}{2}})(\frac{3}{2}t^{-\frac{1}{2}})}{\frac{9}{4}t^{-\frac{1}{2}}} \le t^{-\frac{3}{2}}\left(\frac{3}{4} - \frac{3}{4}\right) = 0
\end{aligned}$$

when r_0 is small, so y is strictly concave. A symmetric calculation holds for the boundary curve where $x \le 0$.

$$\nabla u_4(U_4^\partial \cap \{y = x^2 - r_0^2\}),\ \gamma(t) := (t, x^2 - r_0^2),$$
$$-r_2 \leq t \leq r_2,\ \textbf{region above curve}.$$

$$(f(t), g(t)) = (\frac{9}{2}t^{\frac{1}{2}} - 3t^2, 16(t^2 - r_0^2)^3 + 3(t^2 - r_0^2)),$$
$$(f'(t), g'(t)) = (\frac{9}{4}t^{-\frac{1}{2}} - 6t, 96t(t^2 - r_0^2)^2 + 6t),$$
$$(f''(t), g''(t)) = (-\frac{9}{8}t^{-\frac{3}{2}} - 6, 96(t^2 - r_0^2)^2 + 384t^2(t^2 - r_0^2) + 6),$$

this time for $t = g^{-1}(y)$ and $t > 0$,

$$\begin{aligned} x''(y) &= f''(g^{-1}(y)) - f'(g^{-1}(y))\frac{g''(g^{-1}(y))}{g'(g^{-1}(y))} \\ &= -\frac{9}{8}t^{-\frac{3}{2}} - 6 - \frac{(\frac{9}{4}t^{-\frac{1}{2}} - 6t)(96(t^2 - r_0^2)^2 + 384t^2(t^2 - r_0^2) + 6)}{96t(t^2 - r_0^2)^2 + 6t} \\ &< 0 \end{aligned}$$

when r_0 is small, so x is a strictly concave function of y when $x > 0$. A symmetric argument holds when $x < 0$. When $x = 0$, we find the tangent line to $\nabla u_4(U_4)$ at the boundary point $(0, -16r_0^6 - 3r_0^2)$ is the horizontal line through that point, which is easily seen to lie below $\nabla u_4(U_4)$, touching only at $(0, -16r_0^6 - 3r_0^2)$, with a similar argument for the tangent line at $(0, 0)$. Since ∇u_4 sends vertical lines to vertical lines with the same orientation, this shows $\nabla u_4(U_4)$ is strictly convex, completing the verification.

Finally, we easily see that u_i is strictly convex for each i, and the above calculation of regions shows ∇u is injective on the union of the interiors of the U_i. A quick calculation shows $\det D^2 u_i$ is actually bounded away from zero on U_i for each i, this gives that ν is absolutely continuous with density bounded away from infinity (in fact, this density is actually bounded away from zero on the images of U_1, U_2, and U_3).

Now we can see that $\nabla u_2(U_2)$, $\nabla u_3(U_3)$, and $\nabla u_4(U_4)$ all lie in the half space $\{(x, y) \in \mathbf{R}^2 \mid y \leq 0\}$ and all have nonempty intersections with the x-axis. Fix any $\delta > 0$, we now take the sequence of measures $\{\nu^j\}_{j=1}^\infty$ to be ν, but with the set $\nabla u_4(U_4)$ shifted upward by δ/j, it is clear that $\mathcal{W}_\infty(\nu, \nu^j) \leq \delta/j$. Let u^j be an optimal potential transporting μ to ν^j, and suppose there is a point x that is a singularity of δ-multiplicity 4 for u^j relative to $\{\nabla u_1(U_1), \ldots \nabla u_4(U_4)\}$. Since the extremal points of $\partial u\,(x)$ must be contained in $\operatorname{spt} \nu^j$, this could only happen if $\partial u\,(x)$ intersects both $\nabla u_2(U_2)$ and $\nabla u_3(U_3)$. However this would also force $\partial u\,(x)$ to have nonempty intersection with the interior of $\nabla u_4(U_4) + \frac{\delta}{j}e_2$ and we can derive a contradiction by the same argument as in the proof of Proposition 4.9, thus no point can have a δ-multiplicity of 4. □

References

[1] Giovanni Alberti. On the structure of singular sets of convex functions. *Calc. Var. Partial Differential Equations*, 2(1):17–27, 1994.

