

J. H. MARTEL

ABSTRACT.

## CONTENTS

1. Introduction	1
2. Cost Assumptions	2
3. Optimal Semicouplings and Monge-Kantorovich Duality	4
4. (Twist) Condition	6
5. Statement of Main Results	7
6. Averaged Gradients	10
7. Finite-Time Blow-Up	12
8. Continuity of $\omega$	13
9. Proof of Theorem 1	14
10. Kantorovich's Contravariant Singularity Functor	15
11. Local Topology and Local Dimensions of $Z$	16
12. The Descending Filtration $\{Z_j\}_{j \geq 0}$	19
13. Proof of Theorem 2	22
14. Brief Review of Singularity in the Literature	26
References	27

## 1. INTRODUCTION

Optimal transportation is a subject with a remarkably wide range of applications [Vil09], [Vil03]. In [Mar] some new applications of optimal transport to algebraic topology are proposed. The idea is that the singularities  $Z$  of  $c$ -optimal semicouplings define contravariant topological functors  $Z : 2^Y \rightarrow 2^X$ , where  $Z(Y_I) := \cap_{y \in Y_I} \partial^c \psi(y)$  for  $\psi^{cc} = \psi$  a  $c$ -concave potential on  $Y$ . The functor  $Z = Z(c, \sigma, \tau) = Z(\psi)$  depends on the optimal transport datum  $c, \sigma, \tau$ , and specifically solutions  $\psi : Y \rightarrow \mathbb{R} \cup \{-\infty\}$  to Kantorovich's dual max program. We study the topology of the functor  $Z$ , which

---

*Date:* January 26, 2022.

exists in general conditions whenever the source  $\sigma$  is *abundant* with respect to the target  $\tau$ :

$$(1) \quad \int_X \sigma > \int_Y \tau.$$

The results of this article demonstrate that homotopy-isomorphisms and strong deformation retracts can be constructed between various inclusions  $Z(Y_I) \leftarrow Z(Y_J)$  arising from contravariance and inclusions  $Y_I \hookrightarrow Y_J$ . Following an idea of Dror Bar-Natan [Bar02] the functor  $Z$  leads to a descending filtration of  $X$  into subvarieties  $Z_j$ ,  $j = 0, 1, 2, \dots$ , namely

$$(2) \quad (X = Z_0) \leftarrow (A = Z_1) \leftarrow Z_2 \leftarrow Z_3 \leftarrow \dots$$

Our first theorem 1 identifies local conditions for which the first inclusion  $(X = Z_0) \leftarrow (A = Z_1)$  is a homotopy-isomorphism, even a strong deformation retract. Our second theorem 2 identifies local conditions and maximal index  $J \geq 0$  for which the right-hand inclusions  $Z_j \leftarrow Z_{j+1}$  are homotopy-isomorphisms, and again even strong deformation retracts. The results of 1, 2, 4, 5 identifies a condition, which we call Uniform Halfspace (UHS) conditions 15, and indices  $J \geq 0$  for which the source space  $X$  can be continuously reduced via strong deformation retracts to codimension- $J$  subvarieties  $X \rightsquigarrow Z_{J+1}$ . This leads to the possibility of constructing spines and souls as the singularities of  $c$ -optimal semicouplings, a topological application which we cannot elaborate here.

## 2. COST ASSUMPTIONS

The topology of the locus-of-discontinuities and the singularity  $Z$  of  $c$ -optimal semicouplings is determined by the geometry of the cost  $c$ , and the geometry of  $c$ -concave potentials. The cost  $c = c(x, y)$  represents the cost of transporting a unit mass at source  $x$  to target at  $y$ . The proofs of our Theorems 1, 2, 4, 5 require cost functions  $c : X \times Y \rightarrow \mathbb{R} \cup \{+\infty\}$  satisfying important assumptions labelled (A0), ..., (A6). Abbreviate  $c_y(x) := c(x, y)$ . The assumptions are the following:

- (A0) The cost  $c : X \times Y \rightarrow \mathbb{R} \cup \{+\infty\}$  is continuous throughout  $\text{dom}(c) \subset X \times Y$  and uniformly bounded from below, e.g.  $c \geq 0$ . Moreover we assume the sublevels  $\{x \in X \mid c(x, y) \leq t\}$  are compact subsets of  $X$  for every  $t \in \mathbb{R}$ ,  $y \in Y$ . Thus we assume  $x \mapsto c(x, y)$  is *coercive* for every  $y \in Y$ .
- (A1) The cost is twice continuously differentiable with respect to the source variable  $x$ , uniformly in  $y$  throughout  $\text{dom}(c)$ . So for every  $y \in Y$ , the Hessian function  $x \mapsto \nabla_{xx}^2 c(x, y)$  exists and is continuous throughout  $\text{dom}(c_y)$ .
- (A2) The function  $(x, y) \mapsto \|\nabla_x c(x, y)\|$  is upper semicontinuous throughout  $\text{dom}(c)$ . So for every  $t \in \mathbb{R}$  the superlevel set  $\{\|\nabla_x c(x, y)\| \geq t\}$  is a closed subset of  $\text{dom}(c)$ .

- (A3) For every  $y \in Y$ , we assume  $x' \mapsto \nabla_x c(x', y)$  does not vanish identically on any open subset of  $\text{dom}(c_y)$ .
- (A4) The cost satisfies (Twist) condition with respect to the source variable throughout  $\text{dom}(c)$ . So for every  $x'$  the rule  $y \mapsto \nabla_x c(x', y)$  defines an injective mapping  $\text{dom}(c_{x'}) \rightarrow T_{x'} X$ .
- (A5) For every  $x \in X$ , the function  $y \mapsto c(x, y)$  is continuously differentiable; and for every  $y \in Y$ , the gradients  $\nabla_y c(x, y)$  are bounded on compact subsets  $K \subset X$ .

Consequences of Assumptions (A0), (A1), ... will be clarified as our thesis progresses. In the simplest case where the source and target spaces  $X, Y$  are compact, and  $c$  is smooth and finite-valued throughout  $X \times Y$ , then Assumptions (A0)–(A2) are readily confirmed. Assumption (A3) forbids the cost  $c(x, y)$  from being locally constant on any open subset of  $X$ .

The Assumption (A4) implies the gradients of the cross-differences  $\nabla_x c_\Delta(x, y, y')$  are nonzero for distinct  $y, y'$ , where  $c_\Delta(x; y, y') := c(x, y) - c(x, y')$  is the two-pointed cross difference. In practice, costs  $c$  which have poles, e.g.  $c(x, y) = +\infty$  when  $x = y$ , will more readily satisfy (A4). Indeed the functions  $x \mapsto c_\Delta(x; y, y')$  have critical points if  $X$  is closed compact space and  $c_\Delta$  is everywhere finite.

We will prove (A0)–(A4) implies the general uniqueness of  $c$ -optimal semicouplings. Peculiar to the semicoupling setting is (A4), which requires injectivity of mappings  $y \mapsto \nabla_x c(x', y) : Y \rightarrow T_{x'} X$  for every  $x \in X$ , where  $\dim(Y) < \dim(X)$ . The Assumption (A5) is useful in the construction of our deformation retracts in Theorems 4–5 below.

A further assumption (A6) is necessary for the deformation retracts constructed in the present article. These retracts depend on the nonvanishing of averaged vector fields denoted  $\eta_{avg}(x) \in T_x X$ . The field  $\eta_{avg}$  is an average of a  $Y$ -parameter family of potentials  $\eta(x, y)$  defined with respect to  $c$ -concave potentials  $\psi^{cc} = \psi$ , and parameters  $\beta > 0$ , by formulas

$$(3) \quad \eta(x, y) := |\psi(y_0) - \psi(y) - c_\Delta(x, y_0, y)|^{-\beta} \nabla_x (c(x, y_0) - c(x, y))$$

for  $y_0 \in \partial^c \phi(x)$ ,  $y \in Y$ . The formal definition of  $\eta_{avg}(x)$  depends on the setting. Typically there is a Radon measure  $\bar{\nu}_x$  on  $Y$ , depending on  $x$ , absolutely continuous with respect to  $\mathcal{H}_Y$ , and with average

$$(4) \quad \eta_{avg}(x) := (\bar{\nu}_x[Y])^{-1} \int_Y \eta(x, y) d\bar{\nu}_x(y).$$

With these definitions of  $\eta_{avg}$  we can now state the assumption (A6):

- (A6) The averaged vectors  $\eta_{avg}(x)$  are bounded away from zero, uniformly with respect to  $x \in Z'$  on the relevant subsets  $Z'$  of  $X$ .

See (19), (22), 15 and the hypotheses of 4, 5.

N.B. The assumption (A6) is defined relative to substs  $Z'$  of  $X$ . In practice, the subsets  $Z'$  will be subsets of the form  $Z'(x) := Z(\partial^c \psi^c(x))$ ,  $X - A$ ,  $Z_j - Z_{j+1}$ . More precise formulations of (A6) are given in 9, (16), and the Uniform Halfspace (UHS) Conditions defined in 15. The assumption (A6) depends on properties of  $c$ -concave potentials defined on  $Y$ , and is not an absolute assumption on the geometry of the cost  $c$  like the previous (A0)–(A5).

### 3. OPTIMAL SEMICOUPLINGS AND MONGE-KANTOROVICH DUALITY

Now we briefly introduce the optimal semicoupling program and Kantorovich's dual max program.

**Definition 1** (Semicoupling). A *semicoupling* between source  $(X, \sigma)$  and target  $(Y, \tau)$  is a Borel-Radon measure  $\pi$  on the product space  $X \times Y$  with  $\text{proj}_Y \# \pi = \tau$  and  $\text{proj}_X \# \pi \leq \sigma$ .

The inequality  $\text{proj}_X \# \pi \leq \sigma$  holds if for every Borel subset  $O$ , the numerical inequality  $(\text{proj}_X \# \pi)[O] \leq \sigma[O]$  is satisfied. We remark that  $\pi$  is a coupling between  $\sigma$  and  $\tau$  when  $\text{proj}_X \# \pi = \sigma$ .

**Definition 2** ( $c$ -transforms). If  $\psi : Y \rightarrow \mathbb{R} \cup \{-\infty\}$  is any function on the target  $Y$ , then the  $c$ -Legendre transform  $\psi^c : X \rightarrow \mathbb{R} \cup \{+\infty\}$  is defined by

$$\psi^c(x) := \sup_{y \in Y} [\psi(y) - c(x, y)],$$

for  $x \in X$ .

If  $\phi : X \rightarrow \mathbb{R} \cup \{+\infty\}$  is any function on the source  $X$ , then the  $c$ -Legendre transform  $\phi^c : Y \rightarrow \mathbb{R} \cup \{-\infty\}$  is defined by the rule

$$\phi^c(y) = \inf_{x \in X} [c(x, y) + \phi(x)],$$

for  $y \in Y$ .

**Definition 3** ( $c$ -concavity). A function  $\psi : Y \rightarrow \mathbb{R} \cup \{-\infty\}$  is  $c$ -concave if  $\psi^{cc} = (\psi^c)^c$  coincides pointwise with  $\psi$ . Equivalently  $\psi$  is  $c$ -concave if there exists a lower semicontinuous function  $\phi : X \rightarrow \mathbb{R} \cup \{+\infty\}$  such that  $\phi^c = \psi$  pointwise.

The above definitions imply  $\psi, \psi^c$  satisfy the pointwise inequality

$$(5) \quad -\psi^c(x) + \psi(y) \leq c(x, y)$$

for all  $x \in X, y \in Y$ . The inequality (5) and especially the case of equality is very important for this thesis.

**Definition 4** ( $c$ -subdifferential). Let  $\psi : Y \rightarrow \mathbb{R} \cup \{-\infty\}$  be  $c$ -concave potential  $\psi^{cc} = \psi$ . Select  $y_0 \in Y$  where  $\psi(y_0)$  is finite-valued. The subdifferential  $\partial^c \psi(y_0) \subset X$  consists of those points  $x' \in X$  such that

$$-\psi^c(x') + \psi(y_0) = c(x', y_0).$$

Or equivalently such that for all  $y \in Y$ ,

$$\psi(y) - c(x', y) \leq \psi(y_0) - c(x', y_0).$$

The main technical result we need is Monge-Kantorovich duality. Let  $SC(\sigma, \tau) \subset \mathcal{M}_{\geq 0}(X \times Y)$  denote the set of all semicoupling measures  $\pi$  between source  $\sigma$  and target  $\tau$ . [Vil09, Theorem 5.10, pp.57].

