SIAM J. MATH. ANAL.
Vol. 49, No. 4, pp. 3073–3124

OPTIMAL TRANSPORT VIA A MONGE–AMPÈRE OPTIMIZATION PROBLEM*

MICHAEL LINDSEY† AND YANIR A. RUBINSTEIN‡

Abstract. We rephrase Monge's optimal transportation (OT) problem with quadratic cost—via a Monge–Ampère equation—as an infinite-dimensional optimization problem, which is in fact a convex problem when the target is a log-concave measure with convex support. We define a natural finite-dimensional discretization to the problem and associate a piecewise affine convex function to the solution of this discrete problem. The discrete problems always admit a solution, which can be obtained by standard convex optimization algorithms whenever the target is a log-concave measure with convex support. We show that under suitable regularity conditions the convex functions retrieved from the discrete problems converge to the convex solution of the original OT problem furnished by Brenier's theorem. We demonstrate numerical solutions of these discrete problems and then explain (at the expense of provable convergence) how to modify our method to make it more efficient and more accurate, as well as how to remove the restriction on the target measure via a fixed point method that only involves solving OT problems with constant target densities.

Key words. optimal transport, optimization, Monge–Ampère

AMS subject classifications. 35J60, 65N12

DOI. 10.1137/16M1071560

1. Introduction.

In this article we develop schemes for the numerical solution of the the Monge–Ampère equation governing optimal transport on $\mathbb{R}^n$. Our main theorem provides natural and computationally feasible approximations of optimal transportation (OT) maps as well as of their Brenier convex potential, in the case of quadratic cost.

To achieve this we rephrase Monge's problem as an infinite-dimensional optimization problem, which is, in fact, a convex problem when the target is a measure with convex support whose density g is such that $g^{-1/n}$ is convex. Note that this class of measures *includes* all log-concave measures with convex support. We define a natural finite-dimensional discretization to the problem and associate a piecewise affine convex function to the solution of this discrete problem. The discrete problems always admit a solution, which can be obtained by standard convex optimization algorithms whenever the target is a measure with convex support whose density g is such that $g^{-1/n}$ is convex. We show that under suitable regularity conditions the convex functions retrieved from the discrete problems converge to the convex solution of the original OT problem furnished by Brenier's theorem. While this result yields new insights into optimal transport maps, it also has applications to the numerical solution of OT problems. We illustrate the method treated by our convergence theorem with a number of numerical examples, before turning to some more heuristically

*Received by the editors April 20, 2016; accepted for publication (in revised form) May 22, 2017; published electronically August 17, 2017.
http://www.siam.org/journals/sima/49-4/M107156.html
Funding: This work was initially supported by Stanford University's SURIM and OVPUE during Summer 2012, and subsequently by NSF grants DMS-1206284 and 1515703 and a Sloan Research Fellowship. Part of this work took place at MSRI (supported by NSF grant DMS-1440140) during the Spring 2016 semester. The work of the first named author was also partially supported by the NSF Graduate Research Fellowship Program under grant DGE-1106400.
†Department of Mathematics, University of California, Berkeley, CA 94720-3840 (lindsey@berkeley.edu).
‡Department of Mathematics, University of Maryland, College Park, MD 20742 (yanir@umd.edu).

© 2017 Michael J. Lindsey and Yanir A. Rubinstein

justified modifications of the basic method that enhance its efficiency and accuracy and remove the aforementioned restriction on the target density. We explore these modifications numerically.

1.1. Yet another formulation 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 positive measures μ and ν on Ω and Λ, respectively, by

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

where dx denotes the Lebesgue measure on $\mathbb{R}^n$, and assume that

$$\int_\Omega f\,dx = \int_\Lambda g\,dx. \tag{1}$$

Then by results by Brenier and Caffarelli [8, 9, 40], there exists a unique solution of the corresponding Monge problem for 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 in $C^{1,\alpha}(\Omega)$. Further, the solution is given by $T=\nabla\varphi$ for φ convex and $C^{2,\alpha}(\Omega)$. In addition, φ is the unique (up to an additive constant) Brenier (and, hence, also Alexandrov, or viscosity) solution of the second boundary value problem for the Monge–Ampère equation

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

where $\partial\varphi$ denotes the subdifferential map associated to the convex function φ.

