TWO VIEWS ON OPTIMAL TRANSPORT AND ITS NUMERICAL SOLUTION

MICHAEL LINDSEY

Abstract. We present two insights into certain optimal transportation problems that are accompanied by new approaches to the numerical solution of these problems.

First, we show that Monge's optimal transport problem with quadratic cost can be reformulated as an infinite-dimensional convex optimization problem under certain conditions, most importantly that the target measure has a log-concave density. We define a natural discretization of the problem that can be solved by standard convex optimization methods. We show that under suitable regularity conditions the solutions of these discretized problems converges to the true solution of the optimal transport problem as the discretization becomes increasingly fine.

Second, we exhibit an approach to the numerical solution of optimal transport problems with quadratic cost, essentially only involving the repeated solution of second-order linear elliptic partial differential equations.

We use both methods to numerically compute and obtain visualizations of optimal transport maps. In order to keep this work relatively self-contained, we provide as much as possible of the relevant background of optimal transport and convex analysis.

Contents

Introduction

Optimal transport is a rich field with diverse applications both within and outside of mathematics. The motivating question of optimal transport is the following: how does one move an arrangement of mass to a specified position in a way that minimizes some cost associated with transportation? The solutions to such problems, however, are rarely expressible in closed form, and hence numerical methods for solving optimal transport problems are crucial in applications. For a thorough survey of existing numerical methods we refer the interested reader to the article of Feng–Glowinski–Neilan [15] and the references therein.

In this work we exhibit two novel methods for solving certain optimal transport problems. The first method is accompanied by a proof of accuracy under certain conditions. This proof is surprisingly technical, and a large part of this work is devoted to it. Our second method is justified more heuristically, though it is computationally faster and has empirically proven capable of handling more difficult problems than the first method.

Before discussing these methods, we begin with a review of optimal transport.

Part 1. Background on optimal transport

The theory of optimal transport can be motivated quite concretely. We will give two informal examples of optimal transport problems before defining such problems in generality.

First, suppose that we are given a pile of sand (the shape of which is arbitrary) and a hole of equal volume (also of arbitrary shape) as well as a function $c(x, y)$ specifying the cost of moving mass from point x to point y. How do we move all of the sand into the hole in a way that minimizes the total transportation cost?

Second, suppose that there are N bakeries that together supply a total of B baked goods as well as M restaurants that together have a total demand for B baked goods. Further suppose that there is a cost c_{ij} associated with transporting a single baked good from the i-th bakery to the j-th restaurant. How do we move all of the baked goods from bakeries to restaurants in a way that minimizes the total transportation cost? (For now, we allow for the transportation of non-integer numbers of baked goods.)

1. The Kantorovich problem

Let us now formulate a more general optimal transport problem. We will roughly follow Villani [36] in the following. Since this material will serve as background for the thesis, we shall attempt to avoid excessive formality.

Let X and Y be measure spaces (which we shall call the source and target, respectively), and let μ and ν be probability measures on X and Y, respectively. We want to define an object that fully describes a plan for transporting the mass of μ to ν. This object will be a probability measure π on the product space $X \times Y$. Roughly speaking, we want $d\pi(x, y)$ to specify the amount of mass moved from $x \in X$ to $y \in Y$. Then (again roughly) it must hold that $d\mu(x) = \int_Y d\pi(x, y)$ and $d\nu(y) = \int_X d\pi(x, y)$, i.e., that the total amount of mass transported from x coincide with $d\mu(x)$ and that the total amount of mass transported to y coincide with $d\nu(y)$. This intuition leads us to the following constraint on π: we must have

$$\pi(A \times Y) = \mu(A)$$

for all measurable $A \subset X$ and

$$\pi(X \times B) = \nu(B)$$

for all measurable $B \subset Y$ (so π has marginals μ and ν). A probability measure π on $X \times Y$ satsifying this constraint will be called a transference plan.

Now let $\Pi(\mu, \nu)$ be the set of transference plans associated with μ and ν, and let c (the cost function) be a measurable map $X \times Y \to \mathbb{R} \cup \{+\infty\}$ with $c \geq 0$. Then the Kantorovich optimal transport problem is the following:

$$\text{minimize} \int_{X \times Y} c(x, y)\, d\pi(x, y), \text{ subject to } \pi \in \Pi(\mu, \nu).$$

We are interested in both the minimal cost (the optimal transport cost) and the transference plans π that attain the minimal cost (the optimal transference plans, which may not exist). Notice that it is possible (a priori) that an optimal transference plan 'splits mass.' It makes sense to allow for this possibility; indeed, consider the second example problem above (i.e., the bakery-restaurant problem) in the case that $N < M$. It must be the case that, in any transference plan π, at least one of the bakeries send baked goods to at least two restaurants.

2. The Monge problem

Nonetheless, it is natural to pay special attention to transference plans that do not split mass. Suppose that π does not split mass, i.e., that for any $x \in X$, $d\pi(x,y) = d\mu(x)\,\delta[y = T(x)]$, where T is a function $X \to Y$. In order to proceed, we must additionally assume that T is measurable. Then for any measurable non-negative measurable function $\zeta : X \times Y \to \mathbb{R}$, we have that

$$\int_{X\times Y} \zeta(x,y)\, d\pi(x,y) = \int_X \zeta(x, T(x))\, d\mu(x).$$

In particular, ζ may be characteristic functions of measurable subsets of $X \times Y$. Therefore, given a measurable map $T : X \to Y$, the preceding equation defines a probability measure π_T on $X \times Y$. More precisely, we define π_T by

$$\pi_T(C) = \int_X \chi_C(x, T(x))\, d\mu(x)$$

for measurable $C \subset X \times Y$. We want to know when T yields $\pi_T \in \Pi(\mu, \nu)$. Notice that

$$\pi_T(A \times Y) = \int_X \chi_{A\times Y}(x, T(x))\, d\mu(x) = \mu(A)$$

for measurable $A \subset X$ and

$$\pi_T(X \times B) = \int_X \chi_{X\times B}(x, T(x))\, d\mu(x) = \mu\left(T^{-1}(B)\right)$$

for measurable $B \subset Y$, so in order to have $\pi_T \in \Pi(\mu, \nu)$ we only require that $\nu(B) = \mu\left(T^{-1}(B)\right)$ for all measurable $B \subset Y$. This means precisely that ν is the push-forward measure of μ by T, written $\nu = T_{\#}\mu$.

Now we can formulate the Monge optimal transport problem:

$$\begin{aligned} \text{minimize} \quad & \int_X c(x, T(x))\, d\mu(x) \\ \text{subject to} \quad & T \text{ measurable, } T_{\#}\mu = \nu. \end{aligned}$$

It is clear that Kantorovich's problem is a relaxation of Monge's problem (though Monge's problem preceded Kantorovich's historically). The Kantorovich problem is, in several senses, easier to handle theoretically than the Monge problem. In particular, note that for any μ and ν, while there always exists a transference plan π, there does not always exist T such that $T_{\#}\mu = \nu$. To see the first point, simply note that we can take $d\pi(x,y) = d\mu(x)\, d\nu(y)$. For the second point, consider the case in which μ is a Dirac mass and ν is the combination of two Dirac masses supported at different points.

It is natural to ask when the solutions of the Monge and Kantorovich problems exist and are unique, as well as when the solutions of the two problems coincide (i.e., when T and π_T solve the Monge and Kantorovich problems, respectively). When the two problems can be solved jointly, we will refer to the two problems together as the Monge-Kantorovich problem. This thesis will focus on the setting in which $X = Y = \mathbb{R}^n$, μ and ν are absolutely continuous with respect to the Lebesgue measure, and $c(x, y) = |x-y|^2$. (We shall refer to this cost as the 'quadratic cost.') It happens that in this setting, the Monge and Kantorovich problems admit unique solutions, and their solutions coincide. We will review the results that guarantee this.

It might interest the reader that Gaspard Monge (1746-1818), the originator of the Monge problem, actually focused on the cost function $c(x, y) = |x-y|$ on $\mathbb{R}^n \times \mathbb{R}^n$. This case happens to be much trickier to deal with than the case $c(x, y) = |x - y|^p$, where $p > 1$. In particular, uniqueness results are less readily available. We can illustrate this point with a simple example. Consider the 'bookcase problem,' in which $X = \{-1, 0\}$, $Y = \{0, 1\}$, $2\mu = \delta_{-1} + \delta_0$, $2\nu = \delta_0 + \delta_1$, and $c(x, y) = |x - y|^p$. Then evidently the map $T_1 : X \to Y$ defined by $T_1(-1) = 0$ and $T_1(0) = 1$ and the map T_2 defined by $T_2(-1) = 1$ and $T_2(0) = 0$ both solve the Monge and Kantorovich problems when $p = 1$. However, when $p > 1$, it is clear that T_1 is the unique solution of both problems.

3. The discrete problem

Let $X = \{x_1, \ldots, x_N\}$ and $Y = \{y_1, \ldots, y_M\}$, and let

$$\mu = \sum_{i=1}^{N} a_i \delta_{x_i}, \quad \nu = \sum_{j=1}^{N} b_j \delta_{y_j},$$

be probability measures (so $a_i, b_j \geq 0$ with $\sum_i a_i = 1$ and $\sum_j b_j = 1$). Let $c(x_i, y_j) = c_{ij} \geq 0$.

Notice that for π a probability measure on $X \times Y$, we have that π is completely determined by the $\pi_{ij} := \pi(\{x_i\}, \{y_j\})$, the amount of mass transported from x_i to y_j. The constraint that π be a transference plan can be expressed by the equations $a_i = \sum_j \pi_{ij}$, $b_j = \sum_i \pi_{ij}$.

Then the Kantorovich problem is actually a linear program in this case:

$$\begin{aligned}
\text{minimize} \quad & \sum_{i=1}^{N} \pi_{ij} c_{ij} \\
\text{subject to} \quad & \sum_{j=1}^{N} \pi_{ij} = a_i \qquad (j = 1, \ldots, N) \\
& \sum_{i=1}^{N} \pi_{ij} = b_j \qquad (i = 1, \ldots, N) \\
& \pi_{ij} \geq 0 \qquad (i, j = 1, \ldots, N).
\end{aligned}$$

Indeed, the general Kantorovich problem can be thought of as an infinite-dimensional linear program, and an extension of the duality theory from finite-dimensional linear programs is of great use in the theory of optimal transport.

Let us further narrow our attention to the case in which $N = M$ and $a_i = b_i = \frac{1}{N}$ for all $i = 1, \dots, N$. It is convenient to consider the linear program

$$\begin{aligned} \text{minimize} \quad & \sum_{i=1}^{N} \pi_{ij} c_{ij} \\ \text{subject to} \quad & \sum_{j=1}^{N} \pi_{ij} = 1 \qquad (j = 1, \dots, N) \\ & \sum_{i=1}^{N} \pi_{ij} = 1 \qquad (i = 1, \dots, N) \\ & \pi_{ij} \geq 0 \qquad (i, j = 1, \dots, N). \end{aligned}$$

If $\left(\pi^*_{ij}\right)$ solves this linear program, then $\frac{1}{N}\left(\pi^*_{ij}\right)$ solves the Kantorovich problem (and the reverse implication is also true). Notice that the constraints on π are precisely the constraints that define the set of bistochastic matrices (i.e., one could take this as the definition of the bistochastic matrices). Let $\mathcal{B}_n$ denote the set of bistochastic matrices. Then it is immediate that $\mathcal{B}_n$ is compact and convex, and furthermore we have the following.

Theorem 1. *(Birkhoff.) The $n \times n$ permutation matrices are the extremal points of $\mathcal{B}_n$.*

The proof is not hard, but including it would make for excessive digression. The steps are outlined in Villani [36].

Since $\sum_{i=1}^{N} \pi_{ij} c_{ij}$ is affine in the π_{ij}, by Choquet's theorem this expression attains its minimum at an extremal point of $\mathcal{B}_n$. (We can expect that 'generically' this minimizer is unique.) Thus there exists a permutation matrix $\left(\pi^*_{ij}\right)$ such that $\frac{1}{N}\left(\pi^*_{ij}\right)$ solves the Kantorovich problem. Evidently, the bijection $T^* : \{1, \dots, N\} \to \{1, \dots, N\}$ associated to the permutation $\left(\pi^*_{ij}\right)$ then solves the Monge problem.

The problem of finding a bijection $T : \{1, \dots, N\} \to \{1, \dots, N\}$ that minimizes $\sum_{i=1}^{N} c_{i,T(i)}$ is called the *assignment problem.* Though this is a combinatorial problem, the preceding discussion indicates that the assignment problem can solved by considering its relaxation to the above linear program. Somewhat surprisingly, there do in fact exist combinatorial methods for solving the assignment problem that are faster than solving the corresponding linear program. The Jonker-Volgenant algorithm (see [20]), for example, runs with worst-case complexity $O(N^3)$, which is perhaps surprisingly fast given the combinatorial nature of the problem, though still, practically speaking, not particularly fast. The convenience of solving the assignment problem using this algorithm on a standard computer wanes as N climbs past 1000.

4. The stability of optimal transport

It is very natural to hope that if we change μ or ν slightly, then (roughly speaking) the optimal transference plan π and optimal map T (if they exist) do not change too much. Such stability does indeed hold in a useful sense.

Theorem 2. *(Stability of optimal transport I.) Let X and Y be open subsets of $\mathbb{R}^n$, and let $c : X \times Y \to \mathbb{R}$ be a continuous cost function with $\inf c > -\infty$. Let μ_k and ν_k be sequences of probability measures on X and Y, respectively, such that μ_k converges weakly to μ and ν_k converges weakly to ν. For each k, let π_k be an optimal transference plan between μ_k and ν_k. Lastly assume that*

$$\int c\, d\pi_k < +\infty \text{ for all } k, \ \liminf_k \int c\, d\pi_k < +\infty.$$

Then a subsequence of π_k converges weakly to an optimal transference plan (solving the Kantorovich problem).

Proof. See Theorem 5.20 of Villani [37]. We have stated the theorem here in significantly less generality. □

We also have a notion of stability for the optimal transport maps that solve the Monge problem.

Corollary 3. *(Stability of optimal transport II.) With the same hypotheses as in Theorem 2, further suppose that there exist measurable maps $T_k, T : X \to Y$ such that $\pi_k = (\mathrm{Id}, T_k)_\# \mu_k$ and $\pi = (\mathrm{Id}, T)_\# \mu$ is the unique solution to the Kantorovich problem. Moreover, assume that there exists a constant $C > 0$ such that $\mu_k \leq C\mu$ for all k.*

Then for all $\varepsilon > 0$, we have that

$$\lim_{k\to\infty} \mu_k \left[\{x \in X \,:\, |T_k(x) - T(x)| > \varepsilon\}\right] = 0.$$

In particular, if $\mu_k = \mu$ for al k, then T_k converges to T in μ-probability.

Proof. We are inspired by the proof of Corollary 5.23 of Villani [37], but we make signficant modifications. As was the case for Theorem 2, more general results are available.

First note that by Theorem 2 and the uniqueness of π, we have that $\pi_k \to \pi$ weakly.

Now let $\varepsilon > 0$ and $\delta > 0$. By Lusin's theorem, there exists a compact set $K \subset X$ with $\mu(X\backslash K) < C^{-1}\delta$ (so $\mu_k(X\backslash K) < \delta$) such that the restriction of T to K is continuous. Then let

$$A_\varepsilon = \{(x, y) \in K \times Y \,:\, |T(x) - y| \geq \varepsilon\}.$$

By the continuity of T on K, A_ε is closed in $K \times Y$, hence also in $X \times Y$. Since $\pi = (\mathrm{Id}, T)_\# \mu$, meaning in particular that π is concentrated on the graph of T, we have that $\pi(A_\varepsilon) = 0$. Then by weak convergence and the fact that A_ε is closed, we have that

$$\begin{aligned} 0 = \pi(A_\varepsilon) &\geq \limsup_{k\to\infty} \pi_k(A_\varepsilon) \\ &= \limsup_{k\to\infty} \pi_k\left(\{(x, y) \in K \times Y \,:\, |T(x) - y| \geq \varepsilon\}\right) \end{aligned}$$

$$
\begin{aligned}
&= \limsup_{k\to\infty} \mu_k\left(\{x \in K \,:\, |T(x) - T_k(x)| \geq \varepsilon\}\right) \\
&\geq \limsup_{k\to\infty} \mu_k\left(\{x \in X \,:\, |T(x) - T_k(x)| \geq \varepsilon\}\right) - \mu_k(X \backslash K) \\
&\geq \limsup_{k\to\infty} \mu_k\left(\{x \in X \,:\, |T(x) - T_k(x)| \geq \varepsilon\}\right) - \delta,
\end{aligned}
$$

and the desired result follows. □

Remark 4. The reader can verify by examining the proof that if we remove the assumption of the existence of a constant $C > 0$ with $\mu_k \leq C\mu$ and replace it with the assumption that T is continuous, then the same conclusion holds.

5. Naive method for numerical optimal transport

Our discussion of the discrete problem, together with the stability results just presented, suggest an approach for computing optimal maps numerically. First, take a sequence of discrete measures $\mu_k = \sum_{i=1}^{N_k} a_{ik}\delta_{x_{ik}}$ and $\nu_k = \sum_{j=1}^{M_k} b_{jk}\delta_{y_{jk}}$ converging weakly to μ and ν, respectively. Then compute an optimal transference plan π_k by linear programming, or, in the case that $N_k = M_k$ and $a_{ik} = b_{jk} = \frac{1}{N_k}$, by solving the corresponding assignment problem. In the first case, we get that a subsequence of π_k converges to a solution of the Kantorovich problem associated with μ and ν.

In the second case, as long as we know that there exists a unique continuous solution T to the Monge-Kantorovich problem, then by the second stability result (Corollary 3) we have that T_k converges to T in the following sense: for any $\varepsilon > 0$, the percentage of the x_{ik} at which the error $|T_k - T|$ exceeds ε tends to zero as $k \to \infty$.

The generality of the setting in which this result holds is noteworthy. Indeed, we have assumed little about the cost function c. Also, by using generalizations of the above stability results, we may even work in non-euclidean settings. However, in exchange for the generality of our assumptions, we have received a guarantee of convergence in a fairly weak sense.

This approach is, in addition, quite computationally expensive because the assignment problem does not scale particularly well. Furthermore, it is somewhat difficult to choose discrete μ_k and ν_k that approximate μ and ν, and if the density of μ (or ν) varies considerably, then many points will be needed to do so.

There is another drawback which is more subtle. Suppose for simplicity that μ and ν are measures with uniform density on their respective supports and that T solves the corresponding Monge problem. Suppose that we choose μ_k and ν_k so that, for each k, the Dirac masses at x_{ik} and y_{jk} comprising these discrete measures are distributed 'isotropically' throughout $\operatorname{supp}\mu$ and $\operatorname{supp}\nu$, respectively, in some sense. Now it very well could be the case (and in fact usually is the case) that for fixed k the $T(x_{ik})$ are *not* distributed isotropically throughout $\operatorname{supp}\nu$. (T may distort volume differently in different directions.) But in fixing the positions y_{jk} beforehand we demand that for each k the images $T_k(x_{ik})$ are distributed isotropically. Notionally we then understand that solving the assignment problem does a poor job of pinning down the first-order properties of T (i.e., its Jacobian). In order to have any hope of pinning down these properties at some length scale R, we have to take N_k large enough such that many points lie within any ball of radius R and then somehow smooth T_k using a mollifier supported on the ball of radius R about the origin. This requires N_k to be quite large for a good approximation.

Moreover, in the case in which the cost function has some structure, it is intuitively clear that this method is 'throwing away information' as methods for solving the assignment problem work for an arbitrary cost matrix (c_{ij}). Thus there is hope that approaches more customized to the specific optimal transport problem at hand will yield better results.

6. Convex analysis and optimal transport with the quadratic cost

In this subsection, we indicate the connection between convex analysis and optimal transport with the quadratic cost. We direct the reader to the appendix for useful definitions and facts from convex analysis. We begin with a definition.

Definition 5. *We say that a subset $\Gamma \subset \mathbb{R}^n \times \mathbb{R}^n$ is* cyclically monotone *if for all $m \geq 1$ and all $(x_1, y_1), \ldots, (x_m, y_m) \in \Gamma$, we have*

$$\sum_{i=1}^{m} |x_i - y_i|^2 \leq \sum_{i=1}^{m} |x_i - y_{i-1}|^2$$

(where we understand $y_0 = y_m$), or equivalently

$$\sum_{i=1}^{m} \langle y_i, x_{i+1} - x_i \rangle \leq 0$$

(where we understand $x_{m+1} = x_1$).

In the context of optimal transport, we understand each point $(x, y) \in \Gamma$ to be a pairing of a source point and a target point. There is a cost $|x - y|^2$ associated to this pairing. Intuitively speaking, cyclical monotonicity means that we cannot decrease the total cost of a finite number of pairings in Γ by permuting the target points of these pairings. With this in mind, the following result should not be too surprising.

Theorem 6. *Let μ and ν be probability measures on $\mathbb{R}^n$, and let $\pi \in \Pi(\mu, \nu)$ solve the associated Kantorovich problem with quadratic cost $c(x, y) = |x - y|^2$. Then the support of π is cyclically monotone.*

Proof. For a sketch, see Proposition 2.24 of Villani [36]. □

With the connection between optimal transport and cyclical monotonicity established, we see that understanding cyclically monotone sets will be of great service in characterizing optimal maps. The following theorem is crucial to this endeavor.

Theorem 7. *(Rockafellar.) A nonempty subset $\Gamma \subset \mathbb{R}^n \times \mathbb{R}^n$ is cyclically monotone if and only if it is included in* Graph $(\partial\varphi)$ *for a proper lower semi-continuous convex function φ on $\mathbb{R}^n$.*

Proof. (Following the proof of Theorem 2.27 of Villani [36]) Suppose that φ is a proper lower-semicontinuous convex function. Clearly any subset of a cyclically monotone set is cyclically monotone, so for the 'if' direction it suffices to show that the subdifferential of φ is cyclically monotone.

Let $(x_1, y_1), \ldots, (x_m, y_m) \in \mathrm{Graph}\,(\partial\varphi)$, so for all $z \in \mathbb{R}^n$, we have that $\varphi(z) \geq \varphi(x_i) + \langle y_i, z - x_i \rangle$. Thus we have that

$$\begin{cases} \varphi(x_2) & \geq \varphi(x_1) + \langle y_1, x_2 - x_1 \rangle \\ \varphi(x_3) & \geq \varphi(x_2) + \langle y_2, x_3 - x_2 \rangle \\ & \vdots \\ \varphi(x_1) & \geq \varphi(x_m) + \langle y_m, x_1 - x_m \rangle . \end{cases}$$

Taking the sum of these equalities we obtain $0 \geq \sum_i \langle y_i, x_{i+1} - x_i \rangle$, giving cyclical monotonicity by definition.

Conversely, let $\Gamma \subset \mathbb{R}^n \times \mathbb{R}^n$ be cyclically monotone. Then let $(x_0, y_0) \in \Gamma$ and set

$$\varphi(x) = \sup \{ \langle y_m, x - x_m \rangle + \cdots + \langle y_0, x_1 - x_0 \rangle \, : \, m \in \mathbb{N}, \, (x_i, y_i) \in \Gamma \} .$$

Now φ is the pointwise supremum of affine functions (which are in particular convex), so φ is lower semi-continuous and convex. (The reader may easily verify that the pointwise supremum of convex functions is convex.) Then by cyclical monotonicity, $\varphi(x_0) \leq 0$, so it follows that φ is proper.