**Monge-Kantorovich Duality.** Let  $c : X \times Y \rightarrow \mathbb{R} \cup \{+\infty\}$  be a cost satisfying Assumptions (A0)–(A5). Let  $\sigma, \tau$  be source and target measures on  $X, Y$ , respectively. Assume  $\sigma, \tau$  are absolutely continuous with respect to Hausdorff measures  $\mathcal{H}_X, \mathcal{H}_Y$ , respectively.

Then there exists unique closed domain  $A \hookrightarrow X$  called the activated domain such that the unique  $c$ -optimal semicoupling  $\pi_{opt}$  defines a coupling from  $1_A \cdot \sigma$  to  $\tau$ .

Moreover there exists  $c$ -concave potentials  $\psi^{cc} = \psi$  on  $Y$  and  $c$ -convex potentials  $\phi = \psi^c$  on  $X$  solving Kantorovich's dual max program:

$$(6) \quad \max_{\psi \text{ } c\text{-concave}} \left[ \int_A -\psi^c(x) d\sigma(x) + \int_Y \psi(y) d\tau(y) \right] = \min_{\pi \in SC(\sigma, \tau)} \left[ \int_{X \times Y} c(x, y) d\pi(x, y) \right]$$

Moreover the  $c$ -optimal semicoupling  $\pi$  is supported on the graph of the  $c$ -subdifferentials  $\partial^c \psi^c(x)$ , hence equality  $-\psi^c(x) + \psi(y) = c(x, y)$  holds  $\pi$ -a.e.

Assumptions (A0), ..., (A4), on the cost  $c$  imply various properties of  $c$ -convex potentials and  $c$ -subdifferentials. The first useful property is that  $c$ -subdifferentials are nonempty wherever the potentials  $\phi(x)$  or  $\psi(y)$  are finite, see Lemma 1 below. Recall the domain of  $\phi$  is defined  $dom(\phi) := \{x \in X \mid \phi(x) < +\infty\}$ .

**Lemma 1.** Let  $c : X \times Y \rightarrow \mathbb{R}$  be a cost satisfying Assumptions (A0)–(A2). Let  $\psi : Y \rightarrow \mathbb{R} \cup \{-\infty\}$  be a  $c$ -concave potential  $(\psi^c)^c = \psi$ . Abbreviate  $\phi = \psi^c$ . Suppose there exists  $y' \in Y$  such that  $\psi(y') \neq -\infty$ . Then:

- (i)  $\psi$  is an upper semicontinuous function; and
- (ii)  $\partial^c \psi(y)$  is a nonempty closed subset of  $X$  for every  $y \in dom(\psi)$ ; and
- (iii)  $\phi$  is lower semicontinuous function; and
- (iv)  $\partial^c \psi(y)$  is a nonempty closed subset of  $Y$  for every  $x \in dom(\phi)$ .

*Proof.* We omit the standard arguments. □

## 4. (TWIST) CONDITION

Thus far we have established the existence of optimal semicouplings and uniqueness of active domains. Now we describe the (Twist) hypothesis and the uniqueness of optimal couplings when the source measure  $\sigma$  is absolutely continuous with respect to the reference source measure  $\mathcal{H}_X^d$  in  $X$ .

The next definition elaborates Assumption (A4) from §2.

**Definition 5 ((Twist)).** Let  $c : X \times Y \rightarrow \mathbb{R}$  be cost function satisfying Assumptions (A0)–(A1). Then  $c$  satisfies (Twist) condition if for every  $x' \in X$  the rule

$$y \mapsto \nabla_x c(x', y)$$

defines an injective mapping  $\nabla_x c(x', \cdot) : \text{dom}(c_{x'}) \rightarrow T_{x'}X$ .

Observe that (Twist) condition is equivalent to the function

$$x \mapsto c_\Delta(x; y_0, y_1) := c(x, y_0) - c(x, y_1)$$

admitting no critical points on  $X$ , whenever  $y_0, y_1 \in Y$  are distinct. If  $X$  is closed manifold, then  $x \mapsto c_\Delta(x; y_0, y_1)$  certainly has critical points which violates (Twist). Our settings assume  $X$  is a manifold-with-corners with nontrivial boundary  $\partial X \neq \emptyset$ . The (Twist) condition requires  $c_\Delta$  admit no critical points on the interior of  $X$ , and all maxima/minima exist on the boundary. For instance, the repulsion costs constructed in [Mar] have the property that  $c_\Delta(x; y_0, y_1)$  converges to  $-\infty$  when  $x \rightarrow y_1$ , and converges to  $+\infty$  when  $x \rightarrow y_0$ , and all other level sets  $c_\Delta(-; y_0, y_1)^{-1}(s) \subset X$  are topologically connected and separating  $X$  into two components, for every  $s \in \mathbb{R}$ ,

**Proposition 1.** *Let  $c : X \times Y \rightarrow \mathbb{R}$  be a cost satisfying Assumptions (A0)–(A5). Let  $\psi : Y \rightarrow \mathbb{R} \cup \{-\infty\}$  be a  $c$ -concave potential. Then  $\psi$  is locally Lipschitz on its domain  $\psi : \text{dom}(\psi) \rightarrow \mathbb{R}$ .*

Recall the definition of semiconvexity [Vil09, Definition 10.10, pp.228]:

**Definition 6 (Semiconvexity).** A function  $\phi : X \rightarrow \mathbb{R} \cup \{+\infty\}$  is semiconvex on an open subset  $U$  of  $X$  with modulus  $C > 0$  at  $x_0 \in X$  if for every constant-speed geodesic path  $\gamma(t)$ , for  $0 \leq t \leq 1$  whose image is included in  $U$ , the inequality

$$(7) \quad \phi(\gamma(t)) \leq (1-t)\phi(\gamma(0)) + t\phi(\gamma(1)) + t(1-t)C \text{dist}(\gamma(0), \gamma(1))^2$$

is satisfied for  $0 \leq t \leq 1$ . The function is locally semiconvex if  $\phi$  is semiconvex at every  $x_0 \in U$ , with respect to a modulus  $C > 0$  depending uniformly on  $\gamma(0)$ ,  $\gamma(1)$  varying in compact subsets  $K$  of  $U$ .

**Lemma 2.** *Let  $c : X \times Y \rightarrow \mathbb{R}$  be cost satisfying Assumptions (A0)–(A2). Then every  $c$ -convex potential  $\psi^c : X \rightarrow \mathbb{R} \cup \{+\infty\}$  is  $\mathcal{H}_X$ -almost everywhere locally-Lipschitz on its domain  $\text{dom}(\psi^c) \subset X$ . Furthermore every  $c$ -convex potential is locally-semiconvex on  $\text{dom}(\psi^c)$ .*

*Proof.* Omitted.  $\square$

**Proposition 2.** *Let  $c$  be cost satisfying Assumptions (A0)–(A2). Then  $c$ -convex potentials  $\psi^c$  are  $\mathcal{H}_X$ -almost everywhere differentiable on  $\text{dom}(\psi^c) \subset X$ . Thus  $\text{dom}(D\psi^c)$  is a full  $\mathcal{H}_X$ -measure subset of  $\text{dom}(\psi^c)$ .*

*Proof.* According to Lemma 2, the  $c$ -convex potentials  $\psi^c$  are locally Lipschitz on their domains  $\text{dom}(\psi^c) \subset X$ . Rademacher’s theorem, [Vil09, Thm 10.8, pp.222], says locally Lipschitz functions are almost-everywhere differentiable on  $\text{dom}(\psi^c)$  with respect to  $\mathcal{H}_X$  on their domains. Therefore  $\nabla_x \psi^c$  exists almost everywhere on  $\text{dom}(\psi^c)$  as desired.  $\square$

## 5. STATEMENT OF MAIN RESULTS

The unique  $c$ -optimal semicouplings do not “activate” all of the source measure whenever we have strict inequality  $\int_X \sigma > \int_Y \tau$ . Equivalently the domain  $\text{dom}(\psi^c)$  is a nontrivial subset of the source  $X$ . Informally  $c$ -optimal semicouplings first allocate as much as possible from low-cost regions of the source. The union of all these activated low-cost regions defines a domain designated  $A \hookrightarrow X$ . Specifically we define  $A := \cup_{y \in Y} \partial^c \psi(y)$ .

Our first Theorem 1 describes a criteria to ensure that the canonical inclusion  $A \hookrightarrow X$  is a homotopy-isomorphism.

**Theorem 1.** *Let  $c$  be cost satisfying Assumptions (A0)–(A4), and let  $\sigma, \tau$  be absolutely continuous with respect to  $\mathcal{H}_X, \mathcal{H}_Y$ , respectively and satisfying (1). Let  $\pi$  be a  $c$ -optimal semicoupling from  $\sigma$  to  $\tau$ , with dual  $c$ -concave potential  $\psi^{cc} = \psi$ , 3. Let  $A = \text{dom}(\psi^c)$  be the active domain of the unique  $c$ -optimal semicoupling. Let  $\beta := \dim(Y) + 2$ . Suppose every  $x \in X - A$  has the property that*

$$(8) \quad \eta_{\text{avg}}(x) := (\mathcal{H}_Y[Y])^{-1} \int_Y (c(x, y) - \psi(y))^{-\beta} \cdot \nabla_x c(x, y) \, d\mathcal{H}_Y(y),$$

*is bounded away from zero, uniformly with respect to  $y \in Y$ .*

*Then the inclusion  $A \hookrightarrow X$  is a homotopy-isomorphism, and there exists strong deformation retract  $X \leadsto A$ .*

The condition (8) is an assumption similar to (A6), and is key technical hypothesis for constructing the deformation retract. We remark that  $\eta_{\text{avg}}(x)$  is uniformly bounded away from zero whenever the gradients  $\nabla_x c(x, y)$ ,  $y \in Y$ , occupy a non-trivial halfspace for every  $x$ . Or equivalently, whenever the closed convex hull of  $\nabla_x c(x, y)$ ,  $y \in Y$ , is disjoint from the origin  $\mathbf{0}$  in  $T_x X$ . Without the estimate  $\|\eta_{\text{avg}}(x)\| \geq C > 0$ , our methods could only conclude that the source  $X$  deformation retracts onto  $\epsilon$ -neighborhoods  $A_\epsilon = \{x \in X \mid d(x, A) < \epsilon\} \hookrightarrow X$  of  $A \hookrightarrow X$ , where  $\epsilon > 0$  is a sufficiently small real number.

Now Theorem 1 is an important preliminary result. Our primary motivations are constructing retracts onto large codimension subspaces. Our next Theorem 2 constructs further homotopy-reductions from the active domains  $A \subset X$  to higher codimension subvarieties  $\mathcal{Z}$ .

Let  $Z = Z_\psi : 2^Y \rightarrow 2^X$  be the singularity functor  $Z(Y_I) = \cap_{y \in Y_I} \partial^c \psi(y)$ . For integers  $j \geq 1$ , we define  $Z_j$  to be the subset of  $x \in X$  where the local tangent cone is at least  $j$ -dimensional at  $x$ . The formal definition is provided in §12, c.f. 3, 4. The singularity structure is naturally cellulated by the cells

$$Z'(x) := Z(\partial^c \psi^c(x))$$

and admits a natural filtration

$$(X =: Z_0) \leftarrow (A =: Z_1) \leftarrow Z_2 \leftarrow Z_3 \leftarrow \dots$$

of  $X$  by “subvarieties”  $Z_0, Z_1$ , etc. Under the assumptions (A0)–(A3), the cell  $Z'(x)$  has well-defined tangent space  $T_x Z'(x)$ . Let  $pr_{Z'} : T_x X \rightarrow T_x Z'(x)$  be the orthogonal projection defined relative to the Riemannian structure on  $X$ .

For the following construction we need introduce an auxiliary function  $\nu_x(y)$  which will depend on a source point  $x$  and its  $c$ -subdifferential  $\partial^c \phi(x)$  relative to a  $c$ -convex potential  $\phi = \psi^c$ . The assumptions (A0)–(A5) imply the zeros of  $\psi(y_0) - \psi(y) - c_\Delta(x, y, y_0)$  have a well-defined *order* or *multiplicity* as defined in the complex analysis. The order  $ord(y_0)$  is a positive integer obtained by taking power series expansion of  $\psi(y_0) - \psi(y) - c_\Delta(x, y, y_0)$  with respect to  $d_Y(y, y_0)$  for  $y$  in a sufficiently small neighborhood of  $y_0$ .