The following result rephrases Monge's problem as an infinite-dimensional optimization problem. This problem can be considered "convex" whenever, in particular, the target measure has log-concave (e.g., uniform) density with convex support. We refer the reader to section 2 for more details and intuition on, as well as a proof of, the result. Section 2 can be read independently of all other sections.

PROPOSITION 1.1. *With notation and hypotheses as in the above discussion, φ is the unique (up to addition by a constant) solution of the following optimization problems:*

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

where

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

and

$$\mathcal{G}(\psi,x) := \max\left\{0, -\Big(\det\left(\nabla^2\psi(x)\right)\Big)^{1/n} + \Big(f(x)/g\left(\nabla\psi(x)\right)\Big)^{1/n}\right\}.$$

Remark 1.2. By "ψ convex" we mean that ψ extends to convex function on all of $\mathbb{R}^n$.

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

1.2. A discrete Monge–Ampère optimization problem. Let

$$x_1, \ldots, x_N \in \overline{\Omega}$$

be points, and let

$$S_1, \ldots, S_M \subset \overline{\Omega}$$

be n-dimensional simplices whose set of vertices equals $\{x_j\}_{j=1}^N$ that together form an *almost-triangulation* of $\overline{\Omega}$. By this we mean that the intersection of any two of the S_i is either empty or a common face (of any dimension) and that $\overline{\Omega} \setminus \bigcup_{i=1}^M S_i$ has "small" volume. Note that the triangulation can be made perfect if Ω is a polytope. We denote the vertices of the simplex S_i by

$$x_{i_0}, \ldots, x_{i_n}, \quad i_0, \ldots, i_n \in \{1, \ldots, N\},$$

so that $S_i = \mathrm{co}\{x_{i_0}, \ldots, x_{i_n}\}$, where

$$\mathrm{co}\, A$$

denotes the convex hull of a set A. We denote by

$$\eta_1, \ldots, \eta_N \in \overline{\Lambda}$$

points in the closure of the target domain Λ, and we think of η_j as the "image" of x_j, so that intuitively S_i gets mapped to $\mathrm{co}\{\eta_{i_0}, \ldots, \eta_{i_n}\}$. We associate to the discrete map $x_j \mapsto \eta_j$ and the almost-triangulation $\{S_i\}_{i=1}^M$ a sort of discrete Jacobian defined separately for each simplex. To define this, let A_i, and B_i be n-by-n matrices defined by

$$\begin{aligned} A_i &:= \left[\ (x_{i_1} - x_{i_0}) \quad (x_{i_2} - x_{i_0}) \quad \cdots \quad (x_{i_n} - x_{i_0})\ \right]^T, \\ B_i &:= \left[\ (\eta_{i_1} - \eta_{i_0}) \quad (\eta_{i_2} - \eta_{i_0}) \quad \cdots \quad (\eta_{i_n} - \eta_{i_0})\ \right]^T \end{aligned} \tag{3}$$

(here, x_j and η_j are represented by column vectors). The *discrete Jacobian associated to the simplex* S_i is then the n-by-n matrix

$$H_i := H\Big(S_i, \eta_{i_0}, \ldots, \eta_{i_n}\Big) := \frac{1}{2}\,(A_i)^{-1}\, B_i + \frac{1}{2}\Big((A_i)^{-1}\, B_i\Big)^T. \tag{4}$$

Roughly speaking, H_i is a finite difference approximation to the Jacobian of the transport map on simplex i, whose determinant measures the change in volume under the map. This would indeed be the exact Jacobian of the piecewise-linear map that interpolates $x_j \mapsto \eta_j$, but, as we shall see, we must demand symmetry in order to allow H_i to lie in the positive semidefinite cone, so it is really the symmetrization of this exact Jacobian.