It remains to show that $\Gamma \subset \mathrm{Graph}(\partial\varphi)$, and for this it suffices to show that, given $(x, y) \in \Gamma$, $\varphi(z) \geq \varphi(x) + \langle y, z - x \rangle$ for all $z \in \mathbb{R}^n$.

Let $\delta > 0$. Then there exist (by the definition of φ) m and x_i, y_i such that

$$\varphi(x) - \delta \leq \langle y_m, x - x_m \rangle + \cdots + \langle y_0, x_1 - x_0 \rangle .$$

It follows that

$$\varphi(x) - \delta + \langle y, z - x \rangle \leq \langle y, z - x \rangle + \langle y_m, x - x_m \rangle + \cdots + \langle y_0, x_1 - x_0 \rangle .$$

Let $x_{m+1} = x$ and $y_{m+1} = y$, and note that by applying the definition of φ to the RHS, we obtain

$$\varphi(x) - \delta + \langle y, z - x \rangle \leq \varphi(z).$$

Since $\delta > 0$ was arbitrary, we in fact have $\varphi(z) \geq \varphi(x) + \langle y, z - x \rangle$, as desired. $\square$

Clearly from the preceding two results we have:

Theorem 8. *Let μ and ν be probability measures on $\mathbb{R}^n$, and let $\pi \in \Pi(\mu, \nu)$ solve the associated Kantorovich problem with quadratic cost $c(x, y) = |x - y|^2$. Then the support of π is contained in* $\mathrm{Graph}(\partial\varphi)$ *for a proper lower-semicontinuous convex function φ.*

The preceding was meant to provide some background for the following result, central to this thesis, so that it does not come as a complete surprise. Still, the proof is quite involved, and relating it here would take us too far afield.

Theorem 9. *(Brenier.) Let μ and ν be probability measures on $\mathbb{R}^n$ with finite second-order moments, i.e., satisfying*

$$\int |x|^2 \, d\mu(x) + \int |y|^2 \, d\nu(y) < +\infty,$$

and suppose that μ is absolutely continuous with respect to the Lebesgue measure. Consider the associated Monge and Kantorovich problems with quadratic cost $c(x, y) =$

$|x-y|^2$. There exists a unique transference plan $\pi \in \Pi(\mu, \nu)$ that solves the Kantorovich problem, given by

$$d\pi(x, y) = d\mu(x)\, \delta[y = \nabla\varphi(x)],$$

or equivalently,

$$\pi = (\mathrm{Id} \times \nabla\varphi)_{\#}\, \mu,$$

where $\nabla\varphi$ is the unique (i.e., uniquely determined $d\mu$-almost everywhere) gradient of a convex function φ (called the convex potential) satisfying $\nabla\varphi_{\#}\mu = \nu$. Furthermore, $\mathrm{Supp}(\nu) = \overline{\nabla\varphi\,(\mathrm{Supp}(\mu))}$. It follows that $\nabla\varphi$ is the unique solution to the Monge problem.

If we further assume that ν is absolutely continuous with respect to the Lebesgue measure, then $\nabla\varphi^* \circ \nabla\varphi(x) = x$ for $d\mu$-almost all x, and $\nabla\varphi \circ \nabla\varphi^*(y) = y$ for $d\nu$-almost all y. Furthermore, $\nabla\varphi^*$ is the unique gradient of a convex function that pushes ν forward to μ as well as the solution of the Monge problem with quadratic cost and source and target measures ν and μ, respectively.

Proof. SeeTheorem 2.12 of Villani [36]. □

7. The discontinuity of optimal maps

Though Brenier's theorem is indeed a remarkable and useful result, it is still desirable to know more about the regularity of solutions of optimal transport problems. Indeed, under the conditions of Brenier's theorem, it is not possible to guarantee that the optimal map $T = \nabla\varphi$ is even continuous. Consider for example, a situation in which the support of the source is a disc and the support of the target measure is the union of two disjoint discs. It is clear in this case that T must be discontinuous.

To guarantee the continuity of T, it is actually not enough to take the support of ν to be simply connected (and, in particular, connected). A counterexample due to Caffarelli is the following. Let the support of μ be a disc, and let the density of μ be uniform on its support. Let the support of ν consist of the union of two opposite-facing half-discs, joined by a thin rectangular 'bridge,' and likewise let ν have uniform density on its support. If we take the 'bridge' to be sufficiently thin, then the optimal map $T = \nabla\varphi$ is discontinuous. This result is intuitive: as the width of the bridge tends to zero, by the stability of optimal transport we expect that the optimal map will approach (in some sense) the optimal map from the ball to the disjoint union of two half-discs, which is discontinuous. (See Theorem 12.1 of Villani [37] for details.) In fact, as we shall see, the 'correct' condition to place on the target measure is that its support be convex.

One of the motivations for this research was to investigate discontinuities in optimal maps. What do discontinuity sets look like, and how 'large' are discontinuities? Numerical approximation of optimal maps will, among other things, allow us to examine such discontinuities visually and gain intuition about their behavior.

8. The Monge-Ampère equation and regularity

The regularity theory of optimal transport is achieved through the analysis of a nonlinear partial differential equation, called a Monge-Ampère equation. We outline the derivation of this equation and state the main result we need regarding the regularity of its solutions.

Let $d\mu(x) = f(x)\,dx$ and $d\nu(y) = g(y)\,dy$ be probability measures on $\mathbb{R}^n$ (absolutely continuous with respect to the Lebesgue measure) with finite second-order moments. By Brenier's theorem, there exists a $d\mu$-almost everywhere unique gradient of a convex function $\nabla\varphi$ with $\nabla\varphi_{\#}\mu = \nu$. Then for ζ continuous and bounded, we have that

$$\int \zeta\, d\nu = \int \zeta \circ \nabla\varphi\, d\mu,$$

i.e.,

$$\int \zeta(y) g(y)\, dy = \int \zeta\left(\nabla\varphi(x)\right) f(x)\, dx$$

Assume that $\nabla\varphi$ is C^1 and that φ is strictly convex (i.e., $\nabla^2\varphi$ is positive definite everywhere). It follows easily that $\nabla\varphi$ is injective. Then by change of variables we obtain

$$\int \zeta(y) g(y)\, dy = \int \zeta\left(\nabla\varphi(x)\right) g\left(\nabla\varphi(x)\right) \det\left(\nabla^2\varphi(x)\right)\, dx,$$

so

$$\int \zeta\left(\nabla\varphi(x)\right) f(x)\, dx = \int \zeta\left(\nabla\varphi(x)\right) g\left(\nabla\varphi(x)\right) \det\left(\nabla^2\varphi(x)\right)\, dx,$$

for all bounded continuous ζ. It follows that $f(x) = g\left(\nabla\varphi(x)\right)\det\left(\nabla^2\varphi(x)\right)$, or

$$\det\left(\nabla^2\varphi(x)\right) = \frac{f(x)}{g\left(\nabla\varphi(x)\right)}.$$

This is an example of a Monge-Ampère equation, which are in general equations of the form

$$\det\left(\nabla^2\varphi(x)\right) = F(x, \varphi(x), \nabla\varphi(x)).$$

Our Monge-Ampère equation can be understood as specifying that near a point x, $\nabla\varphi$ distorts volume according to the ratio of the source density at x and the target density at the image point $\nabla\varphi(x)$.

We now state the major regularity result for optimal transport in our setting of interest.

Theorem 10. *(Caffarelli's regularity theory.) Let X and Y be bounded open subsets of $\mathbb{R}^n$. Let $f \in C^{0,\alpha}(X)$ and $g \in C^{0,\alpha}(Y)$ be positive and bounded away from zero and infinity with $\int_X f = \int_Y g$. Suppose that Y is convex, and let φ be the unique convex potential furnished by Brenier's theorem. Then in fact $\varphi \in C^{2,\alpha}(X)$, and φ is a strong solution of the Monge-Ampère equation*

$$\det \nabla^2\varphi(x) = \frac{f(x)}{g\left(\nabla\varphi(x)\right)},\ x \in X.$$

If, in addition, X is convex, then we have that $\varphi \in C^{1,\alpha}\left(\overline{X}\right)$ and φ is strictly convex on $\overline{X}$.

If furthermore X and Y are both uniformly convex (i.e., defined by $\psi < 0$ for some ψ with $\nabla^2\psi \succeq \lambda I_n$ for some $\lambda > 0$) and of class C^2 and $f \in C^{0,\alpha}\left(\overline{X}\right)$ and $g \in C^{0,\alpha}\left(\overline{Y}\right)$, then $\varphi \in C^{2,\alpha}\left(\overline{X}\right)$.

Remark 11. This is a very difficult (and remarkable) theorem. See Chapter 4 of Villani [36] for an introduction to the regularity theory of optimal transport and for further references. For the statement of this theorem, we have combined statements from Villani [36] and Caffarelli [10].

Part 2. The Monge-Ampère optimization problem

9. Overview

9.1. **A reformulation of Monge's problem.** Let Ω and Λ be bounded open sets in $\mathbb{R}^n$ with Λ convex, and let f and g be positive functions on Ω and Λ, respectively, each bounded away from zero and infinity. For simplicity, assume that f and g are in $C^{0,\alpha}(\overline{\Omega})$ and $C^{0,\alpha}(\overline{\Lambda})$, respectively, and that they define probability measures μ and ν on Ω and Λ, respectively, by

$$\mu = f\,dx, \quad \nu = g\,dx,$$

where dx denotes the Lebesgue measure on $\mathbb{R}^n$. Then by results of Brenier (Theorem 9) and Caffarelli (Theorem 10), there exists a unique solution of the corresponding Monge problem with the quadratic cost, i.e.,

$$\underset{\{T:\Omega\to\Lambda\,:\,T_{\#}\mu=\nu\}}{\text{minimize}} \int_\Omega |T(x)-x|^2\,d\mu(x),$$

and, moreover, T is $C^{1,\alpha}$. Further, the solution is given by $T=\nabla\varphi$, for φ convex and $C^{2,\alpha}$. In addition, φ is the unique (up to an additive constant) strong solution of the Monge–Ampère equation

$$\begin{aligned} \det\left(\nabla^2\varphi(x)\right) &= \frac{f(x)}{g\left(\nabla\varphi(x)\right)}, \quad x\in\Omega, \\ \nabla\varphi(\Omega) &\subset \Lambda. \end{aligned} \tag{1}$$

The following result rephrases Monge's problem as an infinite-dimensional optimization problem. This problem can be considered 'convex' whenever the target measure has log-concave (in particular, uniform) density with convex support. We refer to §10 for a proof, as well as more details and intuition on the aforementioned interpretation.

Proposition 12. *With notation and hypotheses as in the above discussion, φ is the unique solution of the following optimization problem:*

$$\underset{\psi\in\mathcal{J}}{\text{minimize}}\ \mathcal{F}(\psi) := \textstyle\int_\Omega \mathcal{G}(\psi,x)dx,$$

where

$$\mathcal{J} := \{\psi\in C^2(\Omega)\,:\,\psi \ \text{ convex, and } \nabla\psi(\Omega)\subset\Lambda\},$$

and

$$\mathcal{G}(\psi,x) := \max\left\{0, -\log\det\left(\nabla^2\psi(x)\right) - \log g\left(\nabla\psi(x)\right) + \log f(x)\right\}.$$

The previous result, while not difficult to prove, provides the key starting point for our discretization method, that we now turn to discuss.

9.2. **The discrete Monge–Ampère optimization problem.** Let $x_1, \ldots, x_N \in \overline{\Omega}$, and let $T_1, \ldots, T_M \subset \overline{\Omega}$ be (n-dimensional) simplices with vertices in $\{x_i\}$ that together form an 'almost-triangulation' of Ω. By this we mean that the intersection of any two of the T_i is either empty or a common face (of any dimension) and that $\overline{\Omega} \backslash \bigcup_{i=1}^{M} T_i$, has small volume. Note that the triangulation can be made perfect if Ω is a polytope. We denote the vertices of the simplex T_i by $x_{i_0}, \ldots, x_{i_n}$, $i_0, \ldots, i_n \in \{1, \ldots, N\}$.

Definition 13. *The discrete Monge–Ampère optimization problem (DMAOP) is:*

$$\begin{aligned}
\underset{\{\psi_i \in \mathbb{R}, \eta_i \in \mathbb{R}^n\}_{i=1}^N}{\text{minimize}} \quad & \max_{i \in \{1,\ldots,M\}} \max \Big\{ 0, -\log\det H_{(i)}\Big(\eta_{i_0}, \ldots, \eta_{i_n}\Big) - \log g\big(\tfrac{1}{n+1} \textstyle\sum_{j=0}^{n} \eta_{i_j}\big) \\
& \qquad\qquad\qquad + \log f\big(\tfrac{1}{n+1} \textstyle\sum_{j=0}^{n} x_{i_j}\big) \Big\} \\
\text{subject to} \quad & \psi_j \geq \psi_i + \langle \eta_i, x_j - x_i \rangle, \ i, j = 1, \ldots, N, \\
& \eta_i \in \overline{\Lambda}, \ i = 1, \ldots, N, \\
& H_i\left(\eta_{i_0}, \ldots, \eta_{i_n}\right) \succeq 0, \ i = 1, \ldots, M,
\end{aligned}$$

where the matrices H_i, A_i, and B_i are defined by

$$\begin{aligned}
A_i &:= \left[\begin{array}{cccc} (x_{i_1} - x_{i_0}) & (x_{i_2} - x_{i_0}) & \cdots & (x_{i_n} - x_{i_0}) \end{array} \right]^T, \\
B_i &:= \left[\begin{array}{cccc} (\eta_{i_1} - \eta_{i_0}) & (\eta_{i_2} - \eta_{i_0}) & \cdots & (\eta_{i_n} - \eta_{i_0}) \end{array} \right]^T, \\
H_i \equiv H_i\Big(\eta_{i_0}, \ldots, \eta_{i_n}\Big) &:= \frac{1}{2} \left(A_i\right)^{-1} B_i + \frac{1}{2} \Big(\left(A_i\right)^{-1} B_i \Big)^T.
\end{aligned} \tag{2}$$

We will show later that the DMAOP defined above is feasible for a fine enough triangulation, so standard results guarantee that the DMAOP always has a (not necessarily unique) solution in the optimization variables $\psi_1, \ldots, \psi_N$ and $\eta_1, \ldots, \eta_N$. These variables are the discrete analogues of the values of the convex potential and its gradient, respectively, at the points x_i. Observe that H_i is the discrete analogue of the Hessian $\nabla^2 \psi$ at simplex T_i and that H_i is linear in $\eta_{i_0}, \ldots, \eta_{i_n}$. One can think of $\det H_i$ as a measure of the volume distortion of simplex T_i under the map $\nabla \psi$.

Next we take a sequence of almost-triangulations indexed by k (so now both N and M are functions of k, although we will usually omit that dependence from our notation) and construct a piecewise linear convex function $\phi^{(k)}$ associated with the solution of the k-th DMAOP, $\{\psi_j^{(k)}, \eta_j^{(k)}\}_{j=1}^N$. Define

$$G_j^{(k)}(x) := \psi_j^{(k)} + \langle \eta_j^{(k)}, x - x_j^{(k)} \rangle, \qquad j = 1, \ldots, N, \tag{3}$$

so $G_j^{(k)}$ is the (unique) affine function with $\nabla G_j^{(k)}(x_j^{(k)}) = \eta_j^{(k)}$ and $G_j^{(k)}(x_j^{(k)}) = \psi_j^{(k)}$, and define the *approximate Brenier potential* (a name that is justified by Theorem 14 below)

$$\phi^{(k)}(x) := b^{(k)} + \max_{j=1,\ldots N} G_j^{(k)}(x), \tag{4}$$

where $b^{(k)} \in \mathbb{R}$ is chosen such that $\phi^{(k)}(0) = 0$ (and we have assumed, without loss of generality, that $0 \in \Omega$). Notice that we have actually defined $\phi^{(k)}$ on all of $\mathbb{R}^n$.

Evidently $\phi^{(k)}$ is convex, and by the constraints of the DMAOP, also

$$\phi^{(k)}(x_j^{(k)}) = \psi_j^{(k)} + b^{(k)}, \tag{5}$$

and

$$\eta_j^{(k)} \in \partial\phi^{(k)}(x_j^{(k)}). \tag{6}$$

Now we state our main theorem, which concerns the convergence of the approximate Brenier potentials.

Theorem 14. *Suppose that* $\max_i \operatorname{diam} T_i^{(k)}$ *tends to zero as* $k \to \infty$ *and that there are open regions regions* $\Omega_\varepsilon \subset \Omega$ *for* $\varepsilon > 0$ *such that: (i) for any* $\varepsilon > 0$*, we have that an* ε*-neighborhood of* Ω_ε *is contained within the* k*-th almost-triangulation for all* k *sufficiently large, (ii)* $\Omega_\varepsilon \subset \Omega_{\varepsilon'}$ *for* $\varepsilon' \leq \varepsilon$*, and (iii)* $\bigcup_{\varepsilon>0} \Omega_\varepsilon = \Omega$*. Furthermore, suppose that the optimal cost of the* k*-th DMAOP tends to zero as* $k \to \infty$*. Then as* $k \to \infty$*, the approximate Brenier potential* $\phi^{(k)}$ *yielded by the solution of the* k*-th DMAOP converges uniformly on* $\overline{\Omega}$ *to the unique solution* φ *of* (1) *with* $\varphi(0) = 0$*.*

The condition that the optimal cost of the k*-th DMAOP tends to zero as* $k \to \infty$ *holds in particular when we additionally assume that the sequence* $\left\{T_i^{(k)}\right\}_{k=0}^{\infty}$ *of almost-triangulations of* Ω *satisfies the determinant regularity property (see Definition 22 below) and that the convex potential* φ *defined by* (1) *is in the class* $C^{2,\alpha}\left(\overline{\Omega}\right)$ *for some* $\alpha > 0$*. (By Caffarelli's regularity theory (see Theorem 10), we are guaranteed that* $\varphi \in C^{2,\alpha}\left(\overline{\Omega}\right)$ *when we additionally assume that* Ω *and* Λ *are uniformly convex and of class* C^2*.)*

Remark 15. The conditions on the sequence of triangulations are necessary for technical reasons, but they are fulfilled easily in practice.

Remark 16. We see from the statement of the theorem that even in situations in which we cannot guarantee convergence, we can acquire good heuristic evidence in favor of convergence if the optimal cost of the k-th DMAOP becomes small as $k \to \infty$.

Remark 17. For the proof we never require that Λ is convex (though we require that $\partial\Lambda$ has Lebesgue measure zero, which follows from convexity), nor that g is log-concave, though these assumptions ensure that the DMAOP is convex and, thus, feasibly solvable.

10. The continuous setting

We carry over the notation from §9. Proposition 12 is a special case of the following result.

Lemma 18. *With notation and hypotheses as in the above discussion,* φ *is the unique solution of the following optimization problem:*

$$\begin{aligned} \underset{\psi \in \mathrm{Cvx}(\Omega) \cap C^2(\Omega)}{\text{minimize}} \quad & \mathcal{F}(\psi) := \int_\Omega h \circ \mathcal{G}(\psi, x) \cdot \rho(x) dx \\ \text{subject to} \quad & \nabla\psi(\Omega) \subseteq \Lambda, \end{aligned}$$

where

$$\mathcal{G}(\psi, x) := \max\left\{0, -\log\det\left(\nabla^2\psi(x)\right) - \log g\left(\nabla\psi(x)\right) + \log f(x)\right\},$$

and $h : [0,\infty) \to \mathbb{R}$ *is convex and increasing with* $h(0) = 0$*, and* ρ *is a positive function on* Ω*, bounded away from zero and infinity.*

Remark 19. Notice that if g is log-concave, this optimization problem can be thought of as an 'infinite-dimensional convex optimization problem' (where the value of ψ at each point x is an optimization variable). To see that the problem can indeed be thought of as 'convex,' notice/recall that

- $\nabla\psi(x)$ and $\nabla^2\psi(x)$ are linear in ψ
- $\log \circ \det$ is concave on the set of positive semidefinite (symmetric) matrices
- the pointwise maximum of two convex functions is convex
- the composition of a convex increasing function with a convex function is convex
- the set of convex functions is a convex cone
- the specification that $\nabla\psi(\Omega) \subseteq \Lambda$ is a convex constraint since Λ is convex.

These points also demonstrate that the discretized version of the problem (the DMAOP) outlined above is a convex problem in the usual sense.

Remark 20. (Intuitive explanation of lemma.) We can think of the objective function in the statement of the lemma as penalizing 'excessive contraction' of volume by the map $\nabla\psi$ (relative to the 'desired' distortion given by the ratio of f and g) while ignoring 'excessive expansion.' However, since we constrain $\nabla\psi$ to map Ω into Λ, we expect that excessive expansion at any point will result in excessive contraction at another, causing the value of the objective function to be positive. Thus we expect that the optimal ψ must in fact be φ.

Proof. It is clear that $F(\varphi) = 0$ (since φ solves the Monge–Ampère equation). Furthermore, it is clear that $F(\psi) \geq 0$ always. Thus letting ψ be such that $F(\psi) = 0$, it only remains to show that $\psi = \varphi$. For a contradiction, suppose that $\psi \neq \varphi$. Since φ is the unique solution to the Monge–Ampère equation above, there exists some $x_0 \in \Omega$ such that

$$\det\left(\nabla^2\psi(x_0)\right) \neq \frac{f(x_0)}{g\left(\nabla\psi(x_0)\right)}.$$

If we have that the left-hand side is less than the right-hand side in the above, then $G(\psi, x_0) > 0$, so by continuity $G(\psi, x) > 0$ for x in a neighborhood of x_0, and $F(\psi) > 0$. Thus we can assume that in fact

$$\det\left(\nabla^2\psi(x)\right) \geq \frac{f(x)}{g\left(\nabla\psi(x)\right)}$$

for all x, with strict inequality at a point x_0. By continuity, we must also have strict inequality on en entire neighborhood of x_0. In addition, we have that $\det\left(\nabla^2\psi(x)\right)$ is bounded away from zero, so ψ is strictly convex. Thus $\nabla\psi$ is injective, and we obtain by a change of variables

$$\int_{\nabla\psi(\Omega)} g(y)dy \quad = \quad \int_\Omega g\left(\nabla\psi(x)\right)\det\left(\nabla^2\psi(x)\right)dx$$

$$> \int_\Omega f(x)dx.$$

Of course, since $\nabla\psi(\Omega) \subseteq \Lambda$, we have in addition that $\int_\Lambda g(y)dy \geq \int_{\nabla\psi(\Omega)} g(y)dy$. We have arrived at a contradiction because $\int_\Lambda g = \int_\Omega f = 1$. $\square$

11. Discretizing the Monge–Ampère optimization problem

Next we discuss a discretized version of the above problem that can be feasibly solved using standard convex optimization techniques. Once again, we carry over the notation from §9.