**Definition 7.** Let  $d_Y$  be the metric distance on the target  $Y$ . Let  $\phi = \psi^c$  be a  $c$ -convex potential on the source  $X$ . For  $x \in \text{dom}(\phi)$  assume  $\partial^c \phi(x)$  has *finite* cardinality. Define

$$(9) \quad \nu_x(y) := \min(1, \prod_{y_0 \in \partial^c \phi(x)} d_Y(y, y_0)^{ord(y_0)}),$$

where  $ord(y_0)$  is the order of the zero of  $\psi(y_0) - \psi(y) - c_\Delta(x, y, y_0)$  at  $y = y_0$ .

**Definition 8.** For  $x \in X$ , abbreviate  $Y'(x) := \text{dom}(c_x)$ . For a real parameter  $\beta > 0$  and  $x \in X$ , define the collection of tangent vectors

$$\eta(x, y) := |\psi(y_0) - \psi(y) + c_\Delta(x; y, y_0)|^{-\beta} \cdot pr_{Z'}(\nabla_x c_\Delta(x; y, y_0)).$$

For every  $x$ , define

$$(10) \quad \eta_{avg}(x) := (\mathcal{H}_Y[Y'(x)])^{-1} \int_{Y'(x)} \eta(x, y) \cdot \nu_x^\beta(y) \cdot d\mathcal{H}_Y(y).$$



À priori,  $\eta_{avg}(x)$  is a vector in  $T_x Z'$ . We say (UHS) Conditions are satisfied on  $Z'(x)$  if  $\eta_{avg}$  is bounded away from zero uniformly with respect to the source point  $x$  (c.f. Definition 15).

Our next result identifies the maximal index  $J \geq 1$  such that  $(A = Z_1) \hookleftarrow Z_{J+1}$  is a homotopy-isomorphism, and even a strong deformation retract.

**Theorem 2.** *Let  $c$  be a cost satisfying Assumptions (A0)–(A5). Suppose  $\sigma, \tau$  are source, target measures absolutely continuous with respect to  $\mathcal{H}_X, \mathcal{H}_Y$ , respectively and satisfying (1). Let  $\psi^{cc} = \psi$  be a  $c$ -concave potential (3) solving Kantorovich's dual max program, and suppose  $\phi = \psi^c$  is a  $c$ -convex potential such that  $\partial^c \phi(x)$  has finite cardinality for every  $x$ . Let  $j \geq 1$  be an integer.*

*(i) Suppose there exists a parameter  $\beta > 0$  such that  $\eta_{avg}(x')$  is bounded away from zero, uniformly with respect to  $x' \in Z_j - Z_{j+1}$ ; and*

*(ii) Suppose the intersection  $\partial^c \psi^c(x') \cap Z_{j+1} \neq \emptyset$  is nonempty for every  $x' \in Z_j - Z_{j+1}$ .*

*Under the above hypotheses, the inclusion  $Z_{j+1} \hookrightarrow Z_j$  is a homotopy-isomorphism, even a strong deformation retract. Furthermore if  $J \geq 1$  is the maximal integer such that every  $x' \in Z_J - Z_{J+1}$  satisfies conditions (i)–(ii), then the inclusion  $Z_{J+1} \hookrightarrow Z_1$  is a homotopy-isomorphism, even a strong deformation retract.*

**Example.** Theorem 1 has the following application to quadratic costs

$$(11) \quad b(x, y) := d(x, y)^2/2 = \|x - y\|^2/2$$

for closed subsets  $X, Y$  in a Euclidean space  $\mathbb{R}^N$ . The gradients  $\{\nabla_x c(x_0, y)\}_{y \in Y}$  are a subset of  $T_{x_0} X$ . Observe that the closed convex hull  $\text{conv}\{\nabla_x c(x_0, y)\}_{y \in Y}$  contains the origin  $\mathbf{0}$  in  $T_{x_0} X$  if and only if  $x_0$  lies in the convex hull  $\text{conv}(Y)$  of  $Y$  in  $\mathbb{R}^N$ .

If the activated domain  $A$  of a  $d^2/2$ -optimal semicoupling contains the convex hull  $\text{conv}(Y) \subset A$ , then Theorem 1 implies the inclusion  $A \subset X$  is a homotopy-isomorphism. The proof of Theorem 1 constructs an explicit strong deformation retract of  $X$  onto  $A$ . We find  $A$  contains  $\text{conv}(Y)$  whenever the ratio  $\int_X \sigma / \int_Y \tau > 1$  is sufficiently close to  $1^+$ , i.e. whenever the active domain  $A$  is a sufficiently large subset of  $X$ .

The hypotheses of 1 are never satisfied when we restrict  $b$  (11) to a convex subset  $X := F$  and its boundary  $Y := \partial F$ . If  $\sigma = 1_F \mathcal{L}$ , and if the target  $\tau$  is supported on  $\partial F$ , and if (1) is satisfied with strict inequality, then the active domain  $A$  of the unique  $b$ -optimal semicoupling from  $\sigma$  to  $\tau$  will not contain  $\text{conv}(Y)$ . Consequently  $\eta_{avg}$  will vanish somewhere on  $X - A$ , and the hypotheses of 1 are violated. In [Mar, Ch.4] a repulsion cost  $c|\tau$  is defined which will satisfy (UHS) conditions throughout the non active domains and satisfy hypotheses of 1.

## 6. AVERAGED GRADIENTS

Suppose we have a Radon measure  $\nu$  on the target  $Y$ , absolutely continuous with respect to  $\mathcal{H}_Y$ , and we take the  $\nu$ -average of a  $\nu$ -family of potentials  $f$ . The following lemma formally establishes that this averaged-gradient is indeed the gradient field of a continuously differentiable “averaged” potential.

**Lemma 3.** *Let  $\beta > 0$ . Let  $\nu_1, \nu_2, \nu_3, \dots$  be a sequence of empirical probability measures, i.e. renormalized sums of Dirac masses, which converge as  $N \rightarrow +\infty$  in the weak-\* topology to the renormalized probability measure  $(\bar{\nu}[Y])^{-1} \cdot \bar{\nu}$  on  $Y$ . Then:*

(i) *For  $x \in X - A$ , the limit*

$$(12) \quad \lim_{N \rightarrow +\infty} \int_Y (c(x, y) - \psi(y))^{-\beta} d\nu_N(y)$$

*exists and converges to the finite integral*

$$(13) \quad f_{avg}(x) := (\bar{\nu}[Y])^{-1} \int_Y (c(x, y) - \psi(y))^{-\beta} d\bar{\nu}(y).$$

(ii) *The rule  $f_{avg} : X - A \rightarrow \mathbb{R}$  defines a continuously differentiable function with gradient*

$$\nabla_x f_{avg} = (\bar{\nu}[Y])^{-1} \int_Y \nabla_x (c(x, y) - \psi(y))^{-\beta} d\bar{\nu}(y).$$

*Proof.* If  $c$  satisfies (A0)–(A4), then the limit defining  $f_{avg}$  converges uniformly on compact subsets of  $X - A$ . So the limit (12) exists and is finite. Moreover the uniform convergence on compact subsets implies (ii), since the approximants are continuously differentiable on  $X - A$ . Therefore  $\nabla_x f_{avg}$  is the average of  $\nabla_x (c(x, y) - \psi(y))^{-\beta}$  with respect to  $\bar{\nu}$ , as desired.  $\square$

For  $y \in Y$ ,  $x \in X - A$ , and  $\beta > 0$ , we abbreviate

$$(14) \quad f_y(x) := (c(x, y) - \psi(y))^{-\beta}$$

and define

$$(15) \quad f_{avg} : X - A \rightarrow \mathbb{R}, \quad f_{avg}(x) = (\nu[Y])^{-1} \int_Y f_y(x) d\nu(y).$$

**Definition 9 (Property (C)).** The collection of functions  $\{f_y | y \in Y\}$  satisfies Property (C) throughout  $X - A$  with respect to the uniform probability measure  $\bar{\nu} := \frac{1}{\nu[Y]} \nu$  if there exists constant  $C > 0$  such that

$$(16) \quad \|\nabla_x f_{avg}\| \geq C \int_Y \|\nabla_x f_y\| d\bar{\nu}(y)$$

pointwise throughout  $X - A$ .

When  $Y$  is finite,  $\#(Y) < +\infty$ , the estimate (16) requires the ratio

$$\|\nabla_x f_{avg}\| / \max_{y \in Y} \|\nabla_x f_y\|$$

be uniformly bounded away from zero throughout  $X - A$ . In this case, Property (C) and (UHS) conditions ensures the divergence of the average  $\nabla_x f_{avg}$  whenever a gradient summand  $\nabla_x f_y$  diverges.

When  $Y$  is infinite with positive dimension, the pointwise divergence of an integrand  $\|f_y\| \rightarrow +\infty$  need not imply the divergence of the average  $f_{avg}$  for every choice of  $\beta > 0$ : rather the rate at which  $f_y$  diverges must be sufficiently large. When cost  $c$  satisfies Assumptions (A0)–(A5), Lemma 1 proves that  $c$ -concave potentials  $\psi : Y \rightarrow \mathbb{R} \cup \{-\infty\}$  are locally Lipschitz throughout  $\text{dom}(\psi)$ . This implies  $\beta = \dim(Y) + 2$  is sufficient. If  $\{x_k\}_k$  is a countable sequence in  $X - A$ , then  $f_{avg}(x_k)$  diverges to  $+\infty$  if and only if  $f_y(x_k)$  diverges for  $y$  belonging to some subset  $V \subset Y$ .

**Lemma 4.** *Let  $c : X \times Y \rightarrow \mathbb{R}$  be cost satisfying (A0)–(A5), as above, and  $A \subset X$  the active domain of a  $c$ -optimal semicoupling. Fix  $y_0 \in Y$ , and abbreviate  $f_0(x) := f_{y_0}(x) = (c(x, y_0) - \psi(y_0))^{-\beta}$  for some  $\beta > 0$ ,  $x \in X$ . Suppose:*

(a)  $\nabla_x c(x, y_0)$  is uniformly bounded away from the origin; and

(b)  $\nabla_{xx}^2 c(x, y_0)$  is uniformly bounded above with respect to  $x \in X - A$ .

*Then for every  $K > 0$ , there exists  $\epsilon > 0$  such that  $\nabla_{xx}^2 f_0 \geq K \text{Id} > 0$  in the direction of  $\nabla_x c(x, y_0)$  throughout the  $\epsilon$ -neighborhood of  $\{f_0 = +\infty\}$  in  $X - A$ .*

*Proof.* The function  $f_0$  is well-defined on  $\{c(x, y_0) > \psi(y_0)\} \subset X$ . If  $x$  converges to  $x_\infty \in \{c(x, y_0) \leq \psi(y_0)\}$ , then both  $f_0$ ,  $\nabla_x f_0$  diverge to infinity. We find  $\nabla_{xx}^2 f_0$  is equal to

$$\beta(1+\beta)(c(x, y_0) - \psi(y_0))^{-2-\beta} \nabla_x c(x, y_0) \otimes \nabla_x c(x, y_0) - \beta(c(x, y_0) - \psi(y_0))^{-1-\beta} \nabla_{xx}^2 c(x, y_0).$$

By Proposition 9 the gradients  $\nabla_x c(x, y_0)$  are uniformly bounded away from zero in neighborhoods of  $\{f_0 = +\infty\}$ . Assumption (A2) implies  $\nabla_{xx}^2 c(x, y_0)$  is uniformly bounded above on superlevel sets  $\{f_0 \geq T\}$  for all  $T > 0$ . Factoring out the term  $\beta(c(x, y_0) - \psi(y_0))^{-1-\beta}$ , we find  $\nabla_{xx}^2 f_0$  is positively proportional

$$(17) \quad (1 + \beta)(c(x, y_0) - \psi(y_0))^{-1} \nabla_x c(x, y_0) \otimes \nabla_x c(x, y_0) - \nabla_{xx}^2 c(x, y_0).$$

We observe (17) diverges to  $+\infty$  with  $(c(x, y_0) - \psi(y_0))^{-1}$ . This implies  $\nabla_{xx}^2 f_0$  is positive semidefinite when  $c(x, y_0) - \psi(y_0) > 0$  is sufficiently small, and strongly convex

$$\nabla_{xx}^2 f_0 \geq K > 0$$

in the direction of  $\nabla_x c(x, y_0)$ .