Finally, denote the volume of the ith simplex by

$$V_i := \mathrm{vol}(S_i). \tag{5}$$

We now introduce a discrete analogue of the first optimization problem associated to the Monge–Ampère equation (2) introduced in Proposition 1.1.

Definition 1.3. *The discrete Monge–Ampère optimization problem (DMAOP) associated to the data* $(\Omega, \Lambda, f, g, \{x_j\}_{j=1}^N, \{S_i\}_{i=1}^M)$ *is as follows:*

$$\underset{\{\psi_i \in \mathbb{R}, \eta_i \in \mathbb{R}^n\}_{i=1}^N}{\text{minimize}} \sum_{i=1}^{M} V_i \cdot \max\left\{0, -(\det H_i)^{1/n} + \left(f\left(\tfrac{\sum_{j=0}^n x_{i_j}}{n+1}\right)/g\left(\tfrac{\sum_{j=0}^n \eta_{i_j}}{n+1}\right)\right)^{1/n}\right\}$$

$$\text{(6)} \quad \textit{subject to} \quad \psi_j \geq \psi_i + \langle \eta_i, x_j - x_i \rangle, \ i, j = 1, \ldots, N,$$

$$\text{(7)} \quad \eta_i \in \overline{\Lambda}, \ i = 1, \ldots, N,$$

$$\text{(8)} \quad H_i \geq 0, \ i = 1, \ldots, M.$$

To ease the notation in the following, we define

$$F_i\left(\{\psi_j^{(k)}, \eta_j^{(k)}\}_{j=1}^{N(k)}\right) := \max\left\{0, -(\det H_i)^{1/n} + \left(f\left(\tfrac{\sum_{j=0}^n x_{i_j}}{n+1}\right)/g\left(\tfrac{\sum_{j=0}^n \eta_{i_j}}{n+1}\right)\right)^{1/n}\right\} \tag{9}$$

and

$$F\left(\{\psi_j^{(k)}, \eta_j^{(k)}\}_{j=1}^{N(k)}\right) := \sum_{i=1}^{M(k)} V_i \cdot F_i\left(\{\psi_j^{(k)}, \eta_j^{(k)}\}_{j=1}^{N(k)}\right), \tag{10}$$

so F_i is a per-simplex penalty, and F is the objective function of the DMAOP.

The variables of the DMAOP are $\psi_1, \ldots, \psi_N$ and $\eta_1, \ldots, \eta_N$ (while $x_1, \ldots, x_N, \Omega, \Lambda, f, g$ are given). These variables are the discrete analogues of the values of the convex potential and its gradient, respectively, at the points $x_1, \ldots, x_N$, while H_i is the discrete analogue of the Jacobian, on the simplex S_i, of the map $x_i \mapsto \eta_i$. One can think of $\det H_i$ as a measure of the volume distortion of simplex S_i under the map.

We will see later (see Proposition 3.1(i)) that the DMAOP is feasible (i.e., there exists a point in the variable space $(\psi_1, \ldots, \psi_N, \eta_1, \ldots, \eta_N) \in (\mathbb{R} \times \mathbb{R}^n)^N$ satisfying the constraints (6)–(8)) for a fine enough triangulation. We will argue now that if the problem is feasible, then it admits an optimizer. First note that the satisfaction of the constraints and the value of the objective function are unaffected by shifting all of the ψ_i by the same constant. Thus it is equivalent to considering the DMAOP with the additional constraint that $\psi_1 = 0$. Note that the η_i are bounded over the optimization domain (since Λ is bounded), and consequently, with the additional constraint included, the ψ_i are bounded over the optimization domain by (6). Thus the optimization domain is compact, and the problem must admit a minimizer. In the case that Λ is convex and $g^{-1/n}$ is convex, the DMAOP is in fact a convex optimization problem (see Remark 2.2), so (see, for instance, [7, Chapter 11]), we expect that the DMAOP can be solved efficiently.