Slightly more generally than in Definition 13, we can study optimizers $\{\psi_i, \eta_i\}_{i=1}^N$ of

$$\begin{aligned}
\underset{\{\psi_i\in\mathbb{R},\eta_i\in\mathbb{R}^n\}_{i=1}^N}{\text{minimize}} \quad & F\left(\{\psi_i,\eta_i\}_{i=1}^N\right) := \max_{i\in\{1,\ldots,M\}} \xi\Big(0, -\log\det H_i\Big(\eta_{i_0},\ldots,\eta_{i_n}\Big) \\
& \qquad\qquad - \log g\left(\tfrac{1}{n+1}\textstyle\sum_{j=0}^n \eta_{i_j}\right) \\
& \qquad\qquad + \log f\left(\tfrac{1}{n+1}\textstyle\sum_{j=0}^n x_{i_j}\right)\Big) \\
\text{subject to} \quad & \psi_j \geq \psi_i + \langle \eta_i, x_j - x_i\rangle, \; i,j = 1,\ldots,N, && (7)\\
& \eta_i \in \overline{\Lambda}, \; i = 1,\ldots,N, && (8)\\
& H_i\left(\eta_{i_0},\ldots,\eta_{i_n}\right) \succeq 0, \; i = 1,\ldots,M, && (9)
\end{aligned}$$

where the H_i are as in Definition 13 and where ξ is an non-decreasing convex function with $\xi(x) = 0$ for all $x \leq 0$ and $\xi(x) > 0$ for all $x > 0$.

Further, notice that we can take ξ to be defined by

$$\xi(x) = \begin{cases} 0, & x \leq 0 \\ x^p, & x > 0 \end{cases},$$

for any $p \geq 1$. We may alternatively choose ξ to be smooth. For simplicity in the proofs below we will just assume that $\xi(x) = \max(0, x)$.

Remark 21. Notice that the last constraint, (9), must be expressed more concretely in order to obtain numerical solutions to the optimization problem. Since Λ is convex, it can be well approximated by a convex polytope, so the last constraint could be replaced by affine constraints corresponding to the faces of the polytope. However, it might be more convenient in certain cases to express Λ as the sublevel set of a convex function. For example, if Λ is the unit sphere, we could write the final constraint above as $\|\eta_i\| \leq 1$. In general, of course, the condition (9) is equivalent to η_i being in the unit ball associated to the norm defined by Λ.

12. Convergence of solutions of the DMAOP

In the following we will often consider sequences of DMAOPs indexed by k. We will maintain the notation from §11, adding "(k)" in superscripts as necessary.

12.1. **The objective function.** First, we would like to understand the behavior of the objective function of the DMAOP. In order to formulate a result characterizing this behavior, we need a condition on the regularity of our sequence of almost-triangulations.

Definition 22. *We say that a sequence*

$$\left\{\left(T_1^{(k)},\dots,T_{M_k}^{(k)}\right)\right\}_{k=0}^{\infty}$$

of almost-triangulations of Ω *(in the sense of* §*9.2) satisfies the* determinant regularity property *if there exists* $R>0$ *such that*

$$\det\left(u_{i,1}^{(k)}\ \ u_{i,2}^{(k)}\ \ \cdots\ \ u_{i,n}^{(k)}\right)\geq R$$

for all k *and all* $i=1,\dots,M_k$*, where*

$$u_j^{(k)} := \|x_{i_j}-x_{i_0}\|^{-1}\left(x_{i_j}-x_{i_0}\right).$$

In particular, this property is satisfied in dimension two if the angles of the triangles are bounded below uniformly in k*.*

Proposition 23. *Let*

$$\left\{\left(T_1^{(k)},\dots,T_{M_k}^{(k)}\right)\right\}_{k=0}^{\infty}$$

be a sequence of almost-triangulations of Ω *(in the sense of* §*9.2) satisfying the determinant regularity property of Definition 22 and for which* $\max_i \operatorname{diam} T_i^{(k)}$ *tends to zero as* $k\to\infty$*. Furthermore, suppose that the convex potential* φ *that defines the optimal transport map from* μ *to* ν *is in the class* $C^{2,\alpha}\left(\overline{\Omega}\right)$ *for some* $\alpha>0$*.*
(i) The restriction of φ *to the* k*-th almost-triangulation satisfies the constraints* (7)–(9) *for all* k *sufficiently large, so in particular the* k*-th DMAOP is feasible for all* k *sufficiently large.*
(ii) Let d_k *be the value of the objective function associated to this restriction, i.e.,*

$$d_k := F\left(\left\{\varphi(x_j^{(k)}),\nabla\varphi(x_j^{(k)})\right\}_{j=1}^{N}\right).$$

Then $\lim_k d_k = 0$*.*

Denote by

$$\left\{\psi_j^{(k)},\eta_j^{(k)}\right\}_{j=1}^{N} \tag{10}$$

the solution to the k-th DMAOP. An immediate consequence is:

Corollary 24. *Let* c_k *be the value of the objective function on the solution of the DMAOP associated with the* k*-th almost-triangulation, i.e.,*

$$c_k := F\left(\left\{\psi_j^{(k)},\eta_j^{(k)}\right\}_{i=1}^{N}\right).$$

Then $\lim_k c_k = 0$*.*

Proof of Proposition 23. (i) We claim that the feasibility conditions (7)–(9) are satisfied for $\left\{\varphi(x_j^{(k)}), \nabla\varphi(x_j^{(k)})\right\}$ for all k sufficiently large. First, the convexity of φ implies (7). Second, (8) follows from (1). It remains to check (9). This follows immediately from the strong convexity of $\varphi \in C^{2,\alpha}$ (recall (1) and the fact f, g are positive), together with the following lemma.

Lemma 25. *Let* $H_i^{(k)} \equiv H_i^{(k)}\left(\left\{\nabla\varphi(x_j^{(k)})\right\}\right)$. *Then,*

$$\lim_k \max_{i=1,\ldots,M_k} \|H_i^{(k)} - \nabla^2\varphi(x_{i_0}^{(k)})\|_\infty = 0.$$

Proof. First let $h^{(k)} = \max_i \operatorname{diam} T_i^{(k)}$ (so $h^{(k)} \to 0$ by assumption). We will denote $h^{(k)}$ by h when this convention is clear.

Next we claim that for any $\xi \in C^{1,\alpha}(\overline{\Omega})$, there exists $C > 0$ such that $|D_u\xi(x) - D_u\xi(x')| \le C\|x - x'\|^\alpha$ for any unit vector $u \in \mathbb{R}^n$ and all $x, x' \in \overline{\Omega}$. Indeed, we have that there exist $C_j > 0$ such that $|\partial_j\xi(x) - \partial_j\xi(x')| \le C_j\|x - x'\|^\alpha$ for all $j = 1, \ldots, n$. Now let u be a unit vector. Then

$$\begin{aligned}
|D_u\xi(x) - D_u\xi(x')| &= |u \cdot \nabla\xi(x) - u \cdot \nabla\xi(x')| \\
&\le \|u\|\|\nabla\xi(x) - \nabla\xi(x')\| \\
&= \left(\sum_{j=1}^n |\partial_j\xi(x) - \partial_j\xi(x')|^2\right)^{1/2} \\
&\le \left(\sum_{j=1}^n C_j^2\|x - x'\|^{2\alpha}\right)^{1/2} \\
&= C\|x - x'\|^\alpha,
\end{aligned}$$

where $C = \left(\sum_{j=1}^n C_j^2\right)^{1/2}$, and the claim is proved. The claim implies that in fact there exists C such that $|D_u(\partial_l\varphi)(x) - D_u(\partial_l\varphi)(x')| \le C\|x - x'\|^\alpha$ for all $l = 1, \ldots, n$, all unit vectors $u \in \mathbb{R}^n$, and all $x, x' \in \overline{\Omega}$.

Fix some i. Then with A_i and B_i defined as in Definition 13 (though now dependent on k), notice that the (j,l)-th entry of $A_i\nabla^2\varphi(x_{i_0})$ is $\left(x_{i_j} - x_{i_0}\right)^T(\nabla\partial_l\varphi(x_{i_0}))$, which is of course equal to $D_{v_j}(\partial_l\varphi)(x_{i_0})$, where D_v denotes the directional derivative in the direction v and where $v_j := x_{i_j} - x_{i_0}$. Now $\eta_{i_j} = \nabla\varphi\left(x_{i_j}\right)$, so $\eta_{i_j} - \eta_{i_0} = \nabla\varphi\left(x_{i_j}\right) - \nabla\varphi(x_{i_0})$, and the (j,l)-th entry of B_i is $\partial_l\varphi\left(x_{i_j}\right) - \partial_l\varphi(x_{i_0})$.

Now let $\zeta = \partial_l\varphi$, $x = x_{i_0}$, $y = x_{i_j}$, $\tau_j = \|x_{i_j} - x_{i_0}\|$, and $u_j = \tau_j^{-1}v_j$ (so u is unit length). Then $D_{v_j}(\partial_l\varphi)(x_{i_0}) = \tau_j D_{u_j}\zeta(x)$, and

$$\begin{aligned}
\left|[B_i]_{jl} - \left[A_i\nabla^2\varphi(x_{i_0})\right]_{jl}\right| &= \left|\zeta(y) - \zeta(x) - \tau_j D_{u_j}\zeta(x)\right| \\
&= \left|\int_0^{\tau_j} D_{u_j}\zeta\left((\tau_j - t)x + ty\right)\, dt - \tau_j D_{u_j}\zeta(x)\right|
\end{aligned}$$

$$
\begin{aligned}
&= \left| \int_0^{\tau_j} \left[D_{u_j}\zeta\left((\tau_j - t)x + ty\right) - D_{u_j}\zeta(x) \right] dt \right| \\
&\leq \int_0^{\tau_j} \left| D_{u_j}\zeta\left((\tau_j - t)x + ty\right) - D_{u_j}\zeta(x) \right| dt \\
&\leq \int_0^{\tau_j} C \left\| \left((\tau_j - t)x + ty\right) - x \right\|^\alpha dt \\
&\leq C \int_0^{\tau_j} \tau_j^\alpha \, dt \\
&= C\tau_j^{\alpha+1}.
\end{aligned}
$$

Now write $A_i = DU$, where $D = \operatorname{diag}(\tau_1, \ldots, \tau_n)$ and $U = (u_1 \; u_2 \; \cdots \; u_n)^T$. Then we obtain

$$
\begin{aligned}
\left| \left[D^{-1} B_i \right]_{jl} - \left[D^{-1} A_i \nabla^2 \varphi\left(x_{i_0}\right) \right]_{jl} \right| &= \left| \tau_j^{-1} \left[B_i \right]_{jl} - \tau_j^{-1} \left[A_i \nabla^2 \varphi\left(x_{i_0}\right) \right]_{jl} \right| \\
&\leq C\tau_j^\alpha \leq Ch^\alpha,
\end{aligned}
$$

where C is independent of k, i, j, and l. Now $U^{-1} = \frac{1}{\det U} \left((-1)^{j+l} M_{jl} \right)^T$, where M_{jl} is the (j,l)-th minor of U. Since the u_j are unit vectors, in particular we have that $|U_{jl}| \leq 1$. Since M_{jl} is a polynomial of $(n-1)!$ terms in the U_{jl}, we have that $|M_{jl}| \leq (n-1)!$ for all j, l, and hence $\left| \left[U^{-1} \right]_{jl} \right| \leq \frac{(n-1)!}{\det U}$. By assumption, $\det U$ is bounded below by a constant $R > 0$ (independent of k and i), so we have that $\left| \left[U^{-1} \right]_{jl} \right| \leq R'$ for $R' = R^{-1}(n-1)! > 0$ (independent of k and i). Then it follows that

$$
\begin{aligned}
\left| \left[U^{-1} D^{-1} B_i \right]_{jl} - \left[U^{-1} D^{-1} A_i \nabla^2 \varphi\left(x_{i_0}\right) \right]_{jl} \right| &= \left| \left[U^{-1} \left(D^{-1} B_i - D^{-1} A_i \nabla^2 \varphi\left(x_{i_0}\right) \right) \right]_{jl} \right| \\
&\leq nR' \left| \left[D^{-1} B_i \right]_{jl} - \left[D^{-1} A_i \nabla^2 \varphi\left(x_{i_0}\right) \right]_{jl} \right| \\
&\leq nR'Ch^\alpha,
\end{aligned}
$$

where R' and C are positive constants that are independent of k and i. Of course, since $A_i = DU$, this means precisely that $\max_{i=1,\ldots,M_k} \left\| A_i^{-1} B_i - \nabla^2 \varphi\left(x_{i_0}\right) \right\|_\infty \leq nR'Ch^\alpha$. Recall that $h = h^{(k)}$ is a function of k, and $h^{(k)} \to 0$ as $k \to \infty$ by assumption. Therefore we have shown that $\max_{i=1,\ldots,M_k} \left\| A_i^{-1} B_i - \nabla^2 \varphi\left(x_{i_0}\right) \right\|_\infty \to 0$ as $k \to \infty$. Since $\nabla^2 \varphi\left(x_{i_0}\right)$ is symmetric, we have that

$$
\max_{i=1,\ldots,M_k} \left\| A_i^{-1} B_i - \nabla^2 \varphi\left(x_{i_0}\right) \right\|_\infty = \max_{i=1,\ldots,M_k} \left\| \left(A_i^{-1} B_i \right)^T - \nabla^2 \varphi\left(x_{i_0}\right) \right\|_\infty,
$$

so

$$
\begin{aligned}
\max_{i=1,\ldots,M_k} \left\| H_i - \nabla^2 \varphi\left(x_{i_0}\right) \right\|_\infty &= \max_{i=1,\ldots,M_k} \left\| \frac{1}{2} A_i^{-1} B_i + \frac{1}{2} \left(A_i^{-1} B_i \right)^T - \nabla^2 \varphi\left(x_{i_0}\right) \right\|_\infty \\
&\leq \max_{i=1,\ldots,M_k} \left\| \frac{1}{2} A_i^{-1} B_i - \frac{1}{2} \nabla^2 \varphi\left(x_{i_0}\right) \right\|_\infty \\
&\quad + \max_{i=1,\ldots,M_k} \left\| \frac{1}{2} \left(A_i^{-1} B_i \right)^T - \frac{1}{2} \nabla^2 \varphi\left(x_{i_0}\right) \right\|_\infty
\end{aligned}
$$

$$= \max_{i=1,\ldots,M_k} \left\| A_i^{-1} B_i - \nabla^2 \varphi\left(x_{i_0}\right) \right\|_\infty,$$

and the last expression tends to zero as $k \to \infty$, as was to be shown.

□

(ii) Given that the feasibility conditions (7)–(9) hold, d_k is well-defined. The rest of the proof is devoted to showing that d_k converges to zero.

For any x, let $i^{(k)}(x)$ be the index (an integer in $\{1, \ldots, M\}$) of a simplex containing x in the k-th almost-triangulation. Let $\tilde{\mathcal{H}}^{(k)}$ and $\zeta^{(k)}$ be piecewise constant matrix-valued functions on Ω defined by

$$\tilde{\mathcal{H}}^{(k)}(x) := H^{(k)}_{i^{(k)}(x)} \equiv H^{(k)}_{i^{(k)}(x)}\left(\left\{\nabla\varphi(x_j^{(k)})\right\}\right), \tag{11}$$

and

$$\zeta^{(k)}(x) := \nabla^2\varphi\left(x^{(k)}_{i^{(k)}(x)_0}\right). \tag{12}$$

Then we have from Lemma 25 that $\|\tilde{\mathcal{H}}^{(k)} - \zeta^{(k)}\|_\infty \to 0$. Notice that the entries of $\tilde{\mathcal{H}}^{(k)}$ and $\zeta^{(k)}$ are bounded over k because $\nabla^2\varphi(\overline{\Omega})$ is bounded. It follows that $\lim_k \|\log\det\tilde{\mathcal{H}}^{(k)} - \log\det\zeta^{(k)}\|_\infty = 0$ (note here that $\nabla^2\varphi(\overline{\Omega})$ is compact and entirely contained in the set of positive-definite matrices), and so also

$$\lim_k \max_{i=1,\ldots,M_k} |\log\det H_i^{(k)} - \log\det\nabla^2\varphi(x_{i_0}^{(k)})| = 0. \tag{13}$$

Now since f is uniformly continuous and bounded away from zero on $\overline{\Omega}$, we have that

$$\max_{i=1,\ldots,M_k} |\log f(x_{i_0}^{(k)}) - \log f(y_i^{(k)})| \to 0, \tag{14}$$

where

$$y_i^{(k)} := \frac{1}{n+1}\sum_{j=0}^{n} x_{i_j}^{(k)}$$

denotes the barycenter of T_i.

Next, since g is uniformly continuous and bounded away from zero on $\overline{\Lambda}$ and $\nabla\varphi$ is Lipschitz, it follows that

$$\max_{i=1,\ldots,M_k} |\log g(\nabla\varphi(x_{i_0}^{(k)})) - \log g(z_i^{(k)})| \to 0, \tag{15}$$

where $z_i^{(k)}$ is the barycenter of the simplex formed by the gradients at the vertices of the i-th simplex, i.e.,

$$z_i^{(k)} := \frac{1}{n+1}\sum_{j=0}^{n} \nabla\varphi(x_{i_j}^{(k)}).$$

Now, by (1),

$$\det\nabla^2\varphi(x_{i_0}^{(k)}) = \frac{f(x_{i_0}^{(k)})}{g(\nabla\varphi(x_{i_0}^{(k)}))}. \tag{16}$$

Then,

$$\begin{aligned} d_k &\le \max_{i=1,\dots,M_k} |\log\det H_i^{(k)} - \log f(y_i^{(k)}) + \log g(z_i^{(k)})| \\ &\le \max_i |\log\det H_i^{(k)} - \log f(x_{i_0}^{(k)}) + \log g(\nabla\varphi(x_{i_0}^{(k)}))| \\ &\quad + \max_i |\log f(x_{i_0}^{(k)}) - \log f(y_i^{(k)})| + \max_i |\log g(\nabla\varphi(x_{i_0}^{(k)})) - \log g(z_i^{(k)})|. \end{aligned}$$

The last term tends to zero with k by (15), while the second does so by (14). Finally, the first term tends to zero with k by (13) and (16). □

12.2. **A piecewise affine approximate Brenier potential.** Next we construct a piecewise linear convex function $\phi^{(k)}$ associated with the solution of the k-th DMAOP. Recall from (10) that we denoted by $\{(\psi_j^{(k)}, \eta_j^{(k)})\}$ the solution to the k-th DMAOP. This data gives rise to the functions $\{\phi^{(k)}\}_k$ defined in (4).

We now turn to the proof of our main theorem, stating that the *approximate Brenier potentials* $\phi^{(k)}$ converge to the Brenier potential φ.

Proof of Theorem 14. Let D be a closed disk containing Ω in its interior. For the proof it will suffice to show that every subsequence of $\phi^{(k)}$ that converges uniformly on D converges to φ on $\overline{\Omega}$. Indeed, this follows by the Arzela–Ascoli theorem since $\{\phi^{(k)}\}_k$ is an equicontinuous, uniformly bounded family (since $\eta_j^{(k)} \in \overline{\Lambda}$ for all k, j, with Λ bounded, and $\phi^{(k)}(0) = 0$).

Thus, assume that $\phi^{(k)} \to \phi$ uniformly for some ϕ, and we need only show that $\phi = \varphi$ on Ω. Notice that ϕ is convex as a pointwise limit of convex functions.

12.3. **First order behavior of the barycenteric extension of the gradient of the approximate Brenier potentials.** The subdifferentials $\partial\phi^{(k)}$ are piecewise constant with jump discontinuities. The objective function of the DMAOP provides us with some sort of control over the 'second-order properties' of the $\phi^{(k)}$, but these properties are neither well-defined at this stage nor readily accessible because the $\phi^{(k)}$ are piecewise linear. In order to get a handle on the 'second-order convergence' of the $\phi^{(k)}$, we want to replace these subdifferentials with continuous, piecewise-affine functions that interpolate rather than jump, which we may then differentiate once again.

Thus, with (6) in mind, we define $g^{(k)}$ by barycentrically interpolating the values $\{\eta_{i_j}^{(k)}\}_{j=1}^n$ over the i-th simplex, for all $i = 1, \dots, M_k$, namely, for each x in $T_i = \mathrm{Conv}(x_{i_0}, \dots x_{i_n})$, write $x = \sum \sigma_j x_{i_j}$, with $\sigma_j \in [0,1]$. Then

$$g^k(x) := \sum_{j=0}^{n} \sigma_j \eta_{i_j}^{(k)}, \qquad \text{if } \ x \in T_i \tag{17}$$

(note that this is well-defined also for x lying in more than one simplex). Alternatively, g^k is the unique vector-valued function that is affine on each simplex $T_j^{(k)}$ and satisfies $g^k(x_i^{(k)}) = \eta_i^{(k)}$ for all i.

Lemma 26. $\mathcal{H}^{(k)}(x) := H_{i^{(k)}(x)}^{(k)}\left(\left\{\psi_j^{(k)}\right\}\right) = \mathrm{symm}\left(\nabla g^{(k)}(x)\right)$ *for* $x \in \bigcup_i \mathrm{int}\,(T_i)$.

Proof. We fix some k and then freely omit k from some of our notation in the remainder of the proof. Also we fix i and work within the triangle T_i. Now let $v_j = x_{i_j} - x_{i_0}$. Then evidently $D_{v_j} g^{(k)} \equiv \eta_{i_j} - \eta_{i_0}$ (note that this is an equation of vectors) on $\operatorname{int}(T_i)$ because $g^{(k)}$ is affine on T_i with $g^{(k)}\left(x_{i_j}\right) = \eta_{i_j}$. Now $D_{v_j} g^{(k)} = v_j \cdot \nabla g^{(k)}$, so $D_{v_j} g^{(k)}$ is the j-th row of $A_i \nabla g^{(k)}$, where $\nabla g^{(k)}$ denotes the matrix with j-th row $\frac{\partial}{\partial x_j} g^{(k)}$ and where A_i is as in Definition 13. Since $D_{v_j} g^{(k)} = \eta_{i_j} - \eta_{i_0}$ is also the j-th row of B_i, we have that $B_i = A_i \nabla g^{(k)}$, i.e., $\nabla g^{(k)} = A_i^{-1} B_i$. Now $H_i = \operatorname{symm}\left(A_i^{-1} B_i\right)$, so it follows that $H_i = \operatorname{symm}\left(\nabla g^{(k)}\right)$, as was to be shown. □

We want to show that $g^{(k)}$ approaches $\partial\phi$ in some sense.

Lemma 27. *$g^{(k)} \to \nabla\phi$ almost everywhere.*

Proof. By Theorem 62 (see the appendix), for every x and every $\varepsilon > 0$, there exists a $\delta > 0$ and K such that $\partial\phi^{(k)}(y) \subset \partial\phi(x) + B_\varepsilon(0)$ for all $y \in B_\delta(x)$ and all $k \geq K$. (To use this theorem, we have employed the fact that if a sequence of convex functions converges uniformly on bounded sets to some convex function, then the sequence epi-converges to this function. See Theorem 61.) Fix a point $x \in \Omega$ where ϕ is differentiable and an $\varepsilon > 0$. Additionally, take $\delta > 0$ and K according to the aforementioned result. If necessary, take K even larger, so that for all $k \geq K$ the maximal distance of x to the vertices of the simplices containing it is at most δ. We assume from now on that $k \geq K$. Thus for all vertices $x_j^{(k)}$ of any simplex containing x, we have that $\partial\phi^{(k)}(x_j^{(k)}) \subset \nabla\phi(x) + B_\varepsilon(0)$. In particular, supposing that $g^{(k)}(x)$ has been obtained by interpolation of $\{\eta_{i_j}^{(k)}\}_{j=0}^{n}$ for some simplex T_i containing x, and since $\eta_{i_j}^{(k)} \in \partial\phi^{(k)}(x_{i_j}^{(k)})$, we have that $\eta_{i_j}^{(k)} \in \nabla\phi(x) + B_\varepsilon(0)$. Since $g^{(k)}(x)$ is a convex combination of the $\eta_{i_j}^{(k)}$, we have that $g^{(k)}(x) \in \nabla\phi(x) + B_\varepsilon(0)$. This proves that $g^{(k)} \to \nabla\phi$ almost everywhere. □

12.4. **Second order control.** Unfortunately, we do not have enough regularity to maintain that $\nabla g^{(k)}$ approaches $\nabla^2\phi$ almost everywhere. We can obtain this regularity by convolving everything with a sequence of mollifiers.