By Assumption (A1) the sublevels  $\{x \mid \psi(y_0) \leq c(x, y_0) \leq \psi(y_0) + \epsilon'\}$  are compact subsets of  $X - A$  for every  $\epsilon' > 0$ . This implies a sufficiently small  $\epsilon > 0$  exists

for which  $\nabla_{xx}^2 f_0 \geq K.Id > 0$  throughout the  $\epsilon$ -neighborhood of  $\{f_0 = +\infty\}$  in the direction  $\nabla_x c(x, y_0)$ .  $\square$

**Example.** To illustrate Lemma 4, consider the  $f(x) = x^{-\beta}$  for  $x \geq 0$ ,  $\beta > 0$ . The gradient flow  $x' = (-\beta)x^{-1-\beta}$  is bounded away from zero in neighborhoods of the pole at  $x = 0$ , and indeed diverges to  $+\infty$ . Moreover  $f''(x)$  is obviously bounded away zero and diverging to  $+\infty$  as  $x \rightarrow 0^+$ .

For initial condition  $x_0 > 0$ , the integral curve of the negative gradient flow is equal to  $x(s) = (x_0^{2+\beta} - \beta(\beta+2)s)^{1/(2+\beta)}$ , which converges in finite time to the pole at  $x = 0$  over the interval  $0 \leq s \leq \frac{1}{\beta(\beta+2)}x_0^{2+\beta}$ . Thus we find  $\omega(x_0) = \frac{1}{\beta(\beta+2)}x_0^{2+\beta}$  varies continuously with respect to  $x_0 > 0$ , and is even Lipschitz. Compare Lemma 7 below.

## 7. FINITE-TIME BLOW-UP

The blow-up in finite time is typical property of the gradient flow defined by the potentials  $f_0$ . Actually our applications require verifying these same properties for the averaged potential  $f_{avg}$  and its gradient  $\nabla_x f_{avg}$ .

**Lemma 5** (Asymptotic Convexity of Average Gradient Flow). *Let  $f_{avg}$  be the average defined in Lemma 3, equation (13), with exponent  $\beta = \dim(Y) + 2$ . If the distance from  $x \in X - A$  to the boundary  $\partial A$  is sufficiently small, then  $f_{avg}$  is strongly convex in the direction of  $\nabla_x f_{avg}$ .*

*Proof.* Let  $x_k$  be a sequence in  $X - A$  converging to a point  $x_\infty \in \{f_{avg} = +\infty\}$ . The choice of  $\beta$  says  $f_{avg}$  diverges if and only integrands  $f_y$  diverge, and there exists a subset  $V \subset Y$  such that  $f_y$  diverges to  $+\infty$  for every  $y \in V$ . The divergence of  $f_y, y \in V$  also implies the divergence of the gradients  $\nabla_x f_y$  and Hessians  $D_{xx}^2 f_y$ , (see (17) in proof of 4). Moreover the Hessians  $D^2 f_{avg}|_{x_k}$  are positive semidefinite when  $k$  is sufficiently large, being the asymptotic to the average rank-one quadratic forms  $\langle \nabla_x f_y, - \rangle^2$ . So  $D^2 f_{avg}[\nabla_x f_{avg}]$  is asymptotic to the average of

$$(18) \quad \langle \nabla_x f_y, \nabla_x f_{avg} \rangle^2$$

for  $y \in V$ . Now we claim

$$\lim_{k \rightarrow +\infty} \langle \nabla_x f_y|_{x_k}, \nabla_x f_{avg}|_{x_k} \rangle^2 = +\infty, \text{ unless } \nabla_x f_y, \nabla_x f_{avg} \text{ are orthogonal.}$$

The (UHS) conditions imply  $\nabla_x f_{avg}$  is uniformly bounded away from zero, and therefore the inner products (18) are not identically zero for all  $y \in V$ . This implies the divergence of  $D^2 f_{avg}[\nabla_x f_{avg}]|_{x_k}$  as  $k \rightarrow +\infty$ .  $\square$

**Lemma 6** (Finite-time Blow-up). *Suppose the functions  $\{f_y \mid y \in Y\}$  satisfy (16) as above. Then for every initial value  $x_0 \in \text{dom}(f_{\text{avg}})$ , the gradient flow defined by the average gradient  $x' = \nabla_x f_{\text{avg}}$  diverges to infinity in finite time.*

*Proof.* The estimate (16) shows the gradient  $\nabla_x f_{\text{avg}}$  is uniformly bounded away from zero in the neighborhoods of the poles  $\{f_{\text{avg}} = +\infty\}$  in  $X - A$ . Moreover  $f_{\text{avg}}$  is asymptotically convex in neighborhoods of the poles using Lemma 5 along directions of  $\nabla_x f_{\text{avg}}$ . This implies the gradient flow  $x' = \nabla_x f_{\text{avg}}$ ,  $x(0) = x_0$  blows-up in finite-time for every initial value  $x_0 \in X - A$ .  $\square$

Informally the estimate (16) implies every step in the discretized gradient flow (e.g., Euler scheme) has a definite size. Meanwhile the asymptotic convexity of Lemma 4 implies the discretized gradient flow well approximates the continuous gradient flow. But if step-sizes have a definite magnitude, then we definitely approach the poles after a finite number of steps and the integral curves blow-up in finite time.

## 8. CONTINUITY OF $\omega$

The blow-up in finite time (Lemma 6) implies the maximal forward-time interval of existence for the gradient flow is a bounded interval  $I(x_0) := [0, \omega(x_0)) \subset [0, +\infty)$ . For general ordinary differential equations, it's known that  $\omega(x_0)$  is lower semicontinuous as a function of  $x_0$ : for a sequence of initial values  $x_0, x_1, \dots$  converging to some  $x_\infty$ , we have  $\omega(x_\infty) \leq \liminf_{k \rightarrow +\infty} \omega(x_k)$ . See [Har64, Theorem 2.1, pp.94]. In our particular setting, it is further necessary to establish the continuity of this maximal interval of existence. This is established in Lemma 7 below.

**Lemma 7** (Continuity of Max Interval of Existence). *Let  $c$  satisfy Assumptions (A0)–(A5). Then the maximal intervals of existence  $I(x_0) = [0, \omega(x_0))$  of solutions to the initial value problem 20 vary continuously with respect to the initial point  $x_0$ . In otherwords,  $x_0 \mapsto \omega(x_0)$  varies continuously with  $x \in X - A$ .*

*Proof.* Our assumptions imply the domain  $X - A$  is a complete open set. Furthermore  $\eta_{\text{avg}}(x)$  being uniformly bounded away from zero implies the trajectories  $s \mapsto \Psi(x_0, s)$  are finite on compact subsets of  $X - A$ . Therefore  $\omega(x_0)$  is characterized by the two limits

$$\lim_{s \rightarrow \omega(x_0)^-} \Psi(x_0, s) \in \partial A, \quad \lim_{s \rightarrow \omega(x_0)^-} \|\Psi(x_0, s)\| = +\infty,$$

and here we take advantage of the divergence  $\|\eta_{\text{avg}}(x)\| \rightarrow +\infty$  as  $x \rightarrow \partial A$ .

Moreover the asymptotic concavity (Proposition 5) of the average potential defining  $\eta_{\text{avg}}(x)$  implies the flow defined by 20 is asymptotically contracting. This implies  $\omega(x_0)$  is actually a Lipschitz function of  $x_0$ , i.e. satisfying  $|\omega(x_1) - \omega(x_0)| \leq C \|x_1 - x_0\|$  for some constant  $C = C(x_0)$  depending on  $x_0 \in X - A$ . Hence  $\omega$  is a continuous function, as desired.  $\square$

## 9. PROOF OF THEOREM 1

*Proof of Theorem 1.* The active domain  $A$  can be expressed as  $\cup_{y \in Y} \{x \mid c(x, y) \leq \psi(y)\}$  for the Kantorovich potential  $\psi : Y \rightarrow \mathbb{R}$ . If (UHS) condition is satisfied at  $x \in X - A$ , then the  $\nu$ -average  $\eta_{avg}(x) = \nabla_x f_{avg}(x)$  gradients  $\{\nabla_x c(x, y) \mid y \in Y\}$  of the gradients is nonzero and uniformly bounded away from zero. We divide the argument into two cases.

(Case I) Assume  $Y$  is finite with  $N = \#(Y) < +\infty$ . Define

$$(19) \quad \eta_{avg}(x) := N^{-1} \sum_{y \in Y} [(c(x, y) - \psi(y))^{-\beta} \cdot \nabla_x c(x, y)].$$

Evidently when  $Y$  is finite, the sum (19) is finite vector. Then  $x \mapsto \eta_{avg}(x)$  is a well-defined nonvanishing vector field on  $X - A$  which diverges whenever a denominator converges  $c(x, y) \rightarrow \psi(y)^+$ . Thus  $\eta(x, avg)$  is finite if and only if  $x \in X - A$ .

We propose integrating the vector field  $x \mapsto \eta_{avg}(x)$  throughout  $X - A$  to obtain the desired retraction. For every initial point  $x_0 \in X - A$ , there exists a unique solution  $\Psi(x_0, s)$  to the ordinary differential equation

$$(20) \quad \frac{d}{ds} \big|_{(x_0, s)} \Psi = \eta_{avg}(\Psi(x_0, s)), \quad \Psi(x_0, 0) = x_0,$$

and defined over a maximal interval of existence  $I(x_0) = [0, \omega(x_0))$ . According to Lemma 7 the maximal interval  $I(x_0)$  varies continuously with respect to  $x_0 \in X - A$ . Moreover orbits  $\{\Psi(x_0, s) \mid s \in I(x_0)\}$  converge in finite-time to the boundary  $\partial(X - A) = \partial A$  for every initial value  $x_0 \in X - A$ . Lemma 7 also proves the orbits can be continuously reparameterized to obtain a continuous mapping  $\Psi' : (X - A) \times [0, 1] \rightarrow X$  defined by

$$(21) \quad \Psi'(x_0, s) = \Psi(x_0, s\omega(x_0)).$$

(Case II) Suppose  $Y$  is infinite set, with uniform measure  $\mathcal{H}_Y$ . We define  $\eta_{avg}(x)$  according to the vector-valued Bochner integral

$$(22) \quad \eta_{avg}(x) := \left( \int_{dom(c_x)} d\mathcal{H}_Y \right)^{-1} \int_{dom(c_x)} (c(x, y) - \psi(y))^{-\beta} \cdot \nabla_x c(x, y) \nu_x^\beta(y) d\mathcal{H}_Y(y),$$

where  $dom(c_x)$  is closed compact subset of  $Y$  for every  $x \in X$  by Assumption (A0). Assumption (A5) implies the exponent  $\beta = \dim(Y) + 2$  is suitable according to Proposition 1. Lemma 3 implies the vector field  $\eta_{avg}(x) = \nabla_x f_{avg}$  is the gradient of a continuously differentiable potential  $f_{avg}$  defined on  $X - A$ .

The proof proceeds as in (Case I). The vector field  $x \mapsto \eta_{avg}(x)$  is well-defined nonvanishing vector field on  $X - A$  which diverges to  $+\infty$  whenever some denominator converges  $c(x, y) \rightarrow \psi(y)^+$ . We integrate the gradient fields and obtain the retraction of  $X - A$  onto the poles  $\partial A$ . The flow converges in finite-time by Lemma

FIGURE 1. Singularity structure  $Z$  on active domain  $A = Z_1$ 

6. We reparameterize the flow according to equation (21), and obtain a continuous deformation using Lemma 7.  $\square$

## 10. KANTOROVICH'S CONTRAVARIANT SINGULARITY FUNCTOR

Our thesis proposes a bridge between measure and algebraic topology. The bridge is realized by contravariant functors  $Z : 2^Y \rightarrow 2^X$ , where  $2^Y$  designates the category of closed subsets  $Y_I$  of  $Y$ , and where morphisms are the inclusions  $Y_I \subset Y_J$  between closed subsets  $Y_I, Y_J$  whenever they exist. Likewise for  $2^X$ .