1.3. Convex functions associated to the solution of the DMAOP. Next, we construct a piecewise linear convex function ϕ associated with a solution $\{\psi_j, \eta_j\}_{j=1}^N$ of the DMAOP. Define

$$a_j(x) := \psi_j + \langle \eta_j, x - x_j \rangle, \qquad j = 1, \ldots, N, \tag{11}$$

so a_j is the (unique) affine function with $a_j(x_j) = \psi_j$ and $\nabla a_j(x_j) = \eta_j$. Define the *optimization potential* by

$$\phi(x) := b + \max_{j=1,\ldots N} a_j(x), \tag{12}$$

where $b \in \mathbb{R}$ is chosen such that $\phi(0) = 0$ (and we have assumed, without loss of generality, that $0 \in \Omega$). Notice that we have defined ϕ on all of $\mathbb{R}^n$. As the supremum of affine functions, ϕ is convex. The point is that ϕ still encodes the solution of the DMAOP. Indeed, by the constraints of the DMAOP (specifically (6)),

$$\phi(x_j) = \psi_j + b \tag{13}$$

and

$$\eta_j \in \partial\phi(x_j) \tag{14}$$

(as a_j is an affine function with slope η_j lying below ϕ but touching it at x_j).

1.4. Relationship of the DMAOP with discrete optimal transport. We briefly describe a 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.

Let $\{\psi_j, \eta_j\}_{j=1}^N \in (\mathbb{R} \times \mathbb{R}^n)^N$ be a solution of the DMAOP (with notation as above). The construction of section 1.3 yields, by (13)–(14), 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 [33, Theorem 24.8], the set $\bigcup_{j=1}^N (x_j, \eta_j) \subset \mathbb{R}^{2n}$ is cyclically monotone (see [33, section 24] for definitions) since it is a subset of the graph of the subdifferential of the convex function ϕ. Thus, $x_j \mapsto \eta_j$ solves the optimal transport problem from $\mu_D = \sum_{j=1}^N \delta_{x_j}$ to $\nu_D = \sum_{j=1}^N \delta_{\eta_j}$ [39], where δ_p denotes the Dirac delta measure concentrated at p. We state this result as a proposition.

Proposition 1.4. *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 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. The DMAOP chooses target points η_j in a way that attempts to achieve correct volume distortion.

1.5. Convergence of the discrete solutions. Now, we take a sequence of almost-triangulations $\{x_i\}_{i=1}^{N(k)}, \{S_i^{(k)}\}_{i=1}^{M(k)}$ indexed by k (so now both N and M are functions of k, although we will usually omit that dependence from our notation) satisfying the following assumptions. Denote by $\operatorname{diam} A$ the diameter of a set A in Euclidean space.

Definition 1.5. *We say that the sequence of almost-triangulations $\{\{S_i^{(k)}\}_{i=1}^{M(k)}\}_{k\in\mathbb{N}}$ is admissible if*

$$\lim_{k\to\infty} \max_{i\in\{1,\ldots,M(k)\}} \operatorname{diam} S_i^{(k)} = 0,$$

and there are open sets $\Omega_\varepsilon \subset \Omega$ indexed by $\varepsilon > 0$, with

$$\Omega_\varepsilon \subset \Omega_{\varepsilon'} \quad \textit{for } \varepsilon' \le \varepsilon$$

and

$$\bigcup_{\varepsilon>0} \Omega_\varepsilon = \Omega,$$

and such that for any $\varepsilon > 0$ sufficiently small, we have that an ε-neighborhood of Ω_ε is contained within the kth almost-triangulation for all k sufficiently large, i.e.,

$$\Omega_\varepsilon + B_\epsilon(0) \subset \bigcup_{i=1}^{M(k)} S_i^{(k)} \quad \forall k \gg 1.$$

Given an admissible sequence of almost-triangulations, we construct the optimization potentials

$$\phi^{(k)} \in C^{0,1}(\mathbb{R}^n), \quad k \in \mathbb{N}, \tag{15}$$

associated with the solution $\{\psi_j^{(k)}, \eta_j^{(k)}\}_{j=1}^{N(k)}$ of the kth DMAOP, by the prescription of the previous subsection (specifically, by (12)). Our main theorem concerns the convergence of the optimization potentials to the Brenier potential, i.e., to the solution of the PDE (2).