The motivation for doing so is fairly intuitive. Strictly speaking, the second-order behavior of the $\phi^{(k)}$ is completely trivial. The second-derivatives of the $\phi^{(k)}$ are everywhere either zero or undefined. However, by virtue of solving the DMAOP, the $\phi^{(k)}$ do actually contain second-order information in some sense. Indeed, we may think of the graphs of the $\phi^{(k)}$ as having some sort of curvature that becomes apparent when we 'blur' $\phi^{(k)}$ on a small scale and then take k large enough so that the scale of the discretization is much smaller than the scale of the blurring. This blurring is achieved by convolving with smooth mollifiers.

Let ξ_ε be a standard set of mollifiers (supported on $B_\varepsilon(0)$). Notice that $g^{(k)}$ is only defined on the almost-triangulation of Ω, so we run into trouble near the boundary when convolving with ξ_ε. Thus, define regions $\Omega_\varepsilon \subset \Omega$ for $\varepsilon > 0$ such that: (i) for any $\varepsilon > 0$, we have that an ε-neighborhood of Ω_ε is contained within the k-th almost-triangulation for all k sufficiently large, (ii) $\Omega_\varepsilon \subset \Omega_{\varepsilon'}$ for $\varepsilon' \leq \varepsilon$, and (iii) $\bigcup_{\varepsilon>0} \Omega_\varepsilon = \Omega$ (so $\lim_{\varepsilon\to 0} \mu(\Omega_\varepsilon) = \mu(\Omega)$).

We aim to show the following:

Lemma 28. *$\tilde{\mathcal{H}}_\varepsilon^{(k)} := \nabla g^{(k)} \star \xi_\varepsilon \to \nabla^2(\phi \star \xi_\varepsilon)$ uniformly (in each of the n^2 components).*

First, we require some auxiliary results:

Lemma 29. *$g^{(k)} \star \xi_\varepsilon \to \nabla\phi \star \xi_\varepsilon$ uniformly (in each of the n components).*

Proof. First, we claim that $g^{(k)} \star \xi_\varepsilon \to \nabla\phi \star \xi_\varepsilon$ pointwise on Ω_ε as $k \to \infty$. (Note that $\nabla\phi \star \xi_\varepsilon$ is everywhere defined because $\nabla\phi$ exists almost everywhere.) To check that this is true, note that for $x \in \Omega_\varepsilon$,

$$|g^{(k)} \star \xi_\varepsilon(x) - \nabla\phi \star \xi_\varepsilon(x)| = |\int (g^{(k)}(y) - \nabla\phi(y))\xi_\varepsilon(y-x)\,dy| \leq \int |g^{(k)}(y) - \nabla\phi(y)|\xi_\varepsilon(y-x)\,dy.$$

Since $|g^{(k)}(y) - \nabla\phi(y)|\xi_\varepsilon(x-y) \to 0$ a.e. and the $g^{(k)}$ are uniformly bounded, our claim follows from bounded convergence (note that ε is constant in this limit). Similarly, $\phi^{(k)} \star \xi_\varepsilon \to \phi \star \xi_\varepsilon$ pointwise because the $\phi^{(k)}$ are uniformly bounded.

Now notice that the $g^{(k)} \star \xi_\varepsilon$ are uniformly bounded in k (because the $g^{(k)}$ are uniformly bounded), and furthermore, $\nabla(g^{(k)} \star \xi_\varepsilon) = g^{(k)} \star \nabla\xi_\varepsilon$. Let B be such that $\|g_i^{(k)}\|_\infty \leq B$ for all $i = 1, \ldots, n$, and write

$$|g_i^{(k)} \star \nabla\xi_\varepsilon(x)| = |\int g_i^{(k)}(y)\nabla\xi_\varepsilon(x-y)\,dy| \leq B|\int \nabla\xi_\varepsilon(x-y)\,dy|,$$

so the $g^{(k)} \star \xi_\varepsilon$ have uniformly bounded derivatives (in each component).

The statement now follows from the general fact that if a uniformly bounded sequence of differentiable functions f_n with uniformly bounded derivatives satisfies $f_n \to f$ pointwise, then $f_n \to f$ uniformly. □

Lemma 30. *(i) $\nabla\phi \star \xi_\varepsilon = \nabla(\phi \star \xi_\varepsilon)$ and (ii) $\nabla g^{(k)} \star \xi_\varepsilon = \nabla(g^{(k)} \star \xi_\varepsilon)$.*

Proof. Notice that for $h \neq 0$

$$\begin{aligned}
\frac{1}{h}\left[\phi \star \xi_\varepsilon(x + e_j h) - \phi \star \xi_\varepsilon(x)\right] &= \frac{1}{h}\left[\int \xi_\varepsilon(x + e_j h - y) g_i^{(k)}(y)\,dy - \int \xi_\varepsilon(x-y)\phi(y)\,dy\right] \\
&= \frac{1}{h}\left[\int \xi_\varepsilon(x-y)\phi(y + e_j h)\,dy - \int \xi_\varepsilon(x-y)\phi(y)\,dy\right] \\
&= \int \xi_\varepsilon(x-y) \cdot \frac{1}{h}\left[\phi(y + e_j h)\,dy - \phi(y)\right]\,dy.
\end{aligned}$$

Clearly the integrand in the last expression approaches $\xi_\varepsilon(x-y) \cdot \partial_j\phi$ almost everywhere as $h \to 0$ (since ϕ is differentiable a.e.). Now ϕ is Lipschitz continuous (since its subgradients are bounded), so the integrand is uniformly bounded over h, and by bounded convergence we obtain $\partial_j(\phi \star \xi_\varepsilon) = \partial_j\phi \star \xi_\varepsilon$, so $\nabla\phi \star \xi_\varepsilon = \nabla(\phi \star \xi_\varepsilon)$, as desired.

Evidently $g_i^{(k)}$ is Lipschitz continuous, so by the same reasoning we obtain the second desired equality. □

Proof of Lemma 28. This will follow from another general fact: for one-variable smooth functions $f_n, f : I \to \mathbb{R}$ (where I is a closed interval of positive length) such that (i) $f_n \to f$ uniformly, (ii) the f_n'' are uniformly bounded, and (iii) f'' is bounded, we have that $f_n' \to f'$ uniformly. This follows from an application of the Landau-Kolmogorov inequality $\|g_n'\|_\infty \le C\|g_n\|_\infty^{\frac{1}{2}}\|g_n''\|_\infty^{\frac{1}{2}}$ (see for instance Chui and Smith [13]) for smooth $g : I \to \mathbb{R}$ such that g, g'' are bounded. We can apply this inequality to $g_n = f_n - f$ to obtain $\|f_n' - f'\|_\infty \le C\|f_n - f\|_\infty^{\frac{1}{2}}\|f_n'' - f''\|_\infty^{\frac{1}{2}}$. Since f'' is bounded and the f_n'' are uniformly bounded, we have that $\|f_n'' - f''\|_\infty^{\frac{1}{2}}$ is uniformly bounded in n. And of course $\|f_n - f\|_\infty^{\frac{1}{2}} \to 0$ by the uniform convergence $f_n \to f$. Therefore $\|f_n' - f'\|_\infty \to 0$ as claimed.

Now the functions $\partial_j^2(g_i^{(k)} \star \xi_\varepsilon) = g_i^{(k)} \star \partial_j^2\xi_\varepsilon$ are uniformly bounded over the k (following from the uniform boundedness of $g_i^{(k)}$ as above). Also, since $\nabla\phi \star \xi_\varepsilon = \nabla(\phi \star \xi_\varepsilon)$ by (i) of Lemma 30, we have that $\nabla(\nabla\phi \star \xi_\varepsilon) = \nabla^2(\phi \star \xi_\varepsilon)$, which is smooth because $\phi \star \xi_\varepsilon$ is smooth. Therefore $\nabla(\nabla\phi \star \xi_\varepsilon)$ is bounded in all of its components.

Let $x \in \Omega_\varepsilon$. Then fix i, j and let $\delta > 0$ be small enough such that $I := \{x + te_j : t \in [-\delta, \delta]\} \subset \Omega_\varepsilon$. Recall that (by Lemma 29) $g_i^{(k)} \star \xi_\varepsilon \to \partial_i\phi \star \xi_\varepsilon$ uniformly. Then by restricting to the j-th variable and applying our fact we obtain that $\partial_j(g_i^{(k)} \star \xi_\varepsilon) \to \partial_j(\partial_i\phi \star \xi_\varepsilon)$ uniformly on $I \ni x$. Since x, i, and j were arbitrary, we see that $\nabla(g^{(k)} \star \xi_\varepsilon) \to \nabla(\nabla\phi \star \xi_\varepsilon)$ pointwise (though we cannot yet say that this convergence is uniform). The uniformity of the convergence follows from our earlier fact that pointwise convergence, together with uniform boundedness and uniform boundedness of derivatives, implies uniform convergence. Then (i) and (ii) of Lemma 30 together imply the desired result. □

We have shown that $\tilde{\mathcal{H}}_\varepsilon^{(k)} := \nabla g^{(k)} \star \xi_\varepsilon \to \nabla^2(\phi \star \xi_\varepsilon)$ uniformly. By the symmetry of $\nabla^2(\phi \star \xi_\varepsilon)$, it follows that the symmetrized matrix

$$\mathcal{H}_\varepsilon^{(k)} := \frac{1}{2}\tilde{\mathcal{H}}_\varepsilon^{(k)} + \frac{1}{2}\left(\tilde{\mathcal{H}}_\varepsilon^{(k)}\right)^T$$

converges to $\nabla^2(\phi \star \xi_\varepsilon)$ uniformly. We have shown the following:

Lemma 31. $\mathcal{H}_\varepsilon^{(k)} := \operatorname{symm}\left(\nabla g^{(k)} \star \xi_\varepsilon\right) \to \nabla^2(\phi \star \xi_\varepsilon)$ *uniformly.*

12.5. **Obtaining a density inequality.** In this subsection we will employ the fact that $\{\psi_j^{(k)}, \eta_j^{(k)}\}_{j=1}^N$ solves the k-th DMAOP in order to obtain an inequality that controls (in a certain sense) the deviation of $\det \mathcal{H}_\varepsilon^{(k)}$ *in one direction* from the desired ratio of densities appearing in the Monge-Ampère equation. This will suggest that $\nabla(\phi \star \xi_\varepsilon)$ cannot 'excessively' shrink volume.

Lemma 32.

$$\log\det \mathcal{H}_\varepsilon^{(k)} \ge \log \frac{f \circ \tau^{(k)}}{g \circ \gamma^{(k)}} \star \xi_\varepsilon - c_k.$$

Proof. By the continuity of det, it follows from Lemma 31 that $\det\left(\mathcal{H}_\varepsilon^{(k)}\right) \to \det \nabla^2(\phi \star \xi_\varepsilon)$ pointwise. Let x be any point. Then

$$\begin{aligned}
\log\det\left(\mathcal{H}_\varepsilon^{(k)}(x)\right) &= \log\det\left(\left(\mathcal{H}^{(k)} \star \xi_\varepsilon\right)(x)\right) \\
&= \log\det \int \mathcal{H}^{(k)}(y - x)\xi_\varepsilon(y)\, dy
\end{aligned}$$

$$\begin{aligned} &\geq \int \log\det\left(\mathcal{H}^{(k)}(y-x)\right)\xi_\varepsilon(y)\,dy \\ &= \left(\log\det\mathcal{H}^{(k)}\star\xi_\varepsilon\right)(x), \end{aligned}$$

where we have used Jensen's inequality and the concavity of $\log\det$ on the positive semidefinite cone in the penultimate step. More compactly, we have:

$$\log\det\mathcal{H}_\varepsilon^{(k)} \geq \left(\log\det\mathcal{H}^{(k)}\right)\star\xi_\varepsilon. \tag{18}$$

Now with c_k denoting the optimal cost of the k-th probem, we have by optimality that

$$-\log\det\mathcal{H}^{(k)}(x) - \log g(\gamma^{(k)}(x)) + \log f(\tau^{(k)}(x)) \leq c_k,$$

where $\tau^{(k)}(x)$ is the center of simplex $i^{(k)}(x)$ and $\gamma^{(k)}(x)$ is the mean of the $\eta_j^{(k)}$ at the vertices of simplex $i^{(k)}(x)$. We can rewrite this inequality as

$$\log\det\mathcal{H}^{(k)}(x) \geq \log\frac{f(\tau^{(k)}(x))}{g(\gamma^{(k)}(x))} - c_k.$$

Therefore by (18) we have the desired result.

□

12.6. **Passing to the limit in k.** The goal of this subsection is to 'take the limit as $k\to\infty$' of the result of the preceding subsection so that we can employ Lemma 31. The proof of the following lemma is somewhat technical and relies on convex analysis, though the result should seem intuitive. The main obstacle is controlling the behavior of $\tau^{(k)}$ and $\gamma^{(k)}$.

Lemma 33. *With $\phi_\varepsilon := \phi\star\xi_\varepsilon$, we have*

$$\det\nabla^2(\phi_\varepsilon)(x) \geq \frac{\inf\{f(y) : y\in B_\varepsilon(x)\}}{\sup\{g\left(\nabla\phi(y)\right) : y\in B_\varepsilon(x), \nabla\phi(y)\text{ exists}\}},$$

for $x\in\Omega_\varepsilon$.

Remark 34. Notice that in the case that f and g are uniform densities on Ω and Λ, respectively, the proof of this lemma is considerably easier. Indeed, we can skip right to (25) below. Even in the case that only g is uniform, the proof is much easier. This is true because the most difficult part of the proof is controlling the behavior of $\gamma^{(k)}$, which requires results from convex analysis, most crucially a result on the 'locally uniform' convergence of the subdifferentials of a sequence of convergent convex functions.

Proof. Let $\alpha,\beta>0$. Fixing some x and using [2, Theorem 8.3] again we have that for every $z\in\overline{B}_\alpha(x)$, there exists $\delta_z>0$ and $N_{z,\beta}$ such that $\partial\phi^{(k)}(y)\subset\partial\phi(z)+B_\beta(0)$ for all $y\in B_{\delta_z}(z)$ and $k\geq N_x$. By compactness, there exist $z_1,\ldots,z_M$ such that the B_{δ_i} cover $\overline{B}_\alpha(x)$, where $\delta_i := \delta_{z_i}$. Thus we have that

$$\partial\phi^{(k)}(y) \subset \left[\bigcup_{z\in\overline{B}_\alpha(x)}\partial\phi(z)\right]_\beta \tag{19}$$

for all $k \geq N'_{x,\alpha,\beta} := \max_i N_{z_i,\beta}$ and all $y \in \overline{B}_\alpha(x)$, where $[\cdot]_\beta$ denotes the β-neighborhood of its argument.

For k sufficiently large (i.e., $k \geq N_\alpha$ for some N_α depending only on α), we have that the simplex $i^{(k)}(x)$ containing x is contained in $\overline{B}_\alpha(x)$ by the regularity of our sequence of almost-triangulations, so in particular $\tau^{(k)}(x) \in \overline{B}_\alpha(x)$. Also, $\gamma^{(k)}(x)$ is a convex combination of $n+1$ elements of $\bigcup_{y\in\overline{B}_\alpha(x)} \partial\phi^{(k)}(y)$, so by (19) we have that $\gamma^{(k)}(x) \in \mathrm{Conv}\left(\left[\bigcup_{z\in\overline{B}_\alpha(x)} \partial\phi(z)\right]_\beta\right)$ for all $k \geq N_{x,\alpha,\beta} := \max\{N_\alpha, N'_{x,\alpha,\beta}\}$.

Thus for k sufficiently large (i.e., $k \geq N_{x,\alpha,\beta}$)

$$\log \frac{f(\tau^{(k)}(x))}{g(\gamma^{(k)}(x))} \geq \log \frac{\min\{f(y) : y \in \overline{B}_\alpha(x)\}}{\max\left\{g(z) : z \in \mathrm{Conv}\left(\left[\bigcup_{z\in\overline{B}_\alpha(x)} \partial\phi(z)\right]_\beta\right)\right\}}. \tag{20}$$

For almost every x (see Theorem 58 and Theorem 57), we have that for any $\gamma > 0$, there exists $\alpha > 0$ such that

$$\partial\phi(x+v) \subset \nabla\phi(x) + B_\gamma(0)$$

for all $v \in \overline{B}_\alpha(0)$. Therefore, for positive α sufficiently small (i.e., $0 < \alpha < C_{x,\gamma}$ for some $C_{x,\gamma}$), $z \in \overline{B}_\alpha(x)$ implies that $\partial\phi(z) \subset \nabla\phi(x) + B_\gamma(0) = B_\gamma(\nabla\phi(x))$, so $\bigcup_{z\in\overline{B}_\alpha(x)} \partial\phi(z) \subset B_\gamma(\nabla\phi(x))$. Then it follows that $\left[\bigcup_{z\in\overline{B}_\alpha(x)} \partial\phi(z)\right]_\beta \subset B_{\gamma+\beta}(\nabla\phi(x))$, implying that

$$\mathrm{Conv}\left(\left[\bigcup_{z\in\overline{B}_\alpha(x)} \partial\phi(z)\right]_\beta\right) \subset B_{\gamma+\beta}(\nabla\phi(x)). \tag{21}$$

Therefore for a.e. x and any $\alpha, \beta, \gamma > 0$ with $\alpha \in (0, C_{x,\gamma})$ we have by (20) and (21) that

$$\log \frac{f(\tau^{(k)}(x))}{g(\gamma^{(k)}(x))} \geq \log \frac{\min\{f(y) : y \in \overline{B}_\alpha(x)\}}{\max\{g(z) : z \in B_{\gamma+\beta}(\nabla\phi(x))\}}$$

for $k \geq N_{x,\alpha,\beta}$. It follows that

$$\liminf_{k\to\infty} \log \frac{f(\tau^{(k)}(x))}{g(\gamma^{(k)}(x))} \geq \log \frac{\min\{f(y) : y \in \overline{B}_\alpha(x)\}}{\max\{g(z) : z \in B_{\gamma+\beta}(\nabla\phi(x))\}}. \tag{22}$$

for a.e. x and any $\alpha, \beta, \gamma > 0$ with $\alpha \in (0, C_{x,\gamma})$. Clearly

$$\lim_{\alpha\to 0} \min\{f(y) : y \in \overline{B}_\alpha(x)\} = f(x) \tag{23}$$

(by the continuity of f), and

$$\lim_{(\beta,\gamma)\to 0} \max\{g(z) : z \in B_{\gamma+\beta}(\nabla\phi(x))\} = g(\nabla\phi(x)) \tag{24}$$

(also by continuity) for a.e. x. Taking limits in (38) (first $\alpha \to 0$, followed by $(\beta,\gamma) \to 0$) and applying (23) and (24), we obtain

$$\liminf_{k\to\infty} \log \frac{f(\tau^{(k)}(x))}{g(\gamma^{(k)}(x))} \geq \log \frac{f(x)}{g(\nabla\phi(x))} \tag{25}$$

for a.e. x. Now recalling from Lemma 32 that $\log\det \mathcal{H}_\varepsilon^{(k)} \geq \log \frac{f\circ\tau^{(k)}}{g\circ\gamma^{(k)}} \star \xi_\varepsilon - c_k$ and using the fact that $c_k \to 0$ (Corollary 24), we see for $x \in \Omega_\varepsilon$ that

$$
\begin{aligned}
\liminf_{k\to\infty} \log\det \mathcal{H}_\varepsilon^{(k)}(x) &\geq \liminf_{k\to\infty} \left[\left(\log \frac{f\circ\tau^{(k)}}{g\circ\gamma^{(k)}} \star \xi_\varepsilon \right)(x) - c_k \right] \qquad (26)\\
&= \liminf_{k\to\infty} \left(\log \frac{f\circ\tau^{(k)}}{g\circ\gamma^{(k)}} \star \xi_\varepsilon \right)(x) \\
&= \liminf_{k\to\infty} \int \xi_\varepsilon(x-y)\cdot \log \frac{f(\tau^{(k)}(y))}{g(\gamma^{(k)}(y))}\, dy \\
&\geq \int \xi_\varepsilon(x-y)\cdot \liminf_{k\to\infty} \log \frac{f(\tau^{(k)}(y))}{g(\gamma^{(k)}(y))}\, dy \\
&\geq \int \xi_\varepsilon(x-y)\cdot \log \frac{f(y)}{g(\nabla\phi(y))}\, dy \\
&= \left(\log \frac{f}{g\circ\nabla\phi} \star \xi_\varepsilon \right)(x), \qquad (27)
\end{aligned}
$$

where we have used the Fatou-Lebesgue theorem to pass the $\liminf$ within the integral and (25) in the penultimate step. In particular, it follows that $\det \mathcal{H}_\varepsilon^{(k)} \geq \frac{1}{2}\cdot\frac{\min f}{\max g} > 0$ for k sufficiently large, so recalling from earlier (Lemma 31) that $\det \mathcal{H}_\varepsilon^{(k)} \to \det \nabla^2(\phi\star\xi_\varepsilon)$, we now have in fact that $\log\det \mathcal{H}_\varepsilon^{(k)} \to \log\det \nabla^2(\phi\star\xi_\varepsilon)$ (so the sequence in the LHS of (26) is actually convergent). Combining this with (27) we have

$$\log\det \nabla^2(\phi\star\xi_\varepsilon) \geq \log \frac{f}{g\circ\nabla\phi} \star \xi_\varepsilon,$$

which implies (since ξ_ε is supported on $B_\varepsilon(0)$)

$$\log\det \nabla^2(\phi\star\xi_\varepsilon)(x) \geq \log \frac{\inf\{f(y) : y\in B_\varepsilon(x)\}}{\sup\{g\left(\nabla\phi(y)\right) : y \in B_\varepsilon(x), \nabla\phi(y)\text{ exists}\}},$$

for $x \in \Omega_\varepsilon$, and the lemma follows. □

12.7. **Passing to the limit in ε.** In this subsection we consider the measures ν_ε obtained by pushing forward the restriction of μ to Ω_ε by $\nabla\phi_\varepsilon$. Using our determinant inequality from the preceding subsection, we will show that a subsequence of these measures (roughly speaking) converges weakly to the target measure ν. Intuitively speaking, our inequality from the last subsection says that $\nabla\phi_\varepsilon$ does not shrink volume 'excessively' at any point. Given that the image of $\nabla\phi_\varepsilon$ must lie within Λ, it is understandable that our claimed result should hold.