The singularity functor  $Z : 2^Y \rightarrow 2^X$  is defined relative to a cost  $c$  on  $X \times Y$  satisfying Assumptions (A0)–(A4). Let  $\sigma, \tau$  be absolutely continuous with respect to the Hausdorff measures  $\mathcal{H}_X, \mathcal{H}_Y$  on  $X, Y$  respectively. By Monge-Kantorovich duality 3 there exists unique  $c$ -minimizing measures and  $c$ -concave potentials  $\psi^{cc} = \psi$  on  $Y$ . The  $c$ -subdifferential  $\partial^c \psi$  is uniquely determined.

**Definition 10** (Kantorovich Singularity). Let  $\psi : Y \rightarrow \mathbb{R} \cup \{-\infty\}$  be a  $c$ -concave potential on  $Y$ . The Kantorovich functor  $Z : 2^Y \rightarrow 2^X$  is defined by the rule

$$Y_I \mapsto Z(Y_I) := \cap_{y \in Y_I} \partial^c \psi(y).$$

We declare  $Z(\emptyset_Y) := X$  for the empty subset  $\emptyset_Y$  of  $Y$ .

The definition  $Z(Y_I) = \cap_{y \in Y_I} Z(y)$  yields a *contravariant functor*  $Z : 2^Y \rightarrow 2^X$  where the inclusions  $Y_I \hookrightarrow Y_J$  correspond to reverse inclusions  $Z(Y_I) \leftarrow Z(Y_J)$  in  $X$ . The contravariant functor  $Z$  is uniquely prescribed by the choice of  $(c, \sigma, \tau)$  whenever  $c$  satisfies (A0)–(A4), and  $\sigma, \tau$  are absolutely continuous as defined above. For  $(c, \sigma, \tau)$  satisfying the assumptions above, the singularity  $Z = Z(c, \sigma, \tau)$  will generally admit many closed subsets  $Y_I \subset Y$  for which the cells  $Z(Y_I) \subset X$  are empty  $Z(Y_I) = \emptyset_X$ . It is useful to restrict ourselves to the nontrivial image of  $Z$  and formally define the support.

**Definition 11.** The support of the contravariant functor  $Z : 2^Y \rightarrow 2^X$  is the subcategory of  $2^Y$ , denoted  $spt(Z)$ , whose objects are the closed subsets  $Y_I$  of  $Y$  for which  $Z(Y_I)$  is nonempty subset of  $X$ .

So  $spt(Z) = \{Y_I \subset Y \mid Z(Y_I) \neq \emptyset_X\}$ . Note that  $\emptyset_Y \subset Y$  is object in subcategory  $spt(Z)$ , since  $Z(\emptyset_Y) = X$  according to Definition 10.

For given  $x \in X$ , we interpret  $Z'(x) := Z(\partial^c \psi^c(x))$  as a “cellular neighborhood” of  $x$  in the active domain  $A =: Z_1$ . This cellular neighborhood is not necessarily an open subset of the source  $X$ . The  $c$ -concavity  $\psi^{cc} = \psi$  provides explicit equations describing  $Z'(x)$ . See Section 11 and (26).



FIGURE 2. Singularity structure on  $Z_2$ 11. LOCAL TOPOLOGY AND LOCAL DIMENSIONS OF  $Z$ 

The present section examines the local differential topology and dimensions of Kantorovich's contravariant functor  $Z : 2^Y \rightarrow 2^X$ , where the cost  $c$  satisfies Assumptions (A0)–(A5). If  $\psi : Y \rightarrow \mathbb{R} \cup \{-\infty\}$  is a  $c$ -concave potential, then we first seek a description of the differential topology of the cells

$$Z'(x) := Z(\partial^c \psi^c(x)).$$

The main results are [3](#) and [4](#). The proof of [3](#) is essentially a local adaptation of the recent work of [\[KM18\]](#), and [4](#) is a corollary to a theorem of Alberti [\[Alb94\]](#).

Let the reader recall the definitions of  $c$ -concavity [3](#) and  $c$ -subdifferentials [4](#). The  $c$ -concavity  $\psi^c = \psi$  represents a pointwise inequality on  $Y$ , namely

$$(23) \quad -\psi^c(x) + \psi(y) \leq c(x, y) \text{ for all } (x, y) \in X \times Y.$$

The case of equality  $-\psi^c(x) + \psi(y) = c(x, y)$  is most important, and occurs if and only if  $y \in \partial^c \psi^c(x)$ , and if and only if  $x \in \partial^c \psi(y)$ . We recall that  $c$ -optimal semicouplings  $\pi$  are supported on the graphs of  $c$ -subdifferentials, hence the equality  $-\psi^c(x) + \psi(y) = c(x, y)$  holds  $\pi$ -a.e., when  $\psi$  is a  $c$ -concave maximizer to Kantorovich's dual program.

Recall that if  $\phi : \mathbb{R}^N \rightarrow \mathbb{R} \cup \{+\infty\}$  is a convex lower semicontinuous function on some  $N$ -dimensional vector space, then the following are equivalent [\[ET99, I.5.3, pp.23\]](#):

- $\phi$  is continuous and finite at  $x \in \mathbb{R}^N$ , with  $\partial\phi(x)$  a singleton;
- $\phi$  is differentiable at  $x$  with  $\partial\phi(x) = \{D\phi(x)\}$ .

Thus  $\phi$  is continuous throughout the interior of  $\text{dom}(\phi)$ . So if  $\phi = \psi^c$  is a  $c$ -convex potential on  $X$ , then  $\partial^c \phi(x)$  is multivalued if and only if  $\phi$  is non-differentiable at  $x$ .

Returning to  $c$ -convexity: in case  $\partial^c \phi(x_0)$  is a singleton, say  $\{y_0\}$ , then  $-\psi^c(x_0) + \psi(y) < c(x_0, y)$  for every  $y \neq y_0$ . The set of  $x$ 's with  $y_0 \in \partial^c \phi(x)$  is then characterized as the subset where

$$\psi(y_0) \geq c(x, y_0) + \phi(x)$$

according to [\(23\)](#), i.e. as the superlevel set of  $x \mapsto c(x, y_0) + \phi(x)$  and Assumptions (A0)–(A3) imply these superlevel sets are closed. More concretely, we find  $Z(\{y_0\})$  is characterized by the equations

$$(24) \quad Z(\{y_0\}) = \{x \in X \mid \psi(y_0) = c(x, y_0) + \phi(x)\}.$$

Equivalently, we find  $x \in \partial^c \psi(y_0)$  if and only if

$$(25) \quad y_0 \in \operatorname{argmax}[\{\psi(y) - c(x, y) + c(x, y_0) \mid y \in Y\}].$$



If  $x$  is such that  $\partial^c \psi^c(x)$  is not a singleton, say including two distinct points  $y_0, y_1$ , then  $x$  satisfies the equations

$$\psi^c(x) + c(x, y_0) = \psi(y_0), \quad \psi^c(x) + c(x, y_1) = \psi(y_1).$$

Subtracting these two equations, we can eliminate  $\phi(x)$  and obtain the equation

$$0 = \psi(y_1) - \psi(y_0) + c_\Delta(x, y_0, y_1),$$

where we abbreviate  $c_\Delta(x; y, y') := c(x, y) - c(x, y')$  for the two-pointed cross difference. If  $y_0 \in \partial^c \psi^c(x)$ , and  $y \notin \partial^c \psi^c(x)$ , then

$$0 < \psi(y) - \psi(y_0) + c_\Delta(x, y_0, y).$$

**Definition 12.** Let  $\phi$  be a  $c$ -convex potential on  $X$ ,  $\phi^{cc} = \phi$ . We say  $x \in \text{dom}(\phi)$  is a singular point if  $\phi$  is not differentiable at  $x$ .

If  $x_0 \in X$  is a singular point, then  $\partial^c \phi(x_0)$  is not a singleton and the cell  $Z'(x_0)$  is described by the system of equations

$$(26) \quad Z'(x_0) = \{x \in X \mid 0 = \psi(y_0) - \psi(y_1) + c_\Delta(x; y_1, y_0), \quad y_1 \in \partial^c \psi^c(x_0), y_1 \neq y_0\}$$

according to (25), where  $\psi = \phi^c$ .

From the equations (26) we obtain the following:

**Lemma 8.** *Under Assumptions (A0)–(A3), the cell  $Z'(x) = Z(\partial^c \psi^c(x))$  is a closed locally DC-subvariety in  $X$  for every singular point  $x \in \text{dom}(\psi^c)$ .*

*Proof.* The cell  $Z'(x)$  is the intersection of sets of the form  $\partial^c \psi(y)$ , which are closed [1](#). Assumption (A1) implies  $x \mapsto \nabla_{xx}^2 c(x, y)$  is locally bounded above on  $X$ , uniformly in  $y$ . Therefore every  $x$  admits a neighborhood  $U$  and a constant  $C \geq 0$  such that  $\nabla_{xx}^2 c(x, y) \leq C \cdot \text{Id}$  uniformly with respect to  $y$  throughout  $U$ . This implies  $x \mapsto c(x, y)$  is locally semiconcave function on  $X$ , uniformly in  $y$ . Therefore the cross-differences  $x \mapsto c_\Delta(x; y, y')$  are locally DC-functions, uniformly in  $y, y'$  in  $Y$ . This observation and equation (26) implies  $Z'(x)$  is locally-DC subvariety of  $X$ .  $\square$

To illustrate, consider the Euclidean quadratic cost  $c(x, y) = \|x - y\|^2/2$  on  $X = \mathbb{R}^N$ . The cross-difference  $c_\Delta(x, y, y')$  is an affine function of  $x$ , namely

$$c_\Delta(x, y, y') = \langle x, y' - y \rangle + \|y\|^2/2 - \|y'\|^2/2.$$

Therefore the cells  $Z'(x)$  are locally affine subsets of  $\mathbb{R}^N$  for every singular point  $x$ .

We caution the reader that Lemma 8 does not specify the dimension of  $Z'(x)$  – the dimension of these cells is addressed below [3](#). Again  $Z'(x)$  is viewed as a local cell on the active domain  $x \in Z_1 = A$ . Our next step is to describe the space of directions  $T_x Z'(x)$ , compare [\[Vil09, Def.10.4.6, pp.257\]](#). The definition of the vector field  $\eta_{avg}(x) \in T_x Z'$  involves the orthogonal projection  $pr_{Z'} : T_x X \rightarrow T_x Z'$  of  $T_x X$  onto the space of directions.

**Lemma 9.** *Let  $x \in X$  be a singular point where  $\partial^c \psi^c(x)$  is not a singleton. Then the space of directions  $T_x Z'(x)$  is a subset of the orthogonal complement*

$$\text{orthog}[\{\nabla_x c_\Delta(x; y_0, y_1) \mid y_1, y_0 \in \partial^c \psi^c(x)\}]$$

*in  $T_x X$ .*

*Proof.* Let  $\phi = \psi^c$ . For  $y_0, y_1 \in Y$ , we abbreviate  $A(x, y_0, y_1) = c_\Delta(x, y, y_0) - \psi(y) + \psi(y_0)$ . If  $x$  is a singular point, then  $Z'(x)$  is characterized by the vanishing  $A(\bar{x}, y, y') = 0$  for  $y, y' \in \partial^c \phi(x)$ . A first-order deformation  $\eta$  at  $x$  and tangent to  $Z'(x)$  must preserve the system of equations  $\{A(x, y, y') = 0 \mid y, y' \in \partial^c \phi(x)\}$ . But this only if  $\eta \in T_x X$  satisfies the homogeneous linear equations  $\eta \cdot \nabla_x c_\Delta(x, y, y') = 0$  for every  $y, y' \in \partial^c \phi(x)$ .  $\square$

Lemma 9 indicates the *expected* Hausdorff dimension of  $Z'(x)$ , namely the dimension of the orthogonal complement  $\{\nabla_x c_\Delta(x, y, y') \mid y, y' \in \partial^c \psi^c(x)\}$ . A standard application of Clarke's Nonsmooth Implicit Function theorem confirms this dimension estimate. Some preliminary notation is convenient. Recall  $X$  is a Riemannian manifold-with-corners, equipped with Riemannian exponential mapping  $\exp_x : T_x X \rightarrow X$ . If  $(X, d)$  is complete Cartan-Hadamard manifold, then  $\exp_x$  is diffeomorphism between  $T_x X$  and the universal covering space  $\tilde{X}$ . On general Riemannian manifold-with-corners  $X$  the exponential map is a local diffeomorphism between sufficiently small open neighborhoods  $U$  of  $x_0 \in X$  with open balls  $U'$  in the tangent space  $T_{x_0} X$ . So we have  $C^1$  diffeomorphisms between local neighborhoods of  $x_0 \in X$  with neighborhoods of 0 in euclidean space  $\mathbb{R}^n$ , with  $n = \dim(X) = \dim(T_{x_0} X)$ . Thus for every  $x_0 \in X$  there exists a local diffeomorphism splitting  $B_\epsilon(x_0)$  as a product  $B_\epsilon(x'_0) \times B_\epsilon(x''_0)$ , viewed as subset of  $\mathbb{R}^{n-j} \times \mathbb{R}^j$  with  $x_0 = (x'_0, x''_0)$  for  $j \geq 0$ .