[2] L. Ambrosio and J. Bertrand. DC Calculus. *ArXiv e-prints*, May 2015.
[3] Yann Brenier. Polar factorization and monotone rearrangement of vector-valued functions. *Comm. Pure Appl. Math.*, 44(4):375–417, 1991.
[4] G. J. Butler, J. G. Timourian, and C. Viger. The rank theorem for locally Lipschitz continuous functions. *Canad. Math. Bull.*, 31(2):217–226, 1988.
[5] Luis A. Caffarelli. The regularity of mappings with a convex potential. *J. Amer. Math. Soc.*, 5(1):99–104, 1992.
[6] Luis A. Caffarelli and Robert J. McCann. Free boundaries in optimal transport and Monge-Ampère obstacle problems. *Ann. of Math. (2)*, 171(2):673–730, 2010.
[7] Shibing Chen. Regularity of free boundary in optimal transportation. *Preprint*.
[8] Shibing Chen and Xu-Jia Wang. Strict convexity and $C^{1,\alpha}$ regularity of potential functions in optimal transportation under condition A3w. *J. Differential Equations*, 260(2):1954–1974, 2016.
[9] Otis Chodosh, Vishesh Jain, Michael Lindsey, Lyuboslav Panchev, and Yanir A. Rubinstein. On discontinuity of planar optimal transport maps. *J. Topol. Anal.*, 7(2):239–260, 2015.
[10] F. H. Clarke. On the inverse function theorem. *Pacific J. Math.*, 64(1):97–102, 1976.
[11] F. H. Clarke. *Optimization and nonsmooth analysis*, volume 5 of *Classics in Applied Mathematics*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, second edition, 1990.
[12] Guido De Philippis and Alessio Figalli. Partial regularity for optimal transport maps. *Publ. Math. Inst. Hautes Études Sci.*, 121:81–112, 2015.
[13] Alessio Figalli. Regularity properties of optimal maps between nonconvex domains in the plane. *Comm. Partial Differential Equations*, 35(3):465–479, 2010.
[14] Alessio Figalli and Young-Heon Kim. Partial regularity of Brenier solutions of the Monge-Ampère equation. *Discrete Contin. Dyn. Syst.*, 28(2):559–565, 2010.
[15] Alessio Figalli, Young-Heon Kim, and Robert J. McCann. Hölder continuity and injectivity of optimal maps. *Arch. Ration. Mech. Anal.*, 209(3):747–795, 2013.
[16] Wilfrid Gangbo. *Quelques problemes d'analyse nonconvex*. PhD thesis, Universite de Metz, 1995. Habilitation thesis.
[17] Wilfrid Gangbo and Robert J. McCann. The geometry of optimal transportation. *Acta Math.*, 177(2):113–161, 1996.
[18] Wilfrid Gangbo and Robert J. McCann. Shape recognition via Wasserstein distance. *Quart. Appl. Math.*, 58(4):705–737, 2000.
[19] Nicola Gigli. On the inverse implication of Brenier-McCann theorems and the structure of $(\mathcal{P}_2(M), W_2)$. *Methods Appl. Anal.*, 18(2):127–158, 2011.
[20] Michael Goldman and Felix Otto. A variational proof for partial regularity of optimal transportation maps. *Preprint at arXiv:1704.05339v1*.
[21] Nestor Guillen and Jun Kitagawa. On the local geometry of maps with c-convex potentials. *Calc. Var. Partial Differential Equations*, 52(1-2):345–387, 2015.
[22] Huai-Yu Jian and Xu-Jia Wang. Continuity estimates for the Monge-Ampère equation. *SIAM J. Math. Anal.*, 39(2):608–626, 2007.
[23] Young-Heon Kim and Robert J. McCann. Continuity, curvature, and the general covariance of optimal transportation. *J. Eur. Math. Soc. (JEMS)*, 12(4):1009–1040, 2010.
[24] Vladimir Levin. Abstract cyclical monotonicity and Monge solutions for the general Monge-Kantorovich problem. *Set-Valued Anal.*, 7(1):7–32, 1999.
[25] Jiakun Liu, Neil S. Trudinger, and Xu-Jia Wang. Interior $C^{2,\alpha}$ regularity for potential functions in optimal transportation. *Comm. Partial Differential Equations*, 35(1):165–184, 2010.
[26] Grégoire Loeper. On the regularity of solutions of optimal transportation problems. *Acta Math.*, 202(2):241–283, 2009.
[27] Xi-Nan Ma, Neil S. Trudinger, and Xu-Jia Wang. Regularity of potential functions of the optimal transportation problem. *Arch. Ration. Mech. Anal.*, 177(2):151–183, 2005.
[28] Robert J. McCann. Existence and uniqueness of monotone measure-preserving maps. *Duke Math. J.*, 80(2):309–323, 1995.
[29] Robert J. McCann. Polar factorization of maps on Riemannian manifolds. *Geom. Funct. Anal.*, 11(3):589–608, 2001.
[30] Grisha Perelman. DC structure on Alexandrov space (preliminary version). http://www.math.psu.edu/petrunin/papers/Cstructure.pdf.
[31] R. Tyrrell Rockafellar. *Convex analysis*. Princeton University Press, 1970.

[32] Neil S. Trudinger. Recent developments in elliptic partial differential equations of Monge-Ampère type. In *International Congress of Mathematicians. Vol. III*, pages 291–301. Eur. Math. Soc., Zürich, 2006.

[33] Neil S. Trudinger and Xu-Jia Wang. On the second boundary value problem for Monge-Ampère type equations and optimal transportation. *Ann. Sc. Norm. Super. Pisa Cl. Sci. (5)*, 8(1):143–174, 2009.

[34] L. Veselý and L. Zajíček. Delta-convex mappings between Banach spaces and applications. *Dissertationes Math. (Rozprawy Mat.)*, 289:52, 1989.

[35] Jérôme Vétois. Continuity and injectivity of optimal maps. *Calc. Var. Partial Differential Equations*, 52(3-4):587–607, 2015.

[36] Cédric Villani. *Topics in optimal transportation*, volume 58 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2003.

[37] Cédric Villani. *Optimal transport: Old and new*, volume 338 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, 2009.

[38] Xu-Jia Wang. Remarks on the regularity of Monge-Ampère. pages 257–263. Academic Press, Beijing, 1992.

[39] J. Warga. An implicit function theorem without differentiability. *Proc. Amer. Math. Soc.*, 69(1):65–69, 1978.

[40] Luděk Zajíček. On the differentiation of convex functions in finite and infinite dimensional spaces. *Czechoslovak Math. J.*, 29(104)(3):340–348, 1979.

Department of Mathematics, Michigan State University, 619 Red Cedar Road, East Lansing, MI 48824

E-mail address: `kitagawa@math.msu.edu`

Department of Mathematics, University of Toronto, Toronto, Ontario, Canada, M5S 2E4

E-mail address: `mccann@math.toronto.edu`