Since convolution with a non-negative kernel preserves convexity, ϕ_ε is in fact a convex function. Lemma 33 implies that ϕ_ε has positive definite Hessian on Ω_ε, and it follows that $\nabla\phi_\varepsilon$ is injective on Ω_ε. (Indeed, for any two points $x, y \in \Omega_\varepsilon$, consider the restriction of ϕ_ε to the line containing these two points. The second directional derivative in the direction of a unit vector parallel to this line must be non-negative along this line and strictly positive near both x and y (after possibly flipping the direction of the unit vector). Thus ∇x and ∇y cannot agree.)

By Theorem 11.1 of Big Villani $\nu_\varepsilon := (\nabla\phi_\varepsilon)_\# (\mu|_{\Omega_\varepsilon})$ is absolutely continuous on $\Lambda_\varepsilon := \nabla\phi_\varepsilon(\Omega_\varepsilon)$ with density

$$
\begin{aligned}
g_\varepsilon(x) &= f\left((\nabla\phi_\varepsilon)^{-1}(x)\right) \cdot \det\left[\nabla^2\phi_\varepsilon\left((\nabla\phi_\varepsilon)^{-1}(x)\right)\right]^{-1} \\
&\le f\left((\nabla\phi_\varepsilon)^{-1}(x)\right) \cdot \frac{\sup\left\{g(\nabla\phi(y)) : y \in B_\varepsilon\left((\nabla\phi_\varepsilon)^{-1}(x)\right), \nabla\phi(y)\,\text{exists}\right\}}{\inf\left\{f(y) : y \in B_\varepsilon\left((\nabla\phi_\varepsilon)^{-1}(x)\right)\right\}} \qquad (28)\\
&=: \overline{g}_\varepsilon(x) \qquad (29)
\end{aligned}
$$

for $x \in \Lambda_\varepsilon$, where the inequality follows from Lemma 33.

Now $(\nabla\phi_\varepsilon)^{-1}(\Lambda\backslash\Lambda_\varepsilon)$ does not intersect Ω_ε (because $\nabla\phi_\varepsilon(\Omega_\varepsilon) = \Lambda_\varepsilon$), so $\nu_\varepsilon(\Lambda\backslash\Lambda_\varepsilon) = 0$. Therefore in fact ν_ε is absolutely continuous on all of Λ with density $g_\varepsilon(x) = 0$ for $x \in \Lambda\backslash\Lambda_\varepsilon$. Setting $\overline{g}_\varepsilon := g$ on $\Lambda\backslash\Lambda_\varepsilon$, we have (since $g \ge 0$ everywhere) that

$$g_\varepsilon \le \overline{g}_\varepsilon \qquad (30)$$

on all of Λ. We need the following result:

Lemma 35. *$\overline{g}_\varepsilon \to g$ a.e. on $\Lambda\backslash\partial\phi(\mathrm{Sing}(\phi))$ as $\varepsilon \to 0$, where $\mathrm{Sing}(\phi)$ denotes the set of points at which ϕ is not differentiable.*

Remark 36. Similarly to the proof of Lemma 33 (and for similar reasons), the proof of this lemma becomes considerably easier in the case that g is a uniform density and trivial in the case that both f and g are uniform densities.

To prove the lemma in full generality, we need a preliminary result. Before even stating this result, we make several technical remarks. In what follows we consider ϕ to be defined on D as the pointwise limit of the $\phi^{(k)}$ and extended to all of $\mathbb{R}^n$ by taking the value $+\infty$ outside of D. (Note that thus far, we have only considered ϕ inside of $\Omega \subset D$, and there has been no need to consider its behavior elsewhere). Accordingly, ϕ is convex on $\mathbb{R}^n$ and, in addition, lower semi-continuous. It is clear from the definition of the convex conjugate that ϕ^* (the convex conjugate of ϕ) is finite on all of $\mathbb{R}^n$, so $\operatorname{int}\operatorname{dom}\phi^* = \operatorname{dom}\phi^* = \mathbb{R}^n$.

Lemma 37. *(i) $\nabla\phi_\varepsilon^* = (\nabla\phi_\varepsilon)^{-1}$ on Λ_ε, where ϕ_ε^* is the convex conjugate of ϕ_ε.*
(ii) $\partial\phi_\varepsilon^* \to \nabla\phi^*$ almost everywhere.

Proof. First note that for $x \in \Lambda_\varepsilon$, Lemma 33 implies in particular that ϕ_ε has positive definite Hessian at $(\nabla\phi_\varepsilon)^{-1}(x) \in \Omega_\varepsilon$, so ϕ_ε^* (the convex conjugate of ϕ_ε) is differentiable at x with $\nabla\phi_\varepsilon^*(x) = (\nabla\phi_\varepsilon)^{-1}(x)$.

It remains only to show (ii). Note that ϕ_ε and ϕ are actually defined on all of $\mathbb{R}^n$ and that $\phi_\varepsilon \to \phi$ uniformly on compact subsets (as ϕ is continuous). Then by Theorem 61, the ϕ_ε epi-converge to ϕ. Then by Theorem 66, we have that the ϕ_ε^* epi-converge to ϕ^*. Again using Theorem 62, we have that $\partial\phi_\varepsilon^*(x) \to \nabla\phi^*(x)$ for all $x \in \operatorname{int}\operatorname{dom}\phi^*$ such that $\nabla\phi^*(x)$ exists (i.e., almost everywhere). □

Proof of Lemma 35. Let $x \in \Lambda\backslash\partial\phi(\mathrm{Sing}(\phi))$ be an x for which

$$\partial\phi_\varepsilon^*(x) \to \nabla\phi^*(x)$$

(so in particular $\nabla\phi^*$ is differentiable at x), so by (ii) of Lemma 37 it will suffice to show that $\overline{g}_\varepsilon(x) \to g(x)$.

Let $E = \{\varepsilon > 0 : x \in \Lambda_\varepsilon\}$. Since $\overline{g}_\varepsilon(x) = g(x)$ whenever $\varepsilon \notin E$, it is clear that it will suffice to show that $\overline{g}_{\varepsilon_j}(x) \to g(x)$ for all sequences $\varepsilon_j \in E$ that tend to zero. Let ε_j be such a sequence. Notice that since $x \in \Lambda_{\varepsilon_j}$ for all j, by (i) of Lemma 37 we have that $\nabla\phi^*_{\varepsilon_j}(x) = \left(\nabla\phi_{\varepsilon_j}\right)^{-1}(x)$ for all j, so

$$\left(\nabla\phi_{\varepsilon_j}\right)^{-1}(x) \to \nabla\phi^*(x). \tag{31}$$

Let $\gamma > 0$. Then by (31), there exists N such that for $j \geq N$ we have

$$\left|\left(\nabla\phi_{\varepsilon_j}\right)^{-1}(x) - \nabla\phi^*(x)\right| < \gamma/2.$$

We can assume that in fact N is large enough such that $\varepsilon_j < \gamma/2$ for all $j \geq N$, so then

$$B_{\varepsilon_j}\left(\left(\nabla\phi_{\varepsilon_j}\right)^{-1}(x)\right) \subset B_\gamma\left(\nabla\phi^*(x)\right) \cap \Omega$$

for $j \geq N$. Thus for all such j we have that

$$\inf\left\{f(y) : y \in B_{\varepsilon_j}\left(\left(\nabla\phi_{\varepsilon_j}\right)^{-1}(x)\right)\right\} \geq \inf\left\{f(y) : y \in B_\gamma\left(\nabla\phi^*(x)\right) \cap \Omega\right\}.$$

Then taking the lim inf as $j \to \infty$ followed by the limit as $\gamma \to 0$, we have (using the continuity of f) that

$$\liminf_{j\to\infty} \inf\left\{f(y) : y \in B_{\varepsilon_j}\left(\left(\nabla\phi_{\varepsilon_j}\right)^{-1}(x)\right)\right\} \geq f\left(\nabla\phi^*(x)\right).$$

Also notice that

$$\inf\left\{f(y) : y \in B_{\varepsilon_j}\left(\left(\nabla\phi_{\varepsilon_j}\right)^{-1}(x)\right)\right\} \leq f\left(\left(\nabla\phi_{\varepsilon_j}\right)^{-1}(x)\right) \xrightarrow[j\to\infty]{} f\left(\nabla\phi^*(x)\right)$$

(where the limit follows by the continuity of f). Then we have that

$$\limsup_{j\to\infty} \inf\left\{f(y) : y \in B_{\varepsilon_j}\left(\left(\nabla\phi_{\varepsilon_j}\right)^{-1}(x)\right)\right\} \leq f\left(\nabla\phi^*(x)\right),$$

so in fact

$$\lim_{j\to\infty} \inf\left\{f(y) : y \in B_{\varepsilon_j}\left(\left(\nabla\phi_{\varepsilon_j}\right)^{-1}(x)\right)\right\} = f\left(\nabla\phi^*(x)\right).$$

It follows that

$$\lim_{j\to\infty} \frac{f\left(\left(\nabla\phi_{\varepsilon_j}\right)^{-1}(x)\right)}{\inf\left\{f(y) : y \in B_{\varepsilon_j}\left(\left(\nabla\phi_{\varepsilon_j}\right)^{-1}(x)\right)\right\}} = 1.$$

Therefore for the lemma it remains only to show that

$$\lim_{j\to\infty} \sup\left\{g(\nabla\phi(y)) : y \in B_{\varepsilon_j}\left(\left(\nabla\phi_{\varepsilon_j}\right)^{-1}(x)\right), \nabla\phi(y)\,\text{exists}\right\} = g(x). \tag{32}$$

By the same arguments as above we have that

$$\begin{aligned} &\limsup_{j\to\infty} \sup\left\{g(\nabla\phi(y)) : y \in B_{\varepsilon_j}\left(\left(\nabla\phi_{\varepsilon_j}\right)^{-1}(x)\right), \nabla\phi(y)\,\text{exists}\right\} \\ \leq\ & \sup\left\{g(\nabla\phi(y)) : y \in B_\gamma\left(\nabla\phi^*(x)\right) \cap \Omega, \nabla\phi(y)\,\text{exists}\right\} \end{aligned} \tag{33}$$

for any $\gamma > 0$.

We claim that ϕ is differentiable at $\nabla\phi^*(x)$. By Theorem 65, $y \in \partial\phi(\nabla\phi^*(x))$ if and only if $\nabla\phi^*(x) \in \partial\phi^*(y)$. Thus plugging in x for y, we see that $x \in \partial\phi(\nabla\phi^*(x))$. By assumption, $x \notin \partial\phi\,(\mathrm{Sing}(\phi))$, so it must be that $\nabla\phi^*(x) \notin \mathrm{Sing}(\phi)$, as claimed. In particular, we have also shown that $\nabla\phi\,(\nabla\phi^*(x)) = x$.

Now by Theorem 58, for any $\alpha > 0$ there exists $\gamma > 0$ such that

$$\partial\phi(\nabla\phi^*(x) + v) \subset \nabla\phi(\nabla\phi^*(x)) + B_\alpha(0) = B_\alpha(x) \tag{34}$$

for all $v \in \overline{B}_\gamma(0)$. Together with the continuity of g, this implies that the RHS of (33) converges to $g(x)$ as $\gamma \to 0$, so we have

$$\limsup_{j\to\infty} \sup\left\{ g(\nabla\phi(y)) : y \in B_{\varepsilon_j}\left(\left(\nabla\phi_{\varepsilon_j}\right)^{-1}(x)\right), \nabla\phi(y)\,\text{exists}\right\} \le g(x).$$

Of course, we also have

$$\sup\left\{ g(\nabla\phi(y)) : y \in B_{\varepsilon_j}\left(\left(\nabla\phi_{\varepsilon_j}\right)^{-1}(x)\right), \nabla\phi(y)\,\text{exists}\right\} \ge g\left(\nabla\phi\left(\left(\nabla\phi_{\varepsilon_j}\right)^{-1}(x) + v_j\right)\right),$$

where v_j is a vector with length less than ε_j chosen so that ϕ is differentiable at $\left(\nabla\phi_{\varepsilon_j}\right)^{-1}(x) + v_j$. Then following from (34), the continuity of g, (31), and the fact that $\nabla\phi\,(\nabla\phi^*(x)) = x$, we have that the RHS tends to $g(x)$ as $j \to \infty$. Thus (32) holds, and the proof of the lemma is complete.

□

Lemma 38. *For any sequence $\varepsilon_n \to 0$, $\mu(\Omega_{\varepsilon_n})^{-1}\nu_{\varepsilon_n}$ is a sequence of probability measures converging weakly to $\nu(\Lambda)^{-1}\nu$.*

Proof. Let $\varepsilon_n \to 0$, and let ζ be a bounded continuous function on $\mathbb{R}^n$. Then by the change of variables formula for the pushforward measure,

$$\begin{aligned}\mu(\Omega_{\varepsilon_n})^{-1}\int \zeta\, d\nu_{\varepsilon_n} &= \mu(\Omega_{\varepsilon_n})^{-1}\int \zeta\circ\nabla\phi_{\varepsilon_n}\, d\mu|_{\Omega_{\varepsilon_n}} \\ &= \mu(\Omega_{\varepsilon_n})^{-1}\int_\Omega \zeta\circ\nabla\phi_{\varepsilon_n}\cdot\chi_{\Omega_{\varepsilon_n}}\cdot f\,dx.\end{aligned}$$

Now $\nabla\phi_{\varepsilon_n} \to \nabla\phi$ pointwise almost everywhere, and $\chi_{\Omega_{\varepsilon_n}} \to \chi_\Omega$ pointwise, so (recalling that ζ is bounded and continuous), we have by bounded convergence and the fact that $\mu(\Omega_{\varepsilon_n}) \to \mu(\Omega)$ that

$$\begin{aligned}\lim_{n\to\infty} \mu(\Omega_{\varepsilon_n})^{-1}\int \zeta\, d\nu_{\varepsilon_n} &= \mu(\Omega)^{-1}\int_\Omega \zeta\circ\nabla\phi\, d\mu \\ &= \mu(\Omega)^{-1}\int_\Omega \zeta\, d\left((\nabla\phi)_\#\mu\right).\end{aligned}$$

This proves that $\mu(\Omega)\mu(\Omega_{\varepsilon_n})^{-1}\nu_{\varepsilon_n}$ converges weakly to $\tilde{\nu} := (\nabla\phi)_\#\mu$. Observe that $\tilde{\nu}$ must be absolutely continuous because the densities of the $\mu(\Omega)\mu(\Omega_{\varepsilon_n})^{-1}\nu_{\varepsilon_n}$ are bounded above uniformly in n (see (28)) and supported on the compact set $\overline{\Lambda}$ Therefore $\tilde{\nu}$ has a density $\tilde{g}$.

Define $S := \partial\phi\,(\mathrm{Sing}(\phi))$. Then we claim that $(\nabla\phi^*)^{-1}(\mathrm{Sing}(\phi)) \cap \mathrm{Diff}(\phi^*) = S \cap \mathrm{Diff}(\phi^*)$, where Diff denotes the set of points of differentiability of its argument. Indeed, suppose that $y \in \text{LHS}$. Then $\nabla\phi^*(y) \in \text{Sing}(\phi)$. Now (just as in the proof of Lemma 35) $y \in \partial\phi\left(\nabla\phi^*(y)\right)$, so $y \in S$.

Then suppose that $y \in S \cap \text{Diff}(\phi^*)$. Then $y \in \partial\phi(x)$ for some $x \in \text{Sing}(\phi)$, implying by Theorem 65 that $x \in \partial\phi^*(y)$, i.e., $x = \nabla\phi^*(y)$, and $y \in (\nabla\phi^*)^{-1}(\text{Sing}(\phi))$. This gives the claimed set equality.

Recall that the set of points of differentiability of a continuous function (e.g., ϕ and ϕ^*) is Borel (and hence also is its complement), so $\text{Sing}(\phi)$, $\text{Sing}(\phi^*)$, and their complements are Borel.

By Theorem 58, $\nabla\phi^*$ is continuous on $\text{Diff}(\phi^*)$. Define h to agree with $\nabla\phi^*$ on $\text{Diff}(\phi^*)$ and to be identically equal to z_0 on $\text{Sing}(\phi^*)$. Let U be open. If $z_0 \notin U$, then $h^{-1}(U)$ is open in $\text{Diff}(\phi^*)$, i.e., $h^{-1}(U) = O \cap \text{Diff}(\phi^*)$ for some open O. If $z_0 \in U$, then $h^{-1}(U)$ is the union of $\text{Sing}(\phi^*)$ with some set open in $\text{Diff}(\phi^*)$. In either case, $h^{-1}(U)$ is Borel. Therefore h is a Borel-measurable function. Therefore $h^{-1}(\text{Sing}(\phi))$ is Borel. Notice that $h^{-1}(\text{Sing}(\phi)) = (\nabla\phi^*)^{-1}(\text{Sing}(\phi)) \cap \text{Diff}(\phi^*)$, so in fact $S \cap \text{Diff}(\phi^*)$ is Borel.

We compute:

$$\begin{aligned}\tilde{\nu}(S \cap \text{Diff}(\phi^*)) &= \int \chi_{S\cap\text{Diff}(\phi^*)}\, d\left((\nabla\phi)_{\#}\mu\right) \\ &= \int_\Omega \chi_{S\cap\text{Diff}(\phi^*)} \circ \nabla\phi\, d\mu.\end{aligned}$$

Notice that the integrand in the last expression is only nonzero on a Lebesgue-null set (namely, $\text{Sing}(\phi)$), so since μ is absolutely continuous, we have that $\tilde{\nu}(S\cap\text{Diff}(\phi^*)) = 0$. Therefore $\tilde{\nu} = \tilde{\nu}|_{(S\cap\text{Diff}(\phi^*))^c} = \tilde{\nu}|_T$, where $T := (S \cap \text{Diff}(\phi^*))^c$.

Now $T = S^c \cup \text{Sing}(\phi^*) = S^c \cup E$, where the union in the last expression is disjoint and $E \subset \text{Sing}(\phi^*)$. Since E is contained within a set of Lebesgue-measure zero, E is Lebesgue-measurable with measure zero, and hence S^c is Lebesgue-measurable as well.

By the absolute continuity of $\tilde{\nu}$, for a.e. $x \in \Lambda$ we can write

$$\tilde{g}(x) = \lim_{R\to 0} \frac{\tilde{\nu}(B_R(x))}{\text{vol}(B_R(x))}. \tag{35}$$

Let $\alpha > 0$ and let U be an open set containing T with $m(U\backslash T) < \alpha$ (where m denotes the Lebesgue measure). This is of course possible because T is Borel. Then

$$\tilde{\nu}(B_R(x)) = \tilde{\nu}(B_R(x) \cap T) \leq \tilde{\nu}(B_R(x) \cap U). \tag{36}$$

Now since $B_R(x) \cap U$ is open, by weak convergence we have that

$$\tilde{\nu}(B_R(x) \cap U) \leq \liminf_{n\to\infty} \mu(\Omega)\mu(\Omega_{\varepsilon_n})^{-1}\nu_{\varepsilon_n}(B_R(x) \cap U) \tag{37}$$

for all R.

Observe that since $U = T \cup (U\backslash T)$, we have that

$$B_R(x) \cap U = (B_R(x) \cap T) \cup (B_R(x) \cap (U\backslash T)) \subset (B_R(x) \cap T) \cup (U\backslash T),$$

so then (noting that by (28) there exists a constant C such that $g_\varepsilon \leq C$ for all ε) we see that

$$\nu_{\varepsilon_n}(B_R(x) \cap U) \quad \leq \quad \nu_{\varepsilon_n}(B_R(x) \cap T) + \nu_{\varepsilon_n}(U\backslash T)$$

$$\leq \int_{B_R(x)\cap T} g_{\varepsilon_n} + C \cdot m(U\backslash T)$$
$$\leq \int_{B_R(x)\cap S^c} \overline{g}_{\varepsilon_n} + C\alpha,$$

where in the last step we have used the fact that $T = S^c \cup E$, where E has Lebesgue measure zero, as well as (30). Now by Lemma 35, $\overline{g}_{\varepsilon_n} \to g$ a.e. on S^c, so by bounded convergence (since the g_ε are uniformly bounded) the last expression is convergent and

$$\liminf_{n\to\infty} \mu(\Omega)\mu(\Omega_{\varepsilon_n})^{-1}\nu_{\varepsilon_n}(B_R(x)\cap U) \leq \int_{B_R(x)\cap S^c} g + C\alpha \tag{38}$$

Then by (36), (37), and (38), we have that

$$\tilde{\nu}(B_R(x)) \leq \int_{B_R(x)\cap S^c} g + C\alpha \leq \int_{B_R(x)} g + C\alpha$$

for all $\alpha > 0$, i.e.,

$$\tilde{\nu}(B_R(x)) \leq \int_{B_R(x)} g. \tag{39}$$

Then by (35) and (39), for a.e. x we have

$$\tilde{g}(x) \leq \liminf_{R\to 0} \frac{1}{\mathrm{vol}(B_R(x))} \int_{B_R(x)} g = g(x).$$

Therefore $g \geq \tilde{g}$ a.e., and consequently

$$\begin{aligned}
\int_\Lambda |g-\tilde{g}| &= \int_\Lambda (g-\tilde{g}) \\
&= \int_\Lambda g - \int_\Lambda \tilde{g} \\
&= \nu(\Lambda) - \tilde{\nu}(\Lambda).
\end{aligned} \tag{40}$$

Notice that since $\tilde{\nu} \ll m$ (the Lebesgue measure) and $m(\partial\Lambda) = 0$, we have that $\tilde{\nu}(\partial\Lambda) = 0$, and Λ is a continuity set of $\tilde{\nu}$. Therefore by the weak convergence $\mu(\Omega)\mu(\Omega_{\varepsilon_n})^{-1}\nu_{\varepsilon_n} \Rightarrow \tilde{\nu}$, we see that

$$\tilde{\nu}(\Lambda) = \lim_{n\to\infty} \mu(\Omega)\mu(\Omega_{\varepsilon_n})^{-1}\nu_{\varepsilon_n}(\Lambda) = \mu(\Omega) \lim_{n\to\infty} \mu(\Omega_{\varepsilon_n})^{-1}\mu|_{\Omega_{\varepsilon_n}} \left[(\nabla\phi_{\varepsilon_n})^{-1}(\Lambda) \right].$$

Now $(\nabla\phi_{\varepsilon_n})^{-1}(\Lambda) \supset \Omega_{\varepsilon_n}$, so $\mu|_{\Omega_{\varepsilon_n}} \left[(\nabla\phi_{\varepsilon_n})^{-1}(\Lambda) \right] = \mu(\Omega_{\varepsilon_n})$, and $\tilde{\nu}(\Lambda) = \mu(\Omega)$. Of course, $\mu(\Omega) = \nu(\Lambda)$, so by (40) we have that $\int_\Lambda |g-\tilde{g}| = 0$. This implies that $g = \tilde{g}$ a.e., so $\nu = \tilde{\nu}$, and the lemma follows immediately. □

12.8. **Concluding the proof via the stability of optimal transport.** We are at last in a position to outline the remainder of the proof of the main theorem. Using our result from the preceding subsection, together with the stability of optimal transport, we will show that the maps $\nabla\phi_{\varepsilon_n}$ converge in a certain sense to $\nabla\varphi$. (Though it seems natural that the stability of optimal transport plays a role in this proof, it is perhaps unexpected that we have employed the stability of optimal transport to obtain convergence in ε (rather than in k). As mentioned earlier, we could not take the seemingly more direct route and needed to use mollifiers to obtain regularity.) Then it is easy to show that the $\nabla\phi_{\varepsilon_n}$ converge pointwise to $\nabla\phi$, and it follows after a short argument that $\phi = \varphi$.