Recall  $c$  satisfies Assumptions (A0)–(A4). Let  $\psi : Y \rightarrow \mathbb{R} \cup \{-\infty\}$  be a  $c$ -concave potential, with  $\phi = \psi^c$ . Let  $Z_1 := A$  be the active domain of the semicoupling defined by  $\psi$ , and let  $x_0 \in Z_1$ .

**Proposition 3.** *Under the above hypotheses, let  $x_0$  be a singular point. Suppose the gradients  $\{\nabla_x c_\Delta(x; y_0, y_1) \mid y_0, y_1 \in \partial^c \psi^c(x_0)\}$  span a  $j$ -dimensional subspace of  $T_x X$  for every  $x \in Z'(x_0)$ . Then  $Z'(x_0)$  is a codimension- $j$  local DC-subvariety of the active domain  $Z_1$ .*

*Proof.* Fix some  $y_0 \in \partial \phi(x_0)$ . Abbreviate  $A(x, y) := c(x, y) - c(x, y_0)$ . Consider the mapping  $G : B_\epsilon(x_0) \rightarrow \mathbb{R}^d$  defined by

$$G(x) = (A(x, y_1), A(x, y_2), \dots, A(x, y_j))$$

for  $y_0, y_1, \dots, y_j \in \partial^c \psi^c(x_0)$ . Assume

$$\{\nabla_x A(x_0, y_i, y_0)\}_{1 \leq i \leq j}$$

is a linearly independent subset of  $T_{x_0}X$ . The map  $G : B_\epsilon(x_0) \rightarrow \mathbb{R}^j$  is local DC-function according to Lemma 8.

We use exponential mapping to obtain local  $C^1$ -diffeomorphism between an open neighborhood of  $x_0$  in  $X$ . Next we apply the DC-implicit function theorem as stated in [KM18, Thm 3.8], and conclude there exists  $\epsilon > 0$  and a biLipschitz DC-mapping  $\phi$  from  $B_\epsilon(x'_0) \subset \mathbb{R}^{n-j}$  to  $B_\epsilon(x''_0) \subset \mathbb{R}^j$  such that, for all  $x = (x', x'') \in B_\epsilon(x'_0) \times B_\epsilon(x''_0) \subset \mathbb{R}^{n-j} \times \mathbb{R}^j$  we have  $G(x) = G(x', x'') = 0$  if and only if  $x'' = \phi(x')$ .

The basic subdifferential inequalities (26) imply  $G(x) = 0$  if and only if  $x \in Z'(x_0) \cap B(x_0, \epsilon)$ . Now because  $Z'(x_0)$  can be covered by countably many sufficiently small open balls, we conclude that  $Z'(x_0)$  is a local DC-subvariety with Hausdorff dimension  $\dim_{\mathcal{H}} Z'(x_0) = n - j$ .  $\square$

The above 3 is a type of “constant-rank” theorem for Lipschitz maps, c.f. [Cla76].

N.B. The description of  $T_x Z'(x)$  is symmetric with respect to  $y_1, y_0$ . There is further symmetry from the additive relations between cross-costs

$$c_\Delta(x; y_0, y_1) + c_\Delta(x; y_1, y_2) = c_\Delta(x; y_0, y_2).$$

This implies the obvious estimate  $\text{codim} Z'(x) \leq \#(\partial^c \psi^c(x)) - 1$  under general conditions.

Recently [KM18] obtained an explicit parameterization of singularities arising from Euclidean quadratic costs, employing a hypothesis of affine independence between the connected components of the subdifferentials  $\partial^c \psi^c(x)$ . Their parameterization requires a global splitting  $X = X_0 \times X_1$  of the source domain to express singularities (“tears” in their terminology) as the graphs of DC-functions  $G : X_0 \rightarrow X_1$  as above. Further hypotheses on the convexity of source and target are required for their arguments.

## 12. THE DESCENDING FILTRATION $\{Z_j\}_{j \geq 0}$

Following an idea of Prof. D. Bar-Natan [Bar02] we “skewer the cube”  $2^X$  (more accurately the support  $\text{spt}(Z)$ ) according to local codimensions and obtain a descending filtration. The previous section described various cells  $Z'(x)$  which decompose the activated source  $A$  of optimal semicouplings. Assembling these cells into the Kantorovich functor  $Z : 2^Y \rightarrow 2^X$  leads to a useful filtration of the source  $X$ .

**Definition 13.** For integers  $j = 0, 1, 2, \dots$ , let

$$Z_{j+1} := \{x \in X \mid \dim[\text{span}\{\nabla_x c_\Delta(x; y, y') \mid y, y' \in \partial^c \psi^c(x)\}] \geq j\}$$

where  $Z'(x) = Z(\partial^c \psi^c(x))$  for a  $c$ -concave potential  $\psi^{cc} = \psi$  on target space  $Y$ .

According to the definition,  $Z_j$  is supported on the subcategory  $\text{spt}(Z)$  of  $2^Y$  for every  $j \geq 1$ . The activated support  $A$  coincides with the support of  $Z_1$ . The cells  $Z(Y_I)$  are closed (1). If the support  $\text{spt}(Z)$  of the functor is finite, then the unions

$Z_{j+1}$  are closed. However if  $\text{spt}(Z)$  is infinite, then the union  $Z_2, Z_3, \dots$  is possibly not closed and depends on the cost  $c$ . For example, the Euclidean quadratic cost  $c = d^2/2$  typically has  $Z_2$  not closed when the target  $Y$  has  $\dim(Y) > 0$ . This motivates our study of repulsion costs in [Mar, Ch.4].

Now we replace the  $c$ -subdifferentials with an important “localized” version using (Twist) condition (A4). The  $c$ -subdifferential  $\partial^c \psi^c$  of a  $c$ -convex function  $\psi^c : X \rightarrow \mathbb{R} \cup \{+\infty\}$  is non-local subset of  $Y$ . The subsets  $\partial^c \psi^c \subset Y$  depends on global data, namely the values of  $c(x, y)$  for all  $y \in \text{dom}(c_x)$ . Hence  $\partial^c \psi^c(x_0)$  depends on the values of  $\psi(y)$  and  $\psi^c(x)$  for every  $y \in Y, x \in X$  and not simply the local behaviour of  $\psi^c$  near  $x_0$ . It is useful to introduce a local subdifferential, namely the so-called “subgradients” of a function  $\psi^c : X \rightarrow \mathbb{R} \cup \{+\infty\}$ .

**Definition 14** (Local Subgradients  $\partial_\bullet \phi$ ). Let  $U$  be open set in  $X$ , sufficiently small such  $U$  is  $C^1$ -diffeomorphic to an open subset of  $\mathbb{R}^n$  where  $n = \dim(X)$ . Let  $\phi : U \rightarrow \mathbb{R}$  be a function. Then  $\phi$  is subdifferentiable at  $x$  with subgradient  $v^* \in T_x^* X$  if

$$\phi(z) \geq \phi(x) + v^*(z - x) + o(|z - x|) \text{ for all } z \text{ near } x.$$

Let  $\partial_\bullet \phi(x) \subset T_x^* X$  denote the set of all subgradients to  $\phi$  at  $x$ . Here  $o$  is the “little-oh” notation.

Evidently the subgradient  $\partial_\bullet \phi(x)$  is local, and depending on the values of  $\phi$  near  $x$ . Moreover  $\partial_\bullet \phi(x)$  is a closed convex subset of  $T_x^* X$  for every  $x \in \text{dom}(\phi)$ .

The Assumptions (A0)–(A3) imply  $\psi^c : X \rightarrow \mathbb{R} \cup \{+\infty\}$  is locally semiconvex over its domain  $\text{dom}(\psi^c) \subset X$  (Lemma 2). In otherwords, every  $x \in \text{dom}(\psi^c)$  admits an open neighborhood  $U$  of  $x$  and a constant  $C \geq 0$  such that  $D^2 \psi^c \geq -C \text{Id}$  throughout  $U$ , where  $D^2 \psi^c = D_{xx}^2 \psi^c$  is the distributional Hessian in the source variable  $x$  ([Vil09, Theorem 14.1, pp.363]). So given a  $c$ -convex potential  $\psi^c$  on  $X$ , for every  $x_0 \in X$  there exists open neighborhood  $U$  and  $C \geq 0$  such that  $\psi^c|_U + C||x||^2/2$  is strictly convex throughout  $U$ . This implies the local subdifferential and subgradients  $\partial_\bullet \psi^c(x)$  coincide with the convex-analytic subdifferential of the local function  $\psi^c|_U + C||x||^2/2$ , when restricted to the sufficiently small neighborhood  $x$  in  $X$ .

Having introduced the local subdifferential, there is important comparison between  $\partial^c \psi^c(x_0) \subset Y$  and  $\partial_\bullet \psi^c(x_0)$ , assuming (Twist) condition Assumption (A4). This relation is the inclusion

$$(27) \quad \partial^c \psi^c(x_0) \subset \{y \in Y \mid -\nabla_x c(x_0, y) \in \partial_\bullet \psi^c(x_0)\}.$$

Observe  $\partial^c \psi^c(x_0)$  is nonempty whenever  $x_0 \in \text{dom}(\psi^c)$ . We abuse notation and denote  $\nabla_x c(x, y)$  for the canonical covector in  $T_x^* X$ , with the tangent vector in  $T_x X$  using the ambient Riemannian structure. The inclusion (27) allows us to replace the global  $c$ -subdifferential  $\partial^c \psi^c$  with the local convex set of subgradients  $\partial_\bullet \psi^c$ . The inclusion is generally strict. However it produces basic upper bounds on the Hausdorff dimension of  $\partial^c \psi^c$ . We quote the following theorem of G. Alberti:

**Theorem 3** ([Alb94]). *Let  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  be proper lowersemicontinuous convex function. For every  $0 < k < n$ , let  $S^k(f) \subset \mathbb{R}^n$  be the subset defined by*

$$S^k(f) := \{x \in \mathbb{R}^n \mid \dim_{\mathcal{H}}(\partial_{\bullet}(f)) \geq k\}.$$

*Then  $S^k(f)$  can be covered by countably many  $(n - k)$ -dimensional DC-manifolds.*

In otherwords  $S^k(f)$  is countably  $(n - k)$ -rectifiable and has Hausdorff dimension  $\leq (n - k)$ . If  $f$  is  $+\infty$ -valued on  $\mathbb{R}^n$ , then Alberti's method shows  $S^k(f)$  can be covered by countably many  $(n' - k)$ -dimensional manifolds where  $n' = \dim_{\mathcal{H}}(\text{dom}(f))$ . The domain  $\text{dom}(f)$  is a closed convex subset of  $\mathbb{R}^n$  having well-defined Hausdorff dimension.