THEOREM 1.6. *Let $f \in C^0(\overline{\Omega})$, $g \in C^0(\overline{\Lambda})$, and suppose that Λ is convex. Let $\{\{S_i^{(k)}\}_{i=1}^{M(k)}\}_{k\in\mathbb{N}}$ be an admissible sequence of almost-triangulations (recall Definition 1.5). Suppose that the optimal cost of the kth DMAOP tends to zero as $k \to \infty$. Then, as $k \to \infty$, the optimization potentials $\phi^{(k)}$ (15) converge uniformly on $\overline{\Omega}$ to the unique Brenier solution φ of the Monge–Ampère equation (2) with $\varphi(0) = 0$.*

Remark 1.7. For the proof of Theorem 1.6, we do not actually need $\phi^{(k)}$ to be the optimization potential retrieved from the optimal solution of the kth DMAOP. We only need that the $\phi^{(k)}$ are defined via (11) and (12) by some $\{\psi_j^{(k)}, \eta_j^{(k)}\}_{j=1}^{N}$ which satisfy the constraints of Definition 1.3 and for which the cost of Definition 1.3 tends to zero as $k \to \infty$. See section 1.6 for implications.

The assumption on the optimal cost holds in many interesting cases by assuming a mild regularity condition on the almost-triangulations.

DEFINITION 1.8. *We say that a sequence $\{\{S_i\}_{i=1}^{M(k)}\}_{k\in\mathbb{N}}$ of almost-triangulations of Ω is regular if there exists $R > 0$ such that*

$$\det\left(u_{i,1}^{(k)} \ u_{i,2}^{(k)} \ \cdots \ u_{i,n}^{(k)}\right) \geq R \quad \forall i \in \{1, \dots, M(k)\}, \ \forall k \in \mathbb{N},$$

where

$$u_{i,j}^{(k)} := \|x_{i_j} - x_{i_0}\|^{-1}\left(x_{i_j} - x_{i_0}\right).$$

For example, in dimension $n = 2$, a sequence of almost-triangulations is regular if the angles of the triangles are bounded below uniformly in k. We remark that similar restrictions appear in the finite element method literature.

Based on Theorem 1.6 we prove the following general convergence result that does not make any assumptions on the optimal cost of the discretized problems. The convergence we obtain on the level of subdifferentials can be viewed as optimal since $\phi^{(k)}$ are Lipschitz but no better, i.e., $\phi^{(k)} \notin C^1$.

THEOREM 1.9. *Let $f \in C^{0,\alpha}(\overline{\Omega})$, $g \in C^{0,\alpha}(\overline{\Lambda})$, and suppose that Λ is convex. Let $\{\{S_i^{(k)}\}_{i=1}^{M(k)}\}_{k\in\mathbb{N}}$ be an admissible and regular sequence of almost-triangulations (recall Definitions 1.5 and 1.8). Let φ be the the unique Brenier solution of the Monge–Ampère equation (2) with $\varphi(0) = 0$, and let $\phi^{(k)}$ be the optimization potentials (15) obtained from the DMAOP. Then*

$$\phi^{(k)} \to \varphi \text{ uniformly on } \overline{\Omega},$$

and $\partial\phi^{(k)} \to \nabla\varphi$ pointwise on $\overline{\Omega}$. In particular, $\nabla\phi^{(k)}$ converges pointwise almost everywhere to the optimal transport map pushing forward $\mu = f\,dx$ to $\nu = g\,dx$.