Lemma 39. *Let $T := \nabla\varphi$ be the unique optimal transport map from μ to ν. Then for any $\beta > 0$, $T_n := \nabla\phi_{\varepsilon_n} \to T$ in probability with respect to $\mu|_{\Omega_\beta}$ (or more precisely, with respect to the multiple of this measure with total mass one).*

Proof. Now notice that $T_n := \nabla\phi_{\varepsilon_n}$ is an optimal transport map from $\mu(\Omega_{\varepsilon_n})^{-1}\mu|_{\Omega_{\varepsilon_n}}$ to $\mu(\Omega_{\varepsilon_n})^{-1}\nu_{\varepsilon_n}$ (probability measures) for all n. This follows from the fact that ϕ_{ε_n} is convex and Brenier's theorem (Theorem 9). Also evidently $\mu(\Omega_{\varepsilon_n})^{-1}\mu_{\varepsilon_n}$ converges weakly to $\mu(\Omega)^{-1}\mu$. Then by the stability of optimal transport (more precisely, Theorem 3)

$$\mu(\Omega_{\varepsilon_n})^{-1}\mu|_{\Omega_{\varepsilon_n}}\left[\{x \in \Omega : d\left(T_n(x), T(x)\right) \geq \alpha\}\right] \to 0$$

for any $\alpha > 0$, where $T := \nabla\varphi$ is the unique optimal transport map from μ to ν. Of course this implies that

$$\mu|_{\Omega_{\varepsilon_n}}\left[\{x \in \Omega : d\left(T_n(x), T(x)\right) \geq \alpha\}\right] \to 0 \tag{41}$$

for any $\alpha > 0$. Let $\alpha, \beta > 0$. Then for all n sufficiently large $\Omega_\beta \subset \Omega_{\varepsilon_n} \subset \Omega$, so

$$\begin{aligned}
\mu|_{\Omega_\beta}\left[\{x \in \Omega_\beta : d\left(T_n(x), T(x)\right) \geq \alpha\}\right] &= \mu\left[\Omega_\beta \cap \{x \in \Omega_\beta : d\left(T_n(x), T(x)\right) \geq \alpha\}\right] \\
&\leq \mu\left[\Omega_{\varepsilon_n} \cap \{x \in \Omega : d\left(T_n(x), T(x)\right) \geq \alpha\}\right] \\
&= \mu|_{\Omega_{\varepsilon_n}}\left[\{x \in \Omega : d\left(T_n(x), T(x)\right) \geq \alpha\}\right],
\end{aligned}$$

and of course the last expression approaches zero as $n \to \infty$ by (41), and the lemma is proved. □

Lemma 40. *$T_n \to \nabla\phi$ a.e. on Ω.*

Proof. Note first that $\phi_{\varepsilon_n} \to \phi$ pointwise because ϕ is convex, hence continuous. Then by Theorem 59, for any $x \in \Omega$ and $\alpha > 0$, we have that for n sufficiently large $\partial\phi_{\varepsilon_n}(x) \subset \partial\phi(x) + B_\alpha(0)$. Thus at a point x such that $\partial\phi$ is a singleton, we have $\nabla\phi_{\varepsilon_n}(x) \subset \nabla\phi(x) + B_\alpha(0)$, i.e., $|\nabla\phi(x) - \nabla\phi_{\varepsilon_n}(x)| < \alpha$ for n sufficiently large, i.e., $\nabla\phi_{\varepsilon_n}(x) \to \nabla\phi(x)$. Therefore since $\partial\phi$ is a singleton almost everywhere, we have $T_n \to \nabla\phi$ a.e., as desired. □

Now we can finally conclude the proof of Theorem 14. We claim that $\phi = \varphi$ on Ω. Let $\beta > 0$. Now $T_n \to T$ in probability on Ω_β with respect to μ by 39, so there exists a subsequence n_j such that $T_{n_j} \to T$ μ-almost everywhere on Ω_β, hence a.e. on Ω_β (because f is bounded away from zero). But $T_n \to \nabla\phi$ almost everywhere on Ω_β, so we must have that $T = \nabla\phi$ on Ω_β. Since $\bigcup_{\beta>0}\Omega_\beta = \Omega$, we have that $T = \nabla\phi$ on Ω, i.e., $\nabla\varphi = \nabla\phi$ a.e. Since $\varphi \in C^{2,\alpha}(\overline{\Omega})$ and ϕ is Lipschitz on $\overline{\Omega}$ (because convex functions are Lipschitz on compact domains (Theorem 52), both are absolutely continuous. Therefore since $\phi(0) = \varphi(0)$, we have that $\phi = \varphi$. This completes the proof. □

Corollary 41. *For any $x \in \Omega$ and $\varepsilon > 0$, there exists $\delta > 0$ such that $\partial\phi^{(k)}(y) \subset \partial\varphi(x) + B_\varepsilon(0)$ for all $y \in B_\delta(x)$ and all k sufficiently large.*

Proof. This follows immediately from the preceding theorem together with Theorem 62. □

13. Relationship of the DMAOP with discrete optimal transport

We now investigate the connection between the DMAOP and classical discrete optimal transport problems (DOTPs). In fact, the solution of the DMAOP gives the solution to a corresponding DOTP. The key here is Rockafellar's theorem.

Let $\{\psi_j, \eta_j\}_{j=1}^N$ be a solution of the DMAOP (with notation as above). Then we can repeat the construction described above to obtain a piecewise-linear convex function $\phi : \mathbb{R}^n \to \mathbb{R}$ such that $\phi(x_j) = \psi_j$ and $\partial\phi(x_j) \ni \eta_j$. By Rockafellar's theorem (Theorem 7), $\bigcup_{j=1}^N (x_j, \eta_j)$ is cyclically monotone, so $x_j \mapsto \eta_j$ solves the discrete optimal transport problem from $\mu_D = \sum_{j=1}^N \delta_{x_j}$ to $\nu_D = \sum_{j=1}^N \delta_{\eta_j}$. We state this result as a proposition.

Proposition 42. *Let $\{\psi_j, \eta_j\}_{j=1}^N$ be a solution of the DMAOP. Then $T : \{x_j\} \to \{\eta_j\}$ given by $x_j \mapsto \eta_j$ solves the Monge-Kantorovich problem with source $\mu_D = \sum_{j=1}^N \delta_{x_j}$ and target $\nu_D = \sum_{j=1}^N \delta_{\eta_j}$.*

Of course, the target points η_j are not fixed before the optimization problem is solved. Indeed we expect the results of solving the DMAOP to be better than the results obtained by picking N target points in advance and then solving the resulting DOTP. This expectation is based on the fact that the DMAOP, chooses target points η_j in a way that attempts to achieve correct volume distortion.

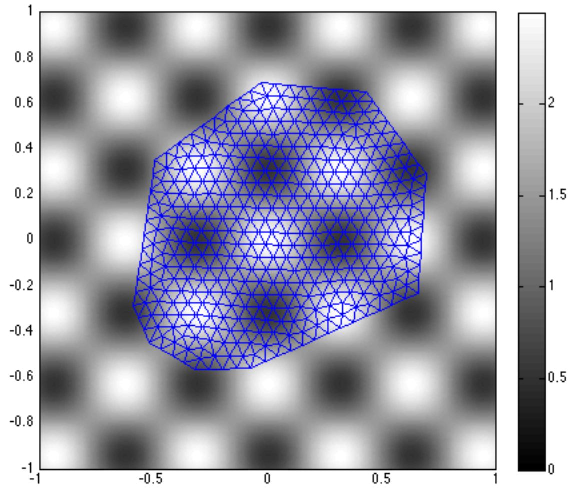
14. Numerical experiments

14.1. **Implementation details.** Only three details of the implementation bear mentioning. First, we used DistMesh for the triangulation of Ω (see [30]). Second, we solved each convex optimization problem using MOSEK (see [26]), called via the modeling language YALMIP (see [23]). Third, our MATLAB program for solving the DMAOP allows the user to hand-draw the support of the source measure, and several of the following examples have source measures with hand-drawn support.

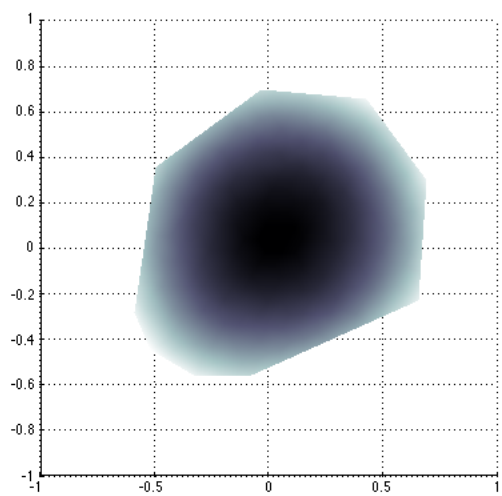
There is significant room for improvement in the efficiency of the implementation. The most expensive inefficiency is that we do not call MOSEK directly. Nonetheless, we are still able to solve the DMAOP over fine triangulations in an acceptable amount of time. It seems that we actually run into problems with numerical stability before the runtime of the algorithm on a standard computer becomes intolerable.

14.2. **Examples.** We will consider only examples in the plane. Furthermore, we will always take the target measure ν to be the measure whose support is the unit ball and having uniform density on its support. It is not difficult to consider other convex target domains or to consider non-uniform log-concave densities (the most prominent examples being Gaussian densities). However, the visualizations that follow are more intuitive in the case that the target measure has uniform density on its support. Recall that it is possible to solve the DMAOP even in situations where we have not guaranteed convergence.

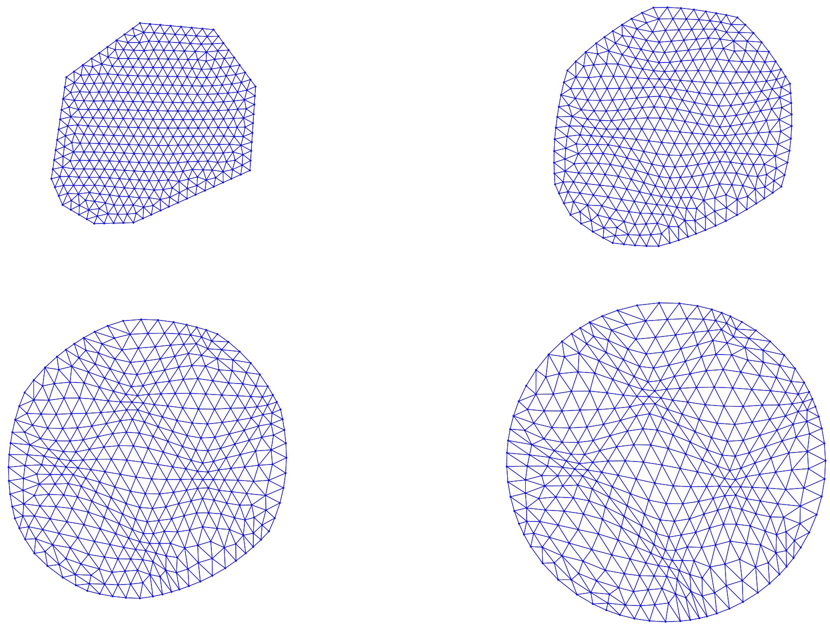
For our first example, consider the following triangulation of a convex polygonal source domain Ω. The shading in the background represents the density of f, though we understand that $f \equiv 0$ outside of Ω.



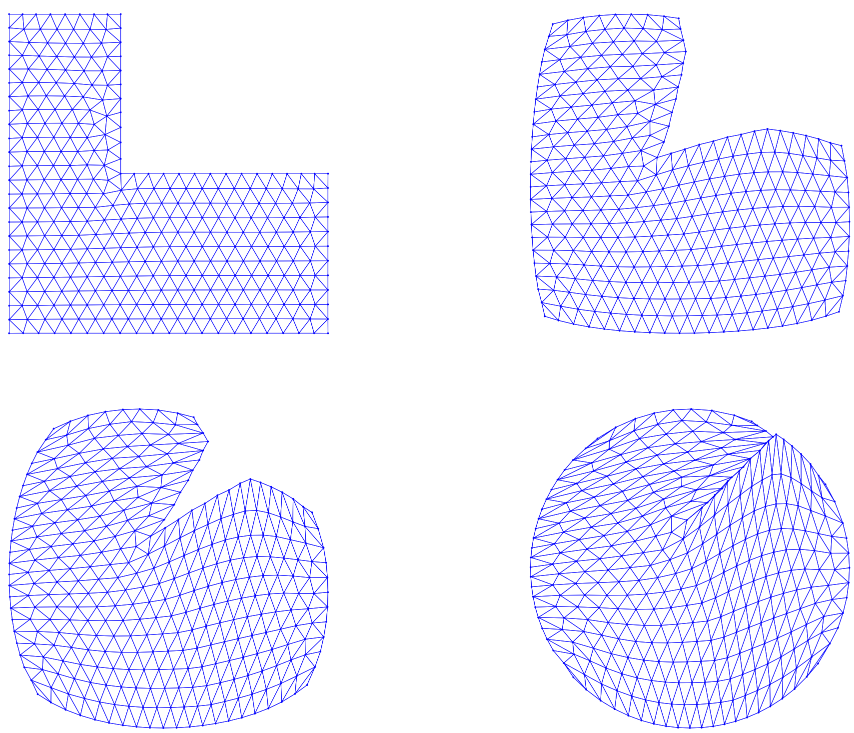
There are 405 points in the triangulation, and the DMAOP took 49.3 seconds to solve on a 2011 MacBook Pro. The following is a plot of the computed convex potential.



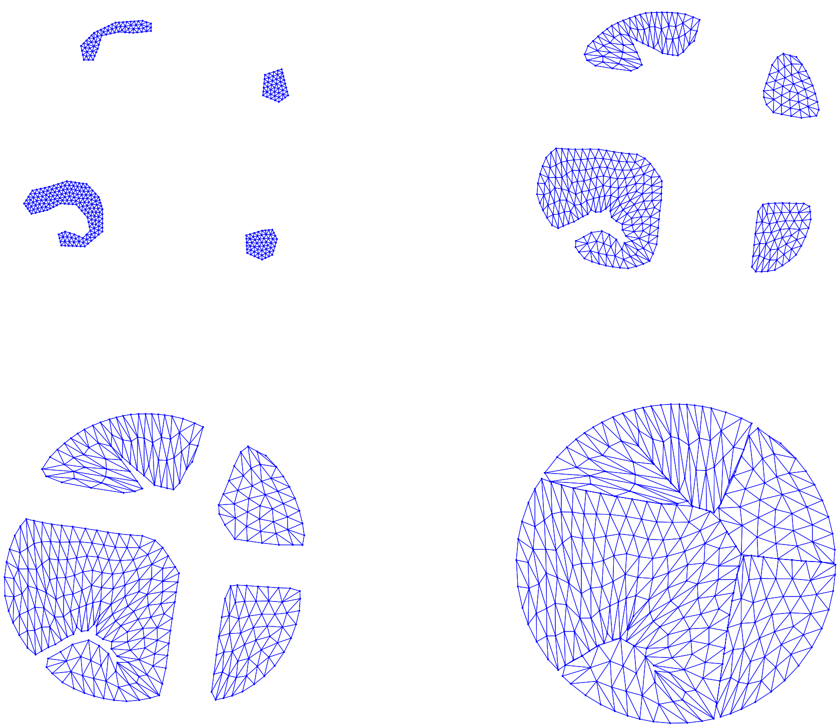
For every point x in the triangulation we can consider the interpolation $T_t(x) := (1-t)x + tT(x)$ for $t \in [0,1]$. We visualize this interpolation at times $t = 0, \frac{1}{3}, \frac{2}{3}, 1$. This interpolation can be understood as the solution of a dynamical optimal transport problem, though we will not discuss this fact further.



Next we consider a source measure with uniform density on a non-convex support. There are 340 points in the triangulation, and solving the DMAOP took 51.2 seconds.



Lastly we consider an example in which the source measure has highly irregular support (with uniform density on its support). There are 359 points in the triangulation, and solving the DMAOP took 58.9 seconds.



Notice that in the last two examples above, the inverse optimal maps are discontinuous. Nonetheless, we are able to approximate them by calculating the (continuous) forward maps and then inverting. Our method is particularly effective for highlighting the discontinuity sets of these inverse maps.

Part 3. Numerical optimal transport via linear elliptic PDE

We now present another method for solving optimal transport problems numerically. Much of the following will be largely formal (though we indicate the assumptions that are made); making the following arguments fully rigorous is an ongoing project. However, we find the arguments suggestive enough–and the associated numerical results compelling enough–to present them here. Even absent a proof that our numerical method must work under certain conditions, it is possible for any given test case to provide strong evidence that it is working (or not working).

Our method stems from answering the following natural question: as we take a one-parameter variation of source and target densities, how does the associated convex potential change? We will see that the 'time'-derivative of the convex potential is in fact (given sufficient regularity) the solution of a second-order linear elliptic PDE. Solving many of these PDEs allows us to 'flow' the solution of one optimal transport problem to another. By choosing an 'easy problem' (with known solution) to begin with, we can flow an easy solution to a hard one!

15. Derivation of the numerical method

15.1. **Deriving a PDE for the time derivative of the convex potential.** Suppose that we know the optimal transport map, $T_0 = \nabla\phi_0$ from (Ω, μ_0) to (Λ, ν_0), where μ_0 and ν_0 are given by Hölder -continuous densities $f_0 : \overline{\Omega} \to \mathbb{R}$ and $g_0 : \overline{\Lambda} \to \mathbb{R}$, respectively, (both positive and bounded away from zero and infinity). Suppose that we desire to know the optimal transport map from (Ω, μ) to (Λ, ν), where $\mu_1 := \mu$ and $\nu_1 := \nu$ are given by densities $f_1 = f$ and $g_1 = g$, respectively (also positive and bounded away from zero and infinity). Define $f_t = (1-t)f_0 + tf_1$ and $g_t = (1-t)g_0 + tg_1$ for $t \in [0, 1]$ (so f_t and g_t are both positive and bounded away from zero for all t). Let $T_t = \nabla\phi_t$ be the (unique) optimal transport map from (Ω, μ_t) to (Λ, ν_t) for all $t \in [0, 1]$, where the existence and uniqueness of such maps follow from Brenier's theorem. Let $\phi = \phi_1$ and $T = T_1$. If we assume that $\nabla\phi_t$ is C^1 (which, as we have seen, can be guaranteed in problems of sufficient regularity), then the following Monge-Ampère equation holds on Ω, for all $t \in [0, 1]$:

$$\det \nabla^2\phi_t = \frac{f_t}{g_t \circ \nabla\phi_t},$$

or equivalently, since the RHS is positive,

$$\log\det \nabla^2\phi_t - \log f + \log(g_t \circ \nabla\phi_t) = 0.$$

Next we take the derivative with respect to t (assuming that ϕ_t is sufficiently smooth as a function of both x and t). Recalling that for an invertible square matrix A depending on u, $\frac{\partial}{\partial u}\log\det A = \operatorname{tr}\left(A^{-1}\frac{\partial A}{\partial u}\right)$, we obtain:

$$\operatorname{tr}\left(\left(\nabla^2\phi_t\right)^{-1}\left(\frac{\partial}{\partial t}\nabla^2\phi_t\right)\right) - \frac{1}{f_t}\frac{\partial f_t}{\partial t} + \frac{1}{g_t \circ \nabla\phi_t}\cdot\left(\frac{\partial g_t}{\partial t}\circ\nabla\phi_t + \sum_{i=1}^{n}\left(\frac{\partial g_t}{\partial x_i}\circ\nabla\phi_t\right)\cdot\frac{\partial^2\phi_t}{\partial t\partial x_i}\right) = 0.$$

Define H_t^{ij} by $(H_t^{ij}) = \left(\nabla^2\phi_t\right)^{-1}$, and let $\eta_t = \frac{\partial}{\partial t}\phi_t$, so then we have

$$\sum_{i,j=1}^{n} H_t^{ij}\frac{\partial\eta_t}{\partial x_i}\frac{\partial\eta_t}{\partial x_j} - \frac{1}{f_t}\frac{\partial f_t}{\partial t} + \frac{1}{g_t\circ\nabla\phi_t}\cdot\left(\frac{\partial g_t}{\partial t}\circ\nabla\phi_t + \sum_{i=1}^{n}\left(\frac{\partial g_t}{\partial x_i}\circ\nabla\phi_t\right)\cdot\frac{\partial\eta_t}{\partial x_i}\right) = 0$$

Let D_t be the differential operator given by

$$D_t = \sum_{i,j=1}^{n} H_t^{ij}\frac{\partial}{\partial x_i}\frac{\partial}{\partial x_j} + \frac{1}{g_t\circ\nabla\phi_t}\sum_{i=1}^{n}\left(\frac{\partial g_t}{\partial x_i}\circ\nabla\phi_t\right)\frac{\partial}{\partial x_i},$$

so our earlier equation is

$$D_t\eta_t = \frac{1}{f_t}\frac{\partial f_t}{\partial t} - \frac{1}{g_t\circ\nabla\phi_t}\cdot\frac{\partial g_t}{\partial t}\circ\nabla\phi_t.$$

By our definition of f_t and g_t, this is the same as

$$D_t\eta_t = \frac{f_1 - f_0}{f_t} + \frac{g_0 - g_1}{g_t}\circ\nabla\phi_t. \tag{42}$$

Since (H^{ij}) is the inverse of a positive definite matrix, (H^{ij}) is positive definite, and it follows that D_t is elliptic, and (42) is a second-order linear elliptic PDE.

15.2. **Discussion and a preview of our numerical approach.** Our above reasoning immediately suggests (modulo the specification of boundary conditions for (42), which we shall discuss below) a method for numerically solving the optimal transport problem from (Ω, μ) to (Λ, ν). First we choose a time-step h. Then, given ϕ_t, we approximate $\phi_{t+h} = \phi_t + h\eta_t$, where η_t is the solution to (42). If we begin with knowledge of ϕ_0, then we can repeat the forward-stepping until we obtain an approximation for ϕ_1.

In practice, how can we initialize f_0 and g_0 so that the associated convex potential ϕ_0 is known to us? If μ and ν have the same support, i.e., $\Omega = \Lambda$, then we can simply set $f_0 = g_0 = \frac{1}{\mathrm{vol}(\Omega)}$ on $\Omega = \Lambda$, so $\phi_0(x) = \frac{1}{2}\|x\|^2$ (and $T_0 = \nabla\phi_0 = \mathrm{id}$).

In fact, we can (almost) do this even when $\Omega \neq \Lambda$. We can take a domain D such that $\Omega, \Lambda \subset D$ and set $f_0 = g_0 = \frac{1}{\mathrm{vol}(D)}\chi_D$. Even though $f = f_1$ and $g = g_1$ may vanish on D, for every $t \in [0,1)$ it is in fact true that f_t and g_t are bounded away from zero on D. Of course, ϕ_t may become singular at $t = 1$, but for $t < 1$, ϕ_t is nice. As $t \to 1$, the stability of optimal transport tells us that T_t should approach $T = \nabla\phi$ in a certain sense, so we should still be able to approximate solutions even in this case.