**Proposition 4.** *Let  $c : X \times Y \rightarrow \mathbb{R} \cup \{+\infty\}$  be cost satisfying Assumptions (A0)–(A4). Let  $\psi : Y \rightarrow \mathbb{R} \cup \{-\infty\}$  be  $c$ -concave potential  $\psi^{cc} = \psi$  with singularity functor  $Z : 2^Y \rightarrow 2^X$ . Define  $n' := \dim_{\mathcal{H}}(\text{dom}(\psi^c))$ . Then for every integer  $j \geq 1$ , the subvariety  $Z_j$  has Hausdorff dimension*

$$\dim_{\mathcal{H}}(Z_j) \leq n' - j + 1.$$

*Proof.* Consider the inclusion (27). For given  $x_0$ , let  $U$  be the open neighborhood and  $C > 0$  such that  $\psi^c|_U + C||x||^2/2$  is strictly convex throughout  $U$ . For  $x \in U$  the  $c$ -subdifferentials  $\partial^c \psi^c(x)$  are contained in the closed convex local subdifferentials  $\partial_{\bullet} \psi^c|_U(x)$ . Now we have equality

$$(28) \quad \partial_{\bullet}(\psi^c|_U(x) + C|x|^2/2) = \partial_{\bullet} \psi^c|_U(x) + C\langle -, x \rangle.$$

So the local subdifferential  $\partial_{\bullet}(\psi^c|_U + C|x|^2/2)$  in  $T_x^*X$  is an affine translate of  $\partial_{\bullet} \psi^c|_U$  by the linear functional  $C\langle -, x \rangle$ .

Next we apply Alberti's theorem to the localized convex function  $\psi^c|_U + C|x|^2/2$ . Thus  $S^k(\psi^c|_U + C|x|^2/2)$  and  $S^k(\psi^c|_U)$  can be covered by countably many  $(n' - k)$ -dimensional manifolds where  $n' := \dim_{\mathcal{H}}(\text{dom} \psi^c|_U)$ .

Finally we relate  $\dim_{\mathcal{H}}(\partial_{\bullet} \psi^c|_U(x))$  to the Hausdorff dimension of  $Z'(x) = Z(\partial^c \psi^c(x)) = \cap_{y \in \partial^c \psi^c(x)} \partial^c \psi(y)$ . From the definition of subgradients, we have

$$\text{conv}(\{\nabla_x c(x, y) \mid y \in \partial^c \psi^c(x)\}) \subset \partial_{\bullet} \psi^c(x).$$

Moreover the closed convex hull  $\text{conv}(\{\nabla_x c(x, y) \mid y \in \partial^c \psi^c(x)\})$  has dimension

$$j = \dim(\text{span}\{\nabla_x c_{\Delta}(x, y, y_0) \mid y \in \partial^c \psi^c(x)\}).$$

For every  $x \in Z_j$  we conclude  $Z_j \cap U \subset S^j(\psi^c|_U + C|x|^2/2)$ , which according to Alberti's theorem 3 yields the upper bound  $\dim_{\mathcal{H}}(Z_j \cap U) \leq n' - j$ . To conclude, we observe that  $Z_j$  can be covered by countably many open neighborhoods  $U$  of points  $x \in Z_j$ .

□

We supplement 4 with the following application of Clarke's Implicit Function Theorem, which gives a criterion for the singularities  $Z_2, Z_3, \dots$  to be *closed* subsets of  $X$ .

**Proposition 5.** *Let  $c : X \times Y \rightarrow \mathbb{R}$  be a cost satisfying Assumptions (A0)–(A4). Let  $\phi = \psi^c$  be a  $c$ -convex potential on  $X$ . Suppose the sup-norm of the gradients  $\nabla_x c_\Delta(x, y_0, y_1)$ ,  $y_0, y_1 \in \partial^c \phi(x)$  is nonzero, uniformly with respect to  $x \in Z_2$ . Then  $Z_2$  is a closed subset of  $X$ .*

*Proof.* Consequence of Clarke's Implicit Function Theorem, c.f. [Vil09, Proof of Theorem 10.50, pp.262–264].  $\square$

The above proposition 5 can be generalized to the following criterion for  $Z_j$  to be a closed subset. If  $I$  is a collection of vectors, then  $\text{span}_{\mathbb{Z}}(I)$  is the group of all finite  $\mathbb{Z}$ -linear combinations  $\sum_{v \in I} n_v v$ , for  $n_v \in \mathbb{Z}$ . We define the “height” of  $I$ , denoted  $ht(I)$ , to be the *volume* of the quotient  $\text{span}(I)/\text{span}_{\mathbb{Z}}(I)$  with respect to a fixed Lebesgue measure.

**Proposition 6.** *Let  $c : X \times Y \rightarrow \mathbb{R}$  be a cost satisfying Assumptions (A0)–(A4). Let  $\phi = \psi^c$  be a  $c$ -convex potential on  $X$ . Fix  $j \geq 1$ . For every  $x \in Z_j$ , let  $I_x$  be the collection of gradients*

$$I_x := \{\nabla_x c_\Delta(x, y_0, y_1) \mid y_0, y_1 \in \partial^c \phi(x)\}.$$

*Suppose the heights  $ht(I_x)$  are bounded away from zero, uniformly with respect to  $x \in Z_j$ . Then  $Z_j$  is a closed subset of  $X$ .*

*Proof.* We leave the proof to the reader, being again a consequence of Clarke's Implicit Function Theorem.  $\square$

### 13. PROOF OF THEOREM 2

This section generalizes the homotopy reduction constructed in 4 and the proof of 1, and establishes a retraction procedure  $Z_1 \rightsquigarrow Z_2 \rightsquigarrow \dots \rightsquigarrow Z_{J+1}$ , defined up to some maximal index  $J \geq 1$  for which the inclusions

$$Z_1 \hookleftarrow Z_2 \hookleftarrow \dots \hookleftarrow Z_{J+1}$$

are simultaneously homotopy isomorphisms. These retractions require our cost  $c$  satisfy Assumptions (A0)–(A5), a specific form of (A6), the  $c$ -subdifferentials have finite cardinality and the Uniform Halfspace condition 15. The retract is obtained by integrating a vector field, denoted  $\eta_{\text{avg}}(x)$ , whose flow under the above assumptions yields continuous deformation retracts (Theorem 5).

An important hypothesis for constructing our retractions is the (UHS) condition. Some notation is necessary. Abbreviate  $Y'(x) := \text{dom}(c_x)$ . Let  $Z = Z(c, \sigma, \tau)$  be the

corresponding Kantorovich functor. For  $x \in Z'$ , let  $pr_{Z'} : T_x X \rightarrow T_x Z'$  denote the orthogonal-projection mapping, where  $T_x Z'$  is the space of directions of  $Z'$  at  $x$ .

**Definition 15.** Let  $\phi = \psi^c$  be the  $c$ -convex potential. For  $x_0 \in \text{dom}(\psi^c)$  abbreviate  $Z' := Z(\partial^c \psi^c(x_0))$ . Select any  $y_0 \in \partial^c \psi^c(x_0)$ . Recall the definition of  $\nu_x(y)$  from 7.

For parameter  $\beta > 2$ ,  $x \in Z'$ ,  $y \in Y$ , define tangent vectors  $\eta(x, y) \in T_x X$  by the equation

$$(29) \quad \eta(x, y) := |\psi(y_0) - \psi(y) + c_\Delta(x; y, y_0)|^{-\beta} \cdot pr_{Z'}(\nabla_x c_\Delta(x; y, y_0)).$$

We say Uniform Halfspace (UHS) conditions are satisfied at  $x$  in  $Z'$  with respect to the parameter  $\beta$  if:

(UHS1) the Bochner integral  $\eta_{avg}(x)$  defined by

$$(30) \quad \eta_{avg}(x) := (\mathcal{H}_Y[Y])^{-1} \int_Y \eta(x, y) \cdot \nu_x^\beta(y) \cdot d\mathcal{H}_Y(y)$$

is nonzero finite vector in  $T_x Z' - \{0\}$ ; and

(UHS2) there exists a constant  $C > 0$ , uniform with  $x \in Z'$ , for which the estimate

$$(31) \quad \|\eta_{avg}(x)\| \geq C \int_Y \|\eta(x, y) \cdot \nu_x^\beta(y)\| d\mathcal{H}_Y(y) > 0$$

holds.

We make some remarks. First, we observe (UHS2) basically implies (UHS1). Second, the definition of  $\eta(x, y)$  is independent of the choice of  $y_0 \in \partial^c \psi^c(x')$ . Third, in practice the parameter  $\beta > 2$  is taken sufficiently large to ensure the divergence of  $\eta_{avg}(x)$  whenever  $x$  converges into  $\partial Z'$ . If  $c$  satisfies Assumption (A5), then  $\psi(y)$  is locally Lipschitz (Prop.1) and the exponent  $\beta = \dim(Y) + 2$  is sufficient. Finally, we observe that symmetry improves (UHS) conditions, i.e. if  $\Gamma$  acts isometrically on  $X \times Y$ , then  $c$ -optimal semicouplings will also be  $\Gamma$ -equivariant (whenever  $\sigma, \tau$  are).

Now we state our theorem.

**Theorem 4** (Local Cellular Retraction  $Z' \cap Z_j \rightsquigarrow Z' \cap Z_{j+1}$ ). *Suppose  $\partial^c \psi^c(x)$  is finite for every  $x \in \text{dom}(\psi^c)$ . Abbreviate  $Z' := Z'(x) = Z(\partial^c \psi^c(x))$ . If (UHS) Conditions are satisfied for all points  $x' \in Z' \cap Z_j$  (15), and  $Z' \cap Z_{j+1} \neq \emptyset$  is nonempty, then there exists a continuous map*

$$\Psi : (Z' \cap Z_j) \times [0, 1] \rightarrow Z' \cap Z_j$$

such that:

- (i)  $\Psi(x, s) = x$  for all  $x \in Z' \cap Z_{j+1}$ ; and
- (ii)  $\Psi(x, 0) = x$  for all  $x \in Z'$ ; and
- (iii)  $\Psi(x'', 1) \in Z' \cap Z_{j+1}$  for all  $x'' \in Z' \cap Z_j$ .



In addition, if the above hypotheses are satisfied, then  $Z' \cap Z_{j+1}$  is a strong deformation retract of  $Z' \cap Z_j$  and the inclusion

$$Z' \cap Z_{j+1} \hookrightarrow Z' \cap Z_j$$

is a homotopy isomorphism.

The vector field  $\eta_{avg}(x)$  defined in Theorem 4 is tangent to the cell  $Z'(x)$ , and therefore the flow is constrained to the cells  $Z'(x)$ . For  $x$  varying over  $Z_j$ , the vector field  $\eta_{avg}(x)$  varies continuously with respect to  $x$  according to Lemma 1. The hypothesis  $Z' \cap Z_{j+1} \neq \emptyset$  cannot be ignored, and is definitely necessary for continuity of the gradient flow of  $\eta_{avg}$ .

Next we claim the mappings

$$\{\Psi : Z'(x) \cap Z_j \rightarrow Z'(x) \cap Z_j\}$$

assemble to a continuous retraction  $Z_j \times [0, 1] \rightarrow Z_j$ , for  $x \in Z_j$ , which establishes that  $Z_j \hookleftarrow Z_{j+1}$  is a homotopy isomorphism.