Note that alternatively, for $\varepsilon > 0$ we could take $f_\varepsilon := (1 + \varepsilon \cdot \mathrm{vol}(D))^{-1}(f + \varepsilon \cdot \chi_D)$ and $g_\varepsilon := (1 + \varepsilon \cdot \mathrm{vol}(D))^{-1}(g + \varepsilon \cdot \chi_D)$ (both bounded away from zero on D) and then use our forward-stepping scheme to find ϕ_ε (solving the optimal transport problem associated with f_ε and g_ε). Then by the stability of optimal transport we see that for $\varepsilon > 0$ small, $\nabla\phi_\varepsilon$ approximates $\nabla\phi$ in a certain sense.

In practice, we find that is is numerically easier to deal with a vanishing source density than a vanishing target density, though the latter can be dealt with by the methods just described. We also find that it is easier to deal with a singular source density than a singular target density. This is intuitive: spacial derivatives of the target density appear in the coefficients of the operator D_t, while spacial derivatives of f_t do not.

There is another way to obtain f_0, g_0, and ϕ_0: we can leverage our previous method, the DMAOP. To do so, we set f_0 and g_0 to be constant on Ω and Λ, respectively, and vanishing everywhere else. If Ω and Λ are convex (or, more generally, if Λ is convex and Ω is such that the associated convex potential ϕ_0 is in $C^{2,\alpha}(\overline{\Omega})$), then we can compute ϕ_0 by the DMAOP (note that g_0 is quite trivially log-concave).

We will not pursue this approach here. We will instead follow the first approach, i.e., consider a domain D containing Ω and Λ. By scaling and translation, we can assume without loss of generality that Ω and Λ are contained in $[-1,1]^n$. This is effectively the same as assuming that $\Omega = \Lambda = [-1,1]^n$ with f and g not necessarily bounded away from zero (and possibly vanishing) on $\Omega = \Lambda$.

By solving our problem on a cube, we can easily employ finite difference techniques for elliptic PDEs. Note that in order to consider an arbitrary domain D, or to follow the approach that leverages the DMAOP, we could use finite element techniques on a suitable triangulation of the domain in question.

15.3. **Deriving the boundary conditions.** What boundary conditions do the η_t satsify? Notice that if Ω and Λ are both convex, then Caffarelli's regularity theory gives that $\nabla\phi_t$ and its inverse $\nabla\phi_t^*$ are in $C^{0,\alpha}(\overline{\Omega})$ and $C^{0,\alpha}(\overline{\Lambda})$, respectively, so $\nabla\phi_t$ is a homeomorphism $\overline{\Omega} \to \overline{\Lambda}$, and hence $\nabla\phi_t(\partial\Omega) = \partial\Lambda$. Then for $x \in \partial\Omega$, we have that $\nabla\phi_t(x) \in \partial\Lambda$ for all t. Consequently, if $\partial\Omega$ and $\partial\Lambda$ are smooth at x, then $\frac{\partial}{\partial t}\nabla\phi_t(x)$ must be tangent to $\partial\Lambda$ at $\nabla\phi_t(x)$, i.e.,

$$\langle \nabla\eta_t(x), \nu\left(\nabla\phi_t(x)\right)\rangle = 0 \tag{43}$$

for all $x \in \partial\Omega$, where $\nu(y)$ is the outward-pointing normal vector to $\partial\Lambda$ at $y \in \partial\Lambda$.

Notice that this yields a Neumann-like (though not precisely Neumann) boundary condition for η_t. However, in the case that $\Omega = \Lambda = [-1,1]^n$, this boundary condition is in fact a Neumann boundary condition (on the interiors of the faces of $[-1,1]^n$).

Lemma 43. *Let $\Omega = \Lambda = [-1,1]^n$. Then $\nabla\phi_t$ preserves the interiors of the faces of $[-1,1]^n$.*

Proof. Since f_t and g_t are bounded away from zero on $[-1,1]^n$, Caffarelli's regularity theory gives that $\phi_t \in C^{2,\alpha}(\Omega) \cap C^{1,\alpha}(\overline{\Omega})$ and that ϕ_t is strictly convex on $[-1,1]^n$. Suppose for contradiction that x and $\nabla\phi_t(x)$ are in the interiors of two distinct faces (F_1 and F_2, respectively) of $[-1,1]^n$.

There are two cases: either F_1 and F_2 are opposite faces, or they are not. In the latter case, there is a vector $v \in \mathbb{R}^n$ which is parallel to F_1 and perpendicular to F_2. By the continuity of $\nabla\phi_t$ on $\overline{\Omega}$ (and recalling that $\nabla\phi_t(\partial\Omega) = \partial\Lambda$), there exists $\delta > 0$ such that $\nabla\phi_t(B_{2\delta}(x) \cap \partial\Omega) \subset F_2$. Then we have that $\nabla\phi_t(x)$, $\nabla\phi_t(x+\delta v) \in F_2$, so $\nabla\phi_t(x+\delta v) - \nabla\phi_t(x)$ is perpendicular to v, hence perpendicular to $\delta v = (x+\delta v) - x$. In other words, with $y := x + \delta v$, we have shown that $\langle y - x, \nabla\phi_t(y) - \nabla\phi_t(x)\rangle = 0$.

Now by the strict convexity of $\nabla\phi_t$ on $[-1,1]^n$, we have that

$$\begin{cases} \phi_t(y) > \phi_t(x) + \langle y - x, \nabla\phi_t(x)\rangle \\ \phi_t(x) > \phi_t(y) + \langle x - y, \nabla\phi_t(y)\rangle, \end{cases}$$

i.e.,

$$\begin{cases} 0 > \phi_t(x) - \phi_t(y) + \langle y - x, \nabla\phi_t(x)\rangle \\ 0 > \phi_t(y) - \phi_t(x) + \langle y - x, -\nabla\phi_t(y)\rangle. \end{cases}$$

Adding these two inequalities, we obtain $\langle y - x, \nabla\phi_t(x) - \nabla\phi_t(y)\rangle < 0$, a contradiction.

It remains to consider the case in which F_1 and F_2 are opposite. Let e_k be the standard unit vector that is perpendicular to both F_1 and F_2. Then $\langle x, e_k\rangle = \pm 1$, and $\langle \nabla\phi_t(x), e_k\rangle = \mp 1$.

Define γ by $\gamma(t) = x + te_k$. Then by the strict convexity of ϕ_t, $\frac{\partial\phi_t}{\partial x_k} \circ \gamma$ is strictly increasing. Then for any $t > 0$, we have that

$$\frac{\partial\phi_t}{\partial x_k}(x - te_k) < \frac{\partial\phi_t}{\partial x_k}(x) = \mp 1$$

and, similarly,

$$\frac{\partial\phi_t}{\partial x_k}(x + te_k) > \frac{\partial\phi_t}{\partial x_k}(x) = \mp 1.$$

Thus if $\langle x, e_k\rangle = 1$, then $\frac{\partial \phi_t}{\partial x_k}(y) < -1$ for some $y \in (-1,1)^n$, and if $\langle x, e_k\rangle = -1$, then $\frac{\partial \phi_t}{\partial x_k}(y) > 1$ for some $y \in (-1,1)^n$. In either case, it follows that $\nabla\phi_t(y) \notin [-1,1]^n$ for some $y \in [-1,1]^n$, giving a contradiction.

□

Then since ν is constant on the interior of eaach face of $[-1,1]^n$, our boundary condition is that $\langle \nabla\eta_t(x), \nu(x)\rangle = 0$ for all x in the interiors of the faces of $[-1,1]^n$. This is of course a Neumann boundary condition. For x in a face that is perpendicular to e_k, our boundary condition dictates that $\frac{\partial \eta_t}{\partial x_k} = 0$.

Notice that instead of providing a boundary condition, we might choose to solve (42) subject to the constraint that $\nabla(\phi_t + h\eta_t)(\partial\Omega) \subset \Lambda$. Numerically, this would amount to solving a linear system subject to convex inequality constraints. This can easily be cast as a convex optimization problem that is feasible to solve. In the case that $\Omega = \Lambda = [-1,1]^n$, the constraints are actually affine, which makes the problem much easier. However, it is still more desirable to solve (42) with boundary conditions because this amounts to solving a single linear system, computationally much cheaper than solving an inequality-constrained optimization problem. The computational savings add up because we must solve (42) at each time step, and a small step size is necessary for an accurate approximation of $\phi = \phi_1$ and $T = T_1$.

16. Details of the numerical scheme

We restrict our attention to the case $n = 2$ (though the generalization to arbitrary dimension is clear) with $\Omega = \Lambda = [-1,1]^2$. Then with the discussion from the preceding section in mind, given ϕ_t we aim to solve

$$\begin{aligned} D_t\eta_t(x) &= \tfrac{f_1(x)-f_0(x)}{f_t(x)} + \tfrac{g_0(\nabla\phi_t(x))-g_1(\nabla\phi_t(x))}{g_t(\nabla\phi_t(x))}, & x = (x_1,x_2) \in (-1,1)^2 \\ \tfrac{\partial}{\partial x_i}\eta_t &= 0, & |x_i| = 1 \end{aligned}$$

where

$$D_t = \sum_{i,j=1}^{2} H_t^{ij} \frac{\partial}{\partial x_i}\frac{\partial}{\partial x_j} + \frac{1}{g_t \circ \nabla\phi_t}\sum_{i=1}^{2}\left(\frac{\partial g_t}{\partial x_i} \circ \nabla\phi_t\right)\frac{\partial}{\partial x_i}$$

with H_t^{ij} the (i,j)-th entry of $\left(\nabla^2\phi_t\right)^{-1}$.

We use a finite difference scheme on $[-1,1]^2$. Fix a positive integer N. Let $x^{(m,n)} = \left(-1 + \frac{2m}{N}, -1 + \frac{2n}{N}\right)$ for $m, n = 0, 1, \ldots, N$. The $x^{(m,n)}$ will be the grid points of our finite difference scheme. Then we have a grid spacing of $d = \frac{2}{N}$.

Furthermore, let M be a positive integer, and let $t_k = \frac{k}{M}$ for $k = 0, \ldots, M$. The t_k reflect our discretization of the time interval $[0,1]$. Note that $h = \frac{1}{M}$ is our time step size.

Inductively, fix $k < M$ and suppose that we have $\phi_k^{(m,n)}$ for all $m, n = 0, 1, \ldots, N$. We understand$\phi_k^{(m,n)}$ as our approximation of $\phi_{t_k}\left(x^{(m,n)}\right)$. Let Φ_k denote the matrix $\left(\phi_k^{(m,n)}\right)$. Then to completely specify our method, we need only specify how to compute $\Phi_{k+1} = \left(\phi_{k+1}^{(m,n)}\right)$.

For a matrix $Z = (\zeta^{(m,n)})$ and for $m, n = 1, \ldots, N-1$, let

$$\begin{aligned}
(\Delta_1 Z)^{(m,n)} &= \frac{Z^{(m+1,n)} - Z^{(m-1,n)}}{2h} \\
(\Delta_2 Z)^{(m,n)} &= \frac{Z^{(m,n+1)} - Z^{(m,n-1)}}{2h} \\
(\Delta_{11} Z)^{(m,n)} &= \frac{Z^{(m+1,n)} + Z^{(m-1,n)} - 2Z^{(m,n)}}{2h^2}, \\
(\Delta_{22} Z)^{(m,n)} &= \frac{Z^{(m,n+1)} + Z^{(m,n-1)} - 2Z^{(m,n)}}{2h^2}, \\
(\Delta_{12} Z)^{(m,n)} &= \frac{Z^{(m+1,n+1)} + Z^{(m-1,n-1)} - Z^{(m+1,n-1)} - Z^{(m-1,n+1)}}{4h^2}.
\end{aligned}$$

It is well-known that Δ_i is the standard second-order finite difference operator corresponding to $\frac{\partial}{\partial x_i}$ and that Δ_{ij} is the standard second-order finite difference operator corresponding to $\frac{\partial^2}{\partial x_i \partial x_j}$.

For $m, n = 1, \ldots, N-1$, let $I_k^{(m,n)}$ be the matrix with (i,j)-th entry $(\Delta_{ij}\Phi_k)^{(m,n)}$, and let $H_k^{(m,n)} = \left(I_k^{(m,n)}\right)^{-1}$. Then let $H_k^{ij;(m,n)}$ be the (i,j)-th entry of $H_k^{(m,n)}$, and let $\mathcal{H}_k^{ij}$ be the multiplication operator defined by

$$\mathcal{H}_k^{ij}\left(Z^{(m,n)}\right)_{m,n=1}^{N-1} = \left(H_k^{ij;(m,n)} Z^{(m,n)}\right)_{m,n=1}^{N-1}.$$

Furthermore, for $m, n = 1, \ldots, N-1$, let

$$G_k^{(m,n)} = \left[(\Delta_1 \Phi_k)^{(m,n)}, (\Delta_2 \Phi_k)^{(m,n)}\right],$$

and let

$$M_k^{i;(m,n)} = \left(g_{t_k}\left(G_k^{(m,n)}\right)\right)^{-1} \cdot \frac{\partial g_{t_k}}{\partial x_i}\left(G_k^{(m,n)}\right).$$

Then let $\mathcal{M}_k^i$ be the multiplication operator defined by

$$\mathcal{M}_k^i\left(Z^{(m,n)}\right)_{m,n=1}^{N-1} = \left(M_k^{i;(m,n)} Z^{(m,n)}\right)_{m,n=1}^{N-1}.$$

Then define

$$\mathcal{D}_k = \sum_{i,j=1}^{2} \mathcal{H}_k^{ij} \Delta_{ij} + \sum_{i=1}^{2} \mathcal{M}_k^i \Delta_i.$$

We understand $\mathcal{D}_k$ as a discretization of the differential operator D_{t_k}.

Next, for $m, n = 1, \ldots, N-1$, let

$$B_k^{(m,n)} = \frac{f_1 - f_0}{f_{t_k}}\left(x^{(m,n)}\right) + \frac{g_0 - g_1}{g_{t_k}}\left(G_k^{(m,n)}\right).$$

Then our discretized equation for $\eta_k = \left(\eta_k^{(m,n)}\right)_{m,n=1}^{N-1}$, modulo the inclusion of boundary conditions, is $\mathcal{D}_k \eta_k = B_k$.

It remains to include boundary conditions in our linear system. For $m, n = 0, \ldots, N$, let

$$\begin{aligned}
(\Delta_1^{(+)} Z)^{(n)} &= \frac{3Z^{(N,n)} - 4Z^{(N-1,n)} + Z^{(N-2,n)}}{2h} \\
(\Delta_1^{(-)} Z)^{(n)} &= \frac{-3Z^{(0,n)} + 4Z^{(1,n)} - Z^{(2,n)}}{2h} \\
(\Delta_2^{(+)} Z)^{(m)} &= \frac{3Z^{(m,N)} - 4Z^{(m,N-1)} + Z^{(m,N-2)}}{2h} \\
(\Delta_2^{(-)} Z)^{(m)} &= \frac{-3Z^{(m,0)} + 4Z^{(m,1)} - Z^{(m,2)}}{2h}.
\end{aligned}$$

$\Delta_i^{(+)}$ simply takes the second-order backward finite difference approximation of $\frac{\partial}{\partial x_i}$ along the boundary $\{x_i = 1\}$, while $\Delta_i^{(-)}$ takes the second-order forward finite difference approximation of $\frac{\partial}{\partial x_i}$ along the boundary $\{x_i = -1\}$.

Let $\Delta_{\text{bd}} = \left(\Delta_1^{(+)}, \Delta_1^{(-)}, \Delta_2^{(+)}, \Delta_2^{(-)}\right)^T$. Then our discretized boundary condition is $\Delta_{\text{bd}} \eta_k = 0$. Then our full linear system for $\eta_k = \left(\eta_k^{(m,n)}\right)_{m,n=1}^{N-1}$ is

$$\mathcal{D}_k \eta_k = B_k, \ \Delta_{\text{bd}} \eta_k = 0.$$

Treating η_k as a vector in $\mathbb{R}^{(N+1)^2}$, there exists $A_k \in \mathbb{R}^{\left[(N+1)^2 + 4(N+1)\right] \times (N+1)^2}$, $b_k \in \mathbb{R}^{(N+1)^2 + 4(N+1)}$ yielding an equivalent system

$$A_k \eta_k = b_k.$$

Notice that A_k is quite sparse by construction.

We solve this system in the least-squares sense, i.e., we find $\eta_k \in \mathbb{R}^{(N+1)^2}$ minimizing $\|A_k \eta_k - b_k\|_2^2$. As is well-known, such a minimizer is obtained by solving the normal equations:

$$\left(A_k^T A_k\right) \eta_k = A_k^T b_k.$$

The numerical solution of the normal equations can be conducted quite rapidly.

Let η_k solve the normal equations, and then consider η_k as a $(N+1) \times (N+1)$ matrix. Then we set $\Phi_{k+1} = \Phi_k + h\eta_k$. We have now fully specified our numerical scheme.

We note that when $\Phi_M = \left(\phi_M^{(m,n)}\right)$ is attained, we consider the $\phi_M^{(m,n)}$ to be the approximate values of $\phi\left(x^{(m,n)}\right)$ for $m, n = 0, 1, \ldots, N$. Furthermore, we can use the finite difference operators described above to compute $T^{(m,n)} \approx \nabla\phi\left(x^{(m,n)}\right)$ for $m, n = 0, 1, \ldots, N$.

Also, there is a simple procedure for obtaining the convex potential and optimal transport map associate with the 'inverse' optimal transport problem with source ν and target μ. Recall that the convex potential ϕ^* solves the inverse problem. Now when $x^* = \nabla\phi(x)$, we have that $\phi^*(x^*) = \phi(x) - \langle x, x^* \rangle$. Then we approximate ϕ^* at $T^{(m,n)}$ by $\phi_M^{(m,n)} - \left\langle x^{(m,n)}, T^{(m,n)} \right\rangle$, and we approximate $\nabla\phi^*$ at $T^{(m,n)}$ by $x^{(m,n)}$. We can obtain a value for ϕ^* or $\nabla\phi^*$ anywhere by taking the Delaunay triangulation of the points $T^{(m,n)}$ and employing barycentric interpolation.

17. An account of theoretical debts and a program for their fulfillment

We will outline some desired results that would be the main steps toward achieving rigorous results pertaining to our numerical method. First, we desire a result that guarantees (in some sense) the smoothness of the convex potential with respect to perturbations of f and g. We expect that such a result exists, and since our variations f_t and g_t are 'nice' in every imaginable sense, we conjecture with some optimism that $\eta_t = \frac{\partial}{\partial t}\phi_t$ indeed exists and that exchanging the spatial and time derivatives of ϕ_t as done earlier is justified. If true, these facts would indicate that a solution to (42) with boundary condition (43) exists.

Next we conjecture that such a solution is in fact unique. We hope that such a result can be obtained from a modification of the standard technique (using the Hopf boundary point lemma and strong maximum principle) for proving the uniqueness of solutions of elliptic equations with Neumann boundary conditions.

Uniqueness is indeed important for our purposes because when we take a step using our forward-stepping scheme, we want to be sure that the step is in the right direction!

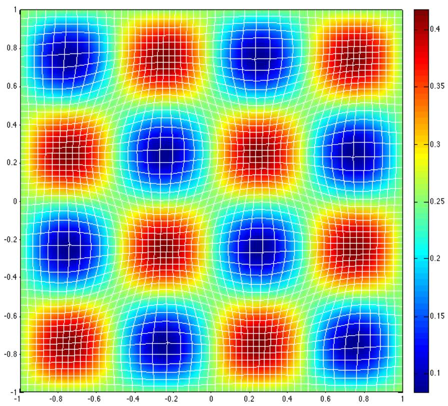
We remark that for existence and uniqueness it may be easiest to assume that $\partial\Omega$ and $\partial\Lambda$ are smooth, but it would also be of interest to examine the case of the cube $[-1, 1]^n$ in detail.

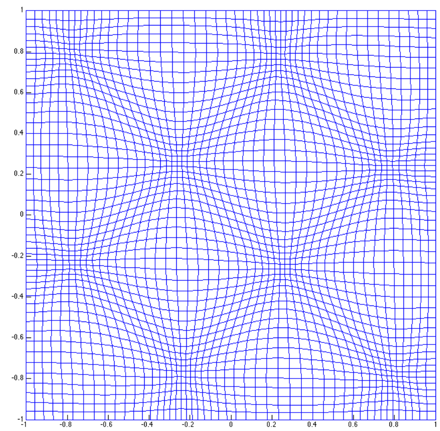
The last major theoretical step remaining would be to bound the error of our numerical solution to (42) with boundary condition (43). The standard techniques used in the numerical analysis of linear elliptic PDEs would be of use, but special attention must be paid to the fact that the coefficients of the differential operator D_t actually change with time. Even worse, these coefficients depend on ϕ_t, so we can only approximate the coefficients using our approximation for ϕ_t. This fact adds significant intricacy to the numerical analysis of our technique.

18. Numerical experiments

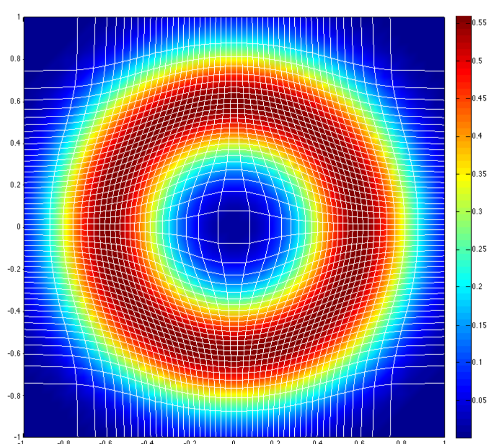
We present several numerical examples. Computation time for each example was less than a minute.

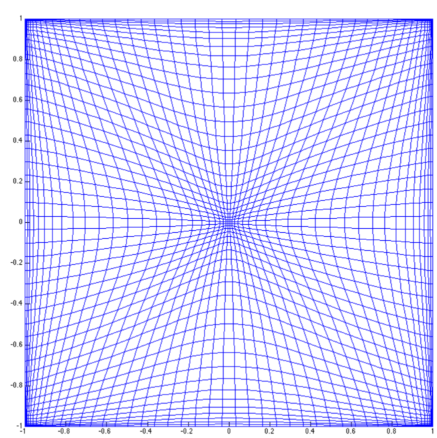
We first consider the optimal map from the density pictured at left below to the uniform density on the square. The lines in the image at right are the image of a grid on the square under the forward optimal map. The lines in the image at left are the image of a grid on the square under the inverse optimal map.



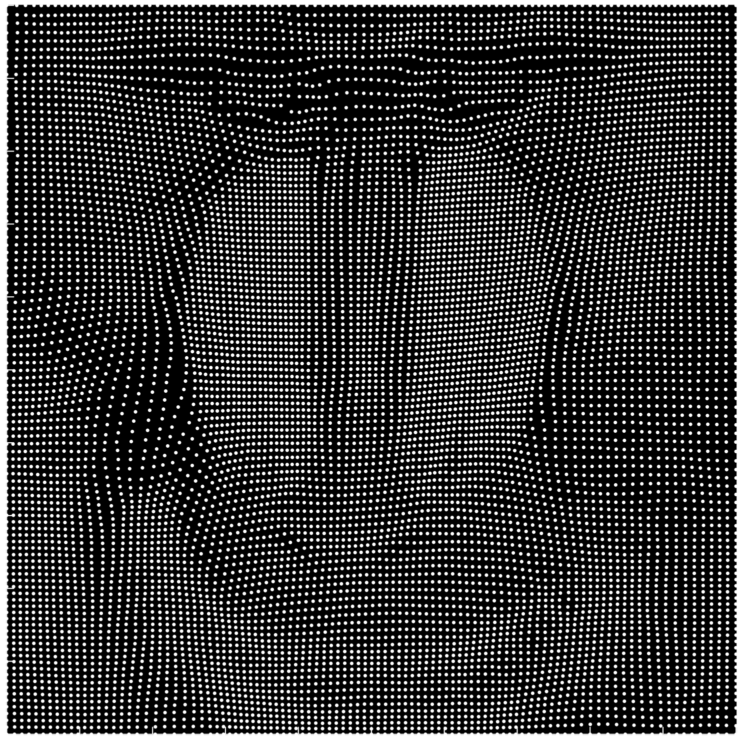


The following two images should be interpreted in the same way as the preceding two.





We let the source density be the famous ‘mandrill image’ and compute the optimal transport map to the uniform density. Then we take the image of grid points under the inverse map to obtain a stippled depiction of the mandrill (when viewing on a computer, zoom in sufficiently to prevent aliasing!).



Acknowledgments

I am pleased to acknowledge and thank my optimal transport friends Yanir Rubinstein, Otis Chodosh, Lyuboslav Panchev, and Vishesh Jain. Since working with them during the summer of 2012, each of these brilliant mathematicians has been a valuable mentor to me. In particular, over the course of the past year, Yanir has encouraged and advised me as I have worked out the ideas in this thesis, and meanwhile Otis has provided valuable written and verbal commentary. I am very grateful to Rafe Mazzeo for advising me and overseeing my reading throughout this year, as well as for being flexible with regard to the changing direction of my thesis. I also thank Akshay Venkatesh for being an unfailingly encouraging presence in the math department for me over the last four years.

Furthermore, I acknowledge financial support from Y.A. Rubinstein's NSF grant DMS-1206284 during the years 2012-2015, as well as from SURIM for funding research during Summer 2012 and from OVPUE for travel to the 2013 Joint Mathematics Meetings in San Diego.

Appendices: Definitions and results from convex analysis

Convex analysis plays a significant role in the theory of optimal transport, perhaps most clearly in the case of the quadratic cost in Euclidean space. We review here several definitions and results from convex analysis that are useful in our exposition as well as in optimal transport more broadly. Much (but not all) of the following is adapted from Rockafellar [33].

A. Convex sets and functions

Definition 44. *A set $C \subset \mathbb{R}^n$ is* convex *if $(1-t)x + ty \in C$ for every $x, y \in C$ and all $t \in [0,1]$.*

Definition 45. *Let $S \subset \mathbb{R}^n$, and let $f : S \to [-\infty, \infty]$. The* epigraph *of f is the set*

$$\{(x,u) \in S \times \mathbb{R} \,|\, u \geq f(x)\},$$

which we denote by $\operatorname{epi} f$*. We say that f is a* convex function *on S if $\operatorname{epi} f \subset \mathbb{R}^{n+1}$ is convex. When we say that f is convex, we take this to mean that f is convex on $\mathbb{R}^n$. The* effective domain *of a convex function f on S, denoted* $\operatorname{dom} f$*, is the set*

$$\operatorname{dom} f = \{x \in S \,|\, \exists u \in \mathbb{R},\, (x,u) \in \operatorname{epi} f\} = \{x \in S \,:\, f(x) < +\infty\}.$$

The following is an immediate consequence of the preceding definitions:

Lemma 46. *Let f be convex on $S \subset \mathbb{R}^n$. Then* $\operatorname{dom} f$ *is convex.*

The following theorem relates our epigraphical definition of convex functions with a more commonly used, and perhaps more familiar, definition:

Theorem 47. *Let $C \subset \mathbb{R}^n$ be a convex set (e.g., $C = \mathbb{R}^n$). $f : C \to (-\infty, +\infty]$ is* convex on C *if and only if $f\left((1-t)x + ty\right) \leq (1-t)f(x) + tf(y)$ for all $x, y \in C$ and $t \in [0,1]$.*

Proof. Evidently f is convex on C if and only if $(1-t)(x,u) + t(y,v) \in \operatorname{epi} f$ (equivalently, $((1-t)x + ty, (1-t)u + tv) \in \operatorname{epi} f$) whenever $(x,u),(y,v) \in \operatorname{epi} f$ and $t \in [0,1]$. Thus f is convex on C if and only if $f\left((1-t)x + ty\right) \leq (1-t)u + tv$ whenever $(x,u),(y,v) \in \operatorname{epi} f$ and $t \in [0,1]$.

If $x, y \in C$ with $f(x), f(y) < +\infty$, then $(x, f(x)), (y, f(y)) \in \operatorname{epi} f$ (since $f(x), f(y) > -\infty$). Thus if f is convex on C, then $f\left((1-t)x + ty\right) \leq (1-t)f(x) + tf(y)$ for all $x, y \in \operatorname{dom} f$. But this inequality also holds trivially when $x \notin \operatorname{dom} f$ or $y \notin \operatorname{dom} f$, so the 'only if' direction holds.

Now whenever $(x,u),(y,v) \in \operatorname{epi} f$, it follows that $f(x) \leq u$ and $f(y) \leq v$. Then if we assume that $f\left((1-t)x + ty\right) \leq (1-t)f(x) + tf(y)$ for all $x, y \in C$ and $t \in [0,1]$, we also have that $f\left((1-t)x + ty\right) \leq (1-t)u + tv$ for all $(x,u),(y,v) \in \operatorname{epi} f$ and $t \in [0,1]$, and the 'if' direction holds. ☐

We note that this theorem provides an alternative definition of convexity for functions that do not take the value $-\infty$. We will in fact only need to consider such functions.

We note that when $f \in C^2(C)$, our definition of convexity coincides with the definition one might learn in multivariate calculus:

Theorem 48. *Let $f \in C^2(C)$, where $C \subset \mathbb{R}^n$ is open and convex. Then f is convex on C if and only if the Hessian matrix $\nabla^2 f(x)$ is positive semi-definite for all $x \in C$.*

Proof. See Theorem 4.5 of Rockafellar [33]. □

Notice that for f convex on a convex set $C \subset \mathbb{R}^n$, we can extend to $\tilde{f}$ to $\mathbb{R}^n$ by defining $\tilde{f}|_{\mathbb{R}^n \setminus C} \equiv +\infty$. It is clear that $\operatorname{epi} f = \operatorname{epi} \tilde{f}$, so $\tilde{f}$ is convex on $\mathbb{R}^n$. Thus *we lose no generality by considering only functions that are convex on $\mathbb{R}^n$.*

The following definitions are helpful for ruling out pathologies:

Definition 49. *A convex function f is called* proper *if* $\operatorname{dom} f \neq \emptyset$ *and $f(x) > -\infty$ for all x.*

We will only ever need to consider proper convex functions.

Definition 50. *A proper convex function f is* closed *if* $\operatorname{epi} f$ *is closed. A non-proper convex function f is closed if either $f \equiv +\infty$ or $f \equiv -\infty$.*

Proposition 51. *If f is a proper convex function, then f is closed if and only if it is lower semi-continuous.*

Proof. First recall that a function g is lower semi-continuous (lsc) if and only if $\{x : g(x) > \alpha\}$ is open for all $\alpha \in \mathbb{R}$, or equivalently, if and only if $\{y : g(y) \leq \alpha\}$ is closed for all $\alpha \in \mathbb{R}$. Suppose that f is a proper convex function. Now let $x_k \in \mathbb{R}^n$ be a sequence with $f(x_k) \leq \alpha$ for some $\alpha \in \mathbb{R}$ and with $x_k \to x$ for some $x \in \mathbb{R}^n$. Notice that $(x_k, \alpha) \in \operatorname{epi} f$ for all k. Since f is closed, $\operatorname{epi} f$ is closed, and therefore $(x, \alpha) \in \operatorname{epi} f$, i.e., $f(x) \leq \alpha$. Thus $\{y : f(y) \leq \alpha\}$ is closed, and f is lsc.

Conversely, suppose that f is lsc, so $\{y : f(y) \leq \beta\}$ is closed for all $\beta \in \mathbb{R}$. Let $(x_k, \alpha_k) \in \operatorname{epi} f$ (so $f(x_k) \leq \alpha_k$) with $x_k \to x$ and $\alpha_k \to \alpha$. Let $\delta > 0$. Then for all k sufficiently large, $\alpha_k \leq \alpha + \delta$, so $f(x_k) \leq \alpha + \delta$, i.e., $x_k \in \{y : f(y) \leq \alpha + \delta\}$ for all k sufficiently large. Then by the closure of this set, we have that $f(x) \leq \alpha + \delta$. Since $\delta > 0$ was arbitrary, in fact we have that $f(x) \leq \alpha$, and $(x, \alpha) \in \operatorname{epi} f$. □

The convexity of a function guarantees its continuity in a certain sense:

Theorem 52. *A convex function f on $\mathbb{R}^n$ is continuous relative to any relatively open convex set in* $\operatorname{dom} f$. *In particular, f is continuous on* $\operatorname{int}\operatorname{dom} f$. *In fact, it holds that a proper convex function f is locally Lipschitz on* $\operatorname{int}\operatorname{dom} f$.

Proof. See Theorems 10.1 and 10.4 of Rockafellar [33]. □

B. First-order properties of convex functions

There is an extension of the notion of differentiability that is fundamental to the analysis of convex functions.

Definition 53. *Let f be a convex function on $\mathbb{R}^n$. $x^* \in \mathbb{R}^n$ is called a* subgradient *of f at x if $f(z) \geq f(x) + \langle x^*, z - x\rangle$ for all $z \in \mathbb{R}^n$. The* subdifferential *of f at x, denoted $\partial f(x)$, is the set of all subgradients of f at x.*

Theorem 54. *Let f be a proper convex function. $\partial f(x)$ is a non-empty bounded set if and only if $x \in \operatorname{int}\operatorname{dom} f$.*

Proof. See Theorem 23.4 of Rockafellar [33]. □

Using this theorem, we can prove the following famous inequality.

Theorem 55. *(Jensen's inequality.) Let f be a proper convex function on $\mathbb{R}^n$, and let μ be a probability measure. Suppose that $g : \mathbb{R}^m \to \mathbb{R}^n$ is μ-measurable and that* $\operatorname{supp}\mu \subset g^{-1}(\operatorname{int}\operatorname{dom} f)$. *Then $\int f \circ g \, d\mu \geq f\left(\int g \, d\mu\right)$.*

Proof. Let $y = \int g \, d\mu$. Now for all $x \in \operatorname{supp}\mu$, $g(x) \in \operatorname{int}\operatorname{dom} f$. The reader can easily verify that the interior of convex set is convex, so $\operatorname{int}\operatorname{dom} f$ is convex. It then follows (omitting some details) that the convex combination $y = \int g \, d\mu \in \operatorname{int}\operatorname{dom} f$.

Now by the preceding theorem, $\partial f(y)$ is nonempty, so we let $y^* \in \partial f(y)$. Then there exists $b \in \mathbb{R}$ such that $f(x) \geq \langle y^*, x\rangle + b$ for all x and $f(y) = \langle y^*, y\rangle + b$. In particular, it holds that $f(g(x)) \geq \langle y^*, g(x)\rangle + b$ for all $x \in \operatorname{dom} f$. Then since μ is supported on $\operatorname{dom} f$, we have that

$$\int f \circ g \, d\mu \geq \int \left(\langle y^*, g(x)\rangle + b\right) \, d\mu(x) = \left\langle y^*, \int g \, d\mu \right\rangle + b \int d\mu = \langle y^*, y\rangle + b = f(y),$$

as desired. □

It is perhaps no surprise that the derivative and the subdifferential of a convex function coincide wherever it is differentiable.

Theorem 56. *Let f be a convex function, and let $x \in \mathbb{R}^n$ such that $f(x)$ is finite. If f is differentiable at x, then $\nabla f(x)$ is the unique subgradient of f at x. Conversely, if f has a unique subgradient at x, then f is differentiable at x.*

Proof. See Theorem 25.1 of Rockafellar [33]. □

It is an important fact that proper convex functions are differentiable almost everywhere:

Theorem 57. *Let f be a proper convex function on $\mathbb{R}^n$, and let D be the set of points where f is differentiable. Then D is a dense subset of* $\operatorname{int}\operatorname{dom} f$, *and* $\operatorname{int}\operatorname{dom} f \backslash D$ *has measure zero. Furthermore, the gradient map ∇f given by $x \mapsto \nabla f(x)$ is continuous on D.*

Proof. See Theorem 25.5 of Rockafellar [33]. □

In fact, the subgradients of a proper convex function enjoy a kind of continuity.

Theorem 58. *If f is a proper convex function on $\mathbb{R}^n$, $x \in \operatorname{int}\operatorname{dom} f$, and $\varepsilon > 0$, then there exists $\delta > 0$ such that*

$$\partial f(z) \subset \partial f(x) + \varepsilon B$$

for all $z \in B_\delta(x)$.

Proof. See Corollary 24.5.1 of Rockafellar [33]. □

C. Convergence of subgradients

Pointwise convergence of convex functions entails a kind of convergence of their subgradients.

Theorem 59. *Let f be a convex function on $\mathbb{R}^n$, and let C be an open convex set on which f is finite. Let $f_1, f_2, \ldots$ be a sequence of convex functions finite on C and converging pointwise to f on C. Let $x \in C$, and let $x_1, x_2, \ldots$ be a sequence of points in C converging to x. Then for any $\varepsilon > 0$, there exists N such that*

$$\partial f_i(x_i) \subset \partial f(x) + B_\varepsilon(0)$$

for all $i \geq N$.

Proof. See Theorem 24.5 of Rockafellar [33]. □

There is in fact a strengthening of this result that will be crucial for us. In order to state the result, we need to define the notion of epigraphical convergence.

Definition 60. *Let f_i be a sequence of functions on $\mathbb{R}^n$. Then the* lower epi-limit $\operatorname{e}\liminf_i f_i$ *is the function having as its epigraph* $\limsup_i (\operatorname{epi} f_i)$ *(the outer limit of the sets* $\operatorname{epi} f_i$*).*

The upper epi-limit $\operatorname{e}\limsup_i f_i$ *is the function having as its epigraph* $\liminf_i (\operatorname{epi} f_i)$ *(the inner limit of the sets* $\operatorname{epi} f_i$*).*

When $\operatorname{e}\liminf_i f_i = \operatorname{e}\limsup_i f_i$*, we call this function the* epi-limit function $\operatorname{e}\lim_i f_i$*, and we say that the f_i* epi-converge *to* $\operatorname{e}\lim_i f_i$*. Therefore we have that the f_i epi-converge to f if and only if* $\operatorname{epi} f_i \to \operatorname{epi} f$ *(in the sense of set convergence).*

There is a useful theorem that relates pointwise convergence and epi-convergence of convex functions.

Theorem 61. *Let f_i be a sequence of convex functions on $\mathbb{R}^n$, and let f be a convex lower semi-continuous function on $\mathbb{R}^n$ such that* $\operatorname{dom} f$ *has non-empty interior. Then* $f = \operatorname{e}\lim_i f_i$ *if and only if the f_i converge uniformly on every compact set C that does not contain a boundary point of* $\operatorname{dom} f$*.*

Proof. See Theorem 7.17 of Rockafellar and Wets [34]. □

Now we have the following sort of 'locally uniform' convergence of the subgradients.

Theorem 62. *Let f, f_i be lower semi-continuous convex functions such that the f_i epi-converge to f. Let $x \in \operatorname{int}\operatorname{dom} f$. Then for all $\varepsilon > 0$, there exist $\delta > 0$ and N such that*

$$\partial f_i(z) \subset \partial f(x) + B_\varepsilon(0)$$

for all $z \in B_\delta(x)$ and all $i \geq N$.

Proof. See Theorem 8.3 of Bagh and Wets [2]. □

D. The convex conjugate

Definition 63. *Let f be a function $\mathbb{R}^n \to [-\infty, +\infty]$. Then the* convex conjugate *(or,* Legendre-Fenchel transform*) $f^* : \mathbb{R}^n \to [-\infty, +\infty]$ is defined by*

$$f^*(x^*) = \sup_x \{\langle x, x^* \rangle - f(x)\} = -\inf_x \{f(x) - \langle x, x^* \rangle\}.$$

Theorem 64. *Let f be a convex function. Then f^* is a closed convex function, proper if and only if f is proper. Furthermore, if f is closed, then $f^{**} = f$.*

Proof. See Theorem 12.2 of Rockafellar [33]. □

It is an important fact that the subgradients of f and f^* are, in a sense, inverse mappings.

Theorem 65. *If f is a closed proper convex function, then $x \in \partial f^*(x^*)$ if and only if $x^* \in \partial f(x)$.*

Proof. See Corollary 23.5.1 of Rockafellar [33]. □

Under certain mild conditions, the epi-convergence of a sequence of convex functions is equivalent to the epi-convergence of the corresponding sequence of conjugate functions.

Theorem 66. *Let f_i and f be proper lower semi-continuous convex functions on $\mathbb{R}^n$. Then the f_i epi-converge to f if and only if the f_i^* epi-converge to f^*.*

Proof. See Theorem 11.34 of Rockafellar and Wets [34]. □

References

[1] S. Angenent, S. Haker, A. Tannenbaum, *Minimizing flows for the Monge–Kantorovich problem,* SIAM J. Math. Anal. 35 (2003), 61–97.

[2] A. Bagh, R. J.-B. Wets, *Convergence of set-valued mappings: equi-outer semicontinuity,* Set-Valued Anal. 4 (1996), 333–360.

[3] A. Baldes, O. Wohlrab, Computer graphics of solutions of the generalized Monge–Ampère equation, in: Geometric analysis and computer graphics (Berkeley, CA, 1988), Math. Sci. Res. Inst. Publ., Vol. 17, Springer, 1991, pp. 19–30,.

[4] G. Barles, P. E. Souganidis, *Convergence of approximation schemes for fully nonlinear second order equations,* Asymptot. Anal., 4 (1991), 271–283.

[5] J.-D. Benamou, Y. Brenier, *A computational fluid mechanics solution to the Monge-Kantorovich mass transfer problem,* Numer. Math. 84 (2000), 375–393.

[6] J.-D. Benamou, Jean-David; B.D. Froese, A.M. Oberman, *Numerical solution of the optimal transportation problem using the Monge–Ampère equation,* J. Comput. Phys. 260 (2014), 107–126.

[7] J.-D. Benamou, B.D. Froese, A.M. Oberman, *A viscosity solution approach to the Monge–Ampère formulation of the optimal transportation problem,* preprint, arxiv:1208.4873.

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

[9] L. Caffarelli, *The regularity of mappings with a convex potential,* J. Amer. Math. Soc. 5 (1992), 99–104.

[10] L. Caffarelli, *Boundary Regularity of Maps with Convex Potentials–II,* Ann. of Math. 144 (1996), 453–496.

[11] O. Chodosh, V. Jain, M. Lindsey, L. Panchev, Y.A. Rubinstein, *Visualizing optimal transportation maps,* Abstracts of papers presented to the Amer. Math. Soc., Vol. 34, No. 1086-28-496, 2013; V. Jain, M. Lindsey, L. Panchev, SURIM project report, Stanford University, 2012.

[12] O. Chodosh, V. Jain, M. Lindsey, L. Panchev, Y.A. Rubinstein, *On discontinuity of planar optimal transport maps,* J. Topology & Analysis (2015).

[13] C. K. Chui and P. W. Smith, *A note on Landau's problem for bounded intervals*, Am. Math. Mon. 82 (1975), 927–929.

[14] G. De Philippis and A. Figalli, *The Monge–Ampére equation and its link to optimal transport,* preprint, available at http://arxiv.org/abs/1310.6167 (2013)

[15] X.-B. Feng, R. Glowinski, M. Neilan, *Recent developments in numerical methods for fully nonlinear second order partial differential equations,* SIAM Rev. 55 (2013), 205–267.

[16] B.D. Froese, A.M. Oberman, *Convergent filtered schemes for the Monge–Ampère partial differential equation,* SIAM J. Numer. Anal. 51 (2013), 423–444.

[17] K. Guittet, *On the time-continuous mass transport problem and its approximation by augmented Lagrangian techniques,* SIAM J. Numer. Anal. 41 (2003), 382–399.

[18] E. Haber, T. Rehman, A. Tannenbaum, *An efficient numerical method for the solution of the L^2 optimal mass transfer problem,* SIAM J. Sci. Comput. 32 (2010), 197–211.

[19] J.-B. Hiriart-Urruty, C. Lemaréchal, *Convex analysis and minimization algorithms I, II,* Springer, 1993.

[20] R. Jonker, A. Volgenant, *A shortest augmenting path algorithm for dense and sparse linear assignment problems*, Computing 38 (1987), 325–340.

[21] J. Kitagawa, *An iterative scheme for solving the optimal transportation problem,* Calc. Var. Partial Differential Equations 51 (2014), 243–263.

[22] G. Loeper, F. Rapetti, *Numerical solution of the Monge–Ampère equation by a Newtons algorithm,* C. R. Acad. Sci. Paris, Ser. I 340 (2005) 319–324.

[23] J. Löfberg, *YALMIP: A toolbox for modeling and optimization in MATLAB*, Proc. CACSD Conf. (2004).

[24] Q. Mérigot, É. Oudet, *Discrete optimal transport: complexity, geometry and applications,* preprint, available at: http://quentin.mrgt.fr/research/

[25] H. Minkowski, *Volumen und Oberflächen,* Math. Ann. 57 (1903), 447–495.

[26] MOSEK ApS, *The MOSEK optimization toolbox for MATLAB manual. Version 7.1 (Revision 28)* (2015), http://docs.mosek.com/7.1/toolbox/index.html.

[27] M. Neilan, *Finite element methods for fully nonlinear second order PDEs based on a discrete Hessian with applications to the Monge–Ampère equation,* J. Comput. Appl. Math. 263 (2014), 351–369.

[28] V.I. Oliker, L.D. Prussner, *On the numerical solution of the equation $(\partial^2 z/\partial x^2)(\partial^2 z/\partial y^2) - ((\partial^2 z/\partial x \partial y))^2 = f$ and its discretizations. I,* Numer. Math. 54 (1988), 271–293.

[29] N. Papadakis, G. Peyré, E. Oudet, *Optimal transport with proximal splitting,* SIAM J. Imaging Sci. 7 (2014), 212–238.

[30] P.-O. Persson, G. Strang, *A simple mesh generator in MATLAB*, SIAM Rev. 46 (2004), 329–345.

[31] A.V. Pogorelov, *The Minkowski multidimensional problem,*, Wiley, 1978.

[32] J. Rauch, B.A. Taylor, *The Dirichlet problem for the multidimensional Monge–Ampère equation,* Rocky Mountain J. Math. 7 (1977), 345–364.

[33] R. T. Rockafellar, *Convex analysis, Princeton Mathematical Series,* no. 28, Princeton University Press, Princeton, N.J. (1970)

[34] R. T. Rockafellar, R. J.-B. Wets, *Variational Analysis,* Springer, 2009.

[35] M.M. Sulman, J.F. Williams, R.D. Russell, *An efficient approach for the numerical solution of the Monge–Ampère equation,* Appl. Numer. Math. 61 (2011), 298–307.

[36] C. Villani, *Topics in optimal transportation,* American Mathematical Society, 2003.

[37] C. Villani, *Optimal transport: old and new*, Springer, 2009.

E-mail address: `lindsey3@stanford.edu`