**Theorem 5** (Global Retraction). *Let  $c, \sigma, \tau, \rho, Z, Z', \psi, \phi$  be defined as in Theorem 4. If  $j \geq 1$  is an integer such that (UHS) Conditions are satisfied for all points  $x \in Z_j$ , and  $Z'(x) \cap Z_{j+1} \neq \emptyset$  for every  $x \in Z_j - Z_{j+1}$ , then the local homotopy equivalences*

$$\{\Psi : (Z' \cap Z_j) \times [0, 1] \rightarrow Z' \cap Z_{j+1}\}$$

*constructed in Theorem 4 assemble to a continuous global deformation retract  $Z_j \rightsquigarrow Z_{j+1}$ .*

*In addition, if  $J \geq 1$  is that maximal integer where (UHS) conditions are satisfied throughout  $Z_J$ , then composing the retractions  $\{\Psi\}$  produces a codimension- $J$  homotopy isomorphism  $Z_1 \simeq Z_{J+1}$ .*

The proof of 5 is analogous to our proof of Theorem 4, which is the “base case” retraction of  $X = Z_0$  onto the activated domain  $A = Z_1$ . The formal proof of 4 required Lemmas 6, 7. Likewise the formal proofs of 5 require analogous lemmas, summarized in:

**Lemma 10.** *In the notations of 4–5 be satisfied. Let  $j \geq 1$  be such that (UHS) conditions are satisfied for all  $x \in Z_j$ , and  $Z'(x) \cap Z_{j+1} \neq \emptyset$  for every  $x \in Z_j - Z_{j+1}$ , and  $\psi^c$  is separated.*

*For  $x_0 \in Z_j - Z_{j+1}$ , consider the initial value problem*

$$(32) \quad x' = \eta_{avg}(x), \quad x(0) = x_0.$$

*(a) For every initial value  $x_0 \in Z_j - Z_{j+1}$ , the flow  $\Psi$  defined by (32) diverges to infinity in finite time.*

*(b) The maximum interval of existence  $I(x_0) := [0, \omega(x_0))$  of solutions to (32) varies continuously with respect to the initial value  $x_0 \in Z_j - Z_{j+1}$ .*



*Proof.* We follow the arguments of 6. First (UHS) implies the flow (32) is well-defined and extendable throughout the interior of the cells  $Z'(x_0) \subset Z_j$ . At the  $\omega$ -limit point

$$\bar{x} := \lim_{t \rightarrow \omega(x_0)^-} x(t)$$

we claim  $\bar{x} \in Z'(x_0) \cap Z_{j+1}$ . The key point is to verify  $\bar{x} \in Z_{j+1}$ . At the  $\omega$ -limit point, we find

$$\partial^c \phi(\bar{x}) \supset \{\bar{y}\} \coprod \partial^c \phi(x_0),$$

and the subdifferential at  $\bar{x}$  contains at least one new target point  $\bar{y} \in Y$ . The (UHS) condition implies the average of the projections  $\nabla_x c_\Delta(\bar{x}, y_0, \bar{y})$  (with respect to  $\bar{\nu}$ ) is a nonzero vector of  $T_{\bar{x}} Z'$ , and linearly independent from the nonzero tangent vectors  $\nabla_x c_\Delta(\bar{x}, y_1, y_0)$ ,  $y_0, y_1 \in \partial^c \phi(x_0)$  (which are orthogonal to  $T_{\bar{x}} Z'$ ). This proves  $\bar{x} \in Z_{j+1}$ . This proves (a).

Next we claim the flow defined by  $\Psi$  is asymptotically contracting, and therefore  $\omega$  is a Lipschitz continuous function. The same arguments from 4, 5 shows the vector field  $\eta_{avg}$  is the gradient of an averaged potential  $f_{avg}$ , and this potential is asymptotically convex in the direction of  $\nabla_x f_{avg}$ . This proves (b).  $\square$

*Sketch of proof for Theorem 4.* We follow the proof of Theorem 4, and construct a continuous vector field  $\eta_{avg}(x')$  on  $Z' \cap Z_j$  which blows-up precisely on  $Z' \cap Z_{j+1} \subset Z' \cap Z_j$ , which we assume is nonempty. For initial points  $x'$ , the maximal interval of existence  $I(x') = [0, \omega(x'))$  varies continuously with respect to  $x'$  (see Lemma 10). The field  $\eta_{avg}(x')$  will generate a global forward-time continuous mapping

$$\Psi : (Z' \cap Z_j) \times I \rightarrow Z' \cap Z_j$$

satisfying the usual ordinary differential equation  $\frac{d}{ds}[\Psi(x', s')] = \eta_{avg}(\Psi(x', s'))$  for all  $s' \in I(x')$ . The flow  $\Psi$  is reparameterized according to the parameter  $\omega(x_0)$  to obtain a continuous deformation retract  $\Psi' : (Z' \cap Z_j) \times [0, 1] \rightarrow Z' \cap Z_j$  as desired.  $\square$

*Proof of Theorem 4.* The (UHS) conditions ensure the cross-differences  $c_\Delta$  have non-vanishing gradient  $\nabla_x c_\Delta \neq 0$  throughout the domain  $Z'$ , and the gradients  $\nabla_x c_\Delta$  vary continuously over  $Z_j$ . Moreover the uniform Halfspace condition (UHS2) ensures  $\|\eta_{avg}(x)\|$  is uniformly bounded away from zero in neighborhoods of the poles  $\{\|\eta_{avg}(x)\| = +\infty\} = Z' \cap Z_{j+1}$  in  $Z' \cap Z_j$ .

The vector field  $\eta_{avg}(x')$  integrates to a global forward-time flow  $\Psi : Z' \cap Z_j \times I \rightarrow Z' \cap Z_j$ , where as usual we have  $d/ds[\Psi(x', s')] = \eta_{avg}(\Psi(x', s'))$  for all  $s' \in I(x')$ . According to Lemma 10, for every choice of  $x = x(0)$  initial value on  $Z'(x) \cap Z_j$ , the trajectory  $x = x(t)$  converges in finite time to  $Z' \cap Z_{j+1}$  with respect to the flow (20). Our Assumptions (A0)–(A6) imply continuous dependence on the choice of initial value. The argument from Proposition 5 with Lemma 10 prove the flow

$\Psi$  can be reparameterized according to the parameter  $\omega(x_0)$  to obtain a continuous deformation retract  $\Psi'$  of  $Z' \cap Z_j$  onto  $Z' \cap Z_j$ , as desired.  $\square$

Finally we observe the “local” homotopy isomorphisms  $\{\Psi : Z'(x) \cap Z_j \rightsquigarrow Z'(x) \cap Z_{j+1}\}$  assemble to a continuous deformation retract of  $Z_j \simeq Z_{j+1}$  and establish Theorem 5:

*Proof of Theorem 5.* The proof of Theorem 4 constructs homotopy deformations  $Z'(x) \cap Z_j \rightarrow Z'(x) \cap Z_{j+1}$ , where  $Z'(x) = Z(\partial\psi^c(x)) \subset Z_1$  for a given  $x$  in  $Z_j$ . When  $x$  varies over  $Z_j$ , the corresponding  $Z' = Z'(x)$  are either coincident or intersect along a subset of  $Z_{j+1}$ . So the local strong deformation retracts  $\{\Psi\}$  constructed in Theorem 5 assemble to a continuous deformation retract  $Z_j \times [0, 1] \rightarrow Z_j$ . Therefore  $Z_j \hookrightarrow Z_{j+1}$  is a homotopy isomorphism. Composing the retractions  $Z_j \rightsquigarrow Z_{j+1}$  for  $j = 1, \dots, J$  yields the deformation retract  $Z_1 \rightsquigarrow Z_{J+1}$ , as desired.  $\square$

The above results establish Theorem 2 from the Introduction.

#### 14. BRIEF REVIEW OF SINGULARITY IN THE LITERATURE

The term “singularity” is evidently overburdened having various interpretations within the literature. We use the term “singularity” as arising in convex geometry: if  $\phi$  is a lower semicontinuous convex function on  $\mathbb{R}^N$ , then the “singularity” of  $\phi$  corresponds to the locus of non-differentiability on  $\text{dom}(\phi)$ , i.e. the set of points  $x \in \text{dom}(\phi)$  where  $\partial\phi(x) \subset (\mathbb{R}^N)^*$  is not single-valued. From the viewpoint of optimal transport, “singularity” then refers to the locus-of-discontinuity of  $c$ -optimal semicouplings.

The Kantorovich singularity (Def. 10) is both categorical and topological, and provides an alternative view to the so-called “regularity theory” of optimal transportation. Regularity in optimal transport usually refers to the Monge map  $y = T(x)$  defined by the (Twist) condition, and

$$T(x') := \nabla_x c(x', -)^{-1}(-\nabla_x \psi^c(x')).$$

There is large volume of research concerning the  $C^{1,\alpha}$  or  $C^2$ ,  $C^\infty$  regularity of  $T$  under various hypotheses on  $c, \sigma, \tau$ . C.f. [Vil09, Ch.12]. Our thesis however is strictly interested in the *discontinuities* of  $T$ . So we pass silently over questions of the type “How regular is the map  $T$  away from the singularities?”. This thesis rather studies the continuity properties of the locus-of-discontinuity of  $T$ .

Several results concerning singularities of optimal transports have been attained in the literature. Alberti’s paper [Alb94] is basic starting point. A. Figalli [Fig10] studies the singularities of optimal transports between two probability measures supported on bounded open domains in the plane  $\mathbb{R}^2$  with respect to the quadratic Euclidean

cost  $c(x, y) = \|x - y\|^2/2$  (equivalently  $c = -\langle x, y \rangle$ ). The main result of [Fig10, §3.2] is that the singularity ( $Z_2$  in our notation) has topological closure  $\overline{Z_2}$  in  $\mathbb{R}^2$  with zero two-dimensional Hausdorff measure,  $\mathcal{H}^2(\overline{Z_2}) = 0$ . (Here we remark that the closure  $\overline{Z_2}$  is not easily topologized from the singularity  $Z_2$ . In other words the closure is not explicit). Figalli's work was extended in [FK10], where the determination  $\mathcal{H}^n(\overline{Z_2}) = 0$  was established for singularities of optimal couplings under the hypothesis that probability measures are supported on bounded open domains of  $\mathbb{R}^n$ , and again with respect to Euclidean quadratic cost. In [PF] a similar result is established with respect to more general costs on  $\mathbb{R}^n$  satisfying basic nondegeneracy conditions, namely:

- the cost  $c \in C^2(X \times Y)$  with  $\|c\|_{C^2} < +\infty$ ,
- the rule  $x \mapsto \nabla_y c(x, y)$  is injective map for every  $y$ ,
- the rule  $y \mapsto \nabla_x c(x, y)$  is injective map for every  $x$ , and
- $\det(D_{xy}^2 c) \neq 0$  for all  $(x, y)$ .

The investigations of Figalli suggest the following: if the source and target measures  $\sigma, \tau$  are absolutely continuous to  $\mathcal{H}_X, \mathcal{H}_Y$ , and the cost  $c$  satisfies (A0)–(A4), then the singularity  $Z_2$  has Hausdorff dimension satisfying

$$\dim_{\mathcal{H}}(Z_2) \leq \dim(Y).$$

We motivate this dimension formula in 4. The recent work [KM18] confirms this estimate under some particular conditions, namely Euclidean quadratic cost, convex source, and convexity and affine-independence of the disjoint target components.

The goal of the present thesis is to establish topological properties of the singularities of optimal transports. This problem has not appeared in the literature, and we believe it is important area of study, and especially for applications to various topological problems (e.g. explicit souls and spines).

## REFERENCES

- [Alb94] G. Alberti. “On the structure of singular sets of convex functions”. In: *Calculus of Variations and Partial Differential Equations* 2.1 (1994), pp. 17–27. URL: <https://doi.org/10.1007/BF01234313>.
- [Bar02] Dror Bar-Natan. “On Khovanov’s categorification of the Jones polynomial.” In: *Algebraic & Geometric Topology* 2 (2002), pp. 337–370. URL: <http://eudml.org/doc/122177>.
- [Cla76] Francis Clarke. “On the inverse function theorem”. In: *Pacific Journal of Mathematics* 64.1 (1976), pp. 97–102.
- [ET99] I. Ekeland and R. Témam. *Convex Analysis and Variational Problems*. Classics in Applied Mathematics, Vol. 28. SIAM, 1999.

- [Fig10] A. Figalli. “Regularity properties of optimal maps between nonconvex domains in the plane”. In: *Comm. Partial Differential Equations* 35 (2010), no. 3 (2010), pp. 465–479.
- [FK10] A. Figalli and Y.-H. Kim. “Partial regularity of Brenier solutions of the Monge-Ampère equation”. In: *Discrete Contin. Dyn. Syst.* 28. no.2 (2010), pp. 559–565.
- [Har64] P. Hartman. *Ordinary Differential Equations*. John Wiley & Sons, 1964.
- [KM18] J. Kitagawa and R.-J. McCann. “Free Discontinuities in Optimal Transport”. In: (2018). URL: <http://www.math.toronto.edu/mccann/papers/KitagawaMcCann.pdf>.
- [Mar] J.H. Martel. “Applications of Optimal Transport to Algebraic Topology: How to Build Spines from Singularity”. PhD thesis. University of Toronto. URL: <https://github.com/marvinMLKUltra/thesis/blob/master/ut-thesis.pdf>.
- [PF] G. De Philippis and A. Figalli. “Partial Regularity for optimal transport maps”. In: (). URL: <https://arxiv.org/abs/1209.5640>.
- [Vil03] C. Villani. *Topics in Optimal transportation*. Graduate Studies in Mathematics Vol. 58. Springer-Verlag, 2003.
- [Vil09] C. Villani. *Optimal transport: old and new*. Grundlehren der mathematischen wissenschaften Vol. 338. Springer-Verlag, 2009.

*Email address:* jhmartel@protonmail.com