Proof. Under the assumption that the sequence of almost-triangulations is admissible and regular and that $\varphi \in C^2(\Omega)$, Corollary 5.5 states that c_k, the optimal cost of the kth DMAOP, tends to zero as $k \to \infty$. Thus, Theorem 1.6 implies $\phi^{(k)}$ converges uniformly to the Brenier solution φ. We will see from the proof of Theorem 1.6 that we can in fact assume that the $\phi^{(k)}$ are uniformly convergent on a closed ball D containing $\overline{\Omega}$ in its interior. The convergence of subgradients then follows from Theorem 4.7 since if a sequence of lower semicontinuous finite convex functions converges uniformly on bounded sets to some convex function, then the sequence epi-converges to this function [34, Theorem 7.17]. ∎

Remark 1.10. The conditions of admissibility and regularity on the sequence of triangulations are necessary for technical reasons, but they are fulfilled easily in practice.

Remark 1.11. We see from the statement of Theorem 1.6 that, in practice, 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 kth DMAOP becomes small as $k \to \infty$. However, it remains an open problem to prove the error estimates that would make this insight rigorous.

Remark 1.12. Although, via Corollary 5.5, we do not require regularity up to the boundary to obtain convergence in the DMAOP case, it can be seen that the proof of Corollary 5.5 does not yield a rate for the convergence $c_k \to 0$. To prove such a rate, we need $f \in C^{0,\alpha}(\overline{\Omega})$, $g \in C^{0,\alpha}(\overline{\Lambda})$, and $\varphi \in C^{2,\alpha}(\overline{\Omega})$, since in this case it follows from the proof of Corollary 3.3 that $c_k = O(h^\alpha)$, where $h = h(k)$ is the maximal simplex diameter in the kth triangulation. Note that a bound on c_k does *not* imply an error bound for the optimization potentials, though it suggests a candidate.

1.6. Contributions of this work. We outline here the theoretical and numerical contributions of this work. First, we have introduced an infinite-dimensional optimization problem equivalent to the Monge–Ampère equation for optimal transport, and we have identified natural discretizations of this problem for which convergence can be established. It is perhaps most appropriate to think of the convergence result (Theorem 1.6) as a convergence result for convex functions. We interpret the theorem as stating that a sort of one-sided convergence (over an increasingly fine set of discretization points) of the discrete Hessian determinants of a sequence of convex functions to the "right-hand side" of a Monge–Ampère equation implies convergence of the convex functions themselves to the solution of the Monge–Ampère equation, as long as the gradients are confined within the target region. We can interpret this as a coercivity-type result allowing us to pass from vanishing violation (even one-sided) of the Monge–Ampère equation to vanishing deviation from the solution.

Furthermore, in establishing Theorem 1.6, we develop a perspective on the numerical analysis of the peculiar boundary condition of (2) that we do not think has been explored in the literature. We also comment that our method treats the boundary condition of (2) in a new way that is simple both in theory and in numerical implementation, relying only on the enforcement of a convex constraint.

Though the practical applicability of our method is a priori limited to target densities g such that $g^{-1/n}$ is convex, we have introduced (in section 8) a fixed point iteration that at least heuristically reduces the solution of an OT problem to the solution of a sequence of OT problems with uniform target density, extending the applicability of our method to a wider class of problems. This iteration could be of use in other methods.

Numerically, we illustrate the features of the DMAOP before introducing (in section 7) a more efficient and accurate method, the revised DMAOP, or RDMAOP. In particular, the RDMAOP achieves second-order accuracy in some settings.

Lastly, we remark that the ideas of this work (specifically, regarding the optimization approach) have applicability more generally to the numerical analysis of nonlinear elliptic PDEs (including other Monge–Ampère equations), and we are actively pursuing applications of these ideas.

1.7. Review of existing methods. Our work relies on a (convex) optimization approach. In this subsection we review other, different, approaches to discretizing the OT problem and, sometimes, more generally the Monge–Ampère equation or even more general fully nonlinear second-order elliptic PDEs. Recently, the literature on numerical methods for Monge–Ampère equations in general, and for optimal transport maps in particular, has grown considerably. Therefore, we do not attempt a comprehensive review of existing methods in the literature, but rather concentrate on briefly mentioning those approaches for which both a numerical algorithm has been implemented *and* a convergence result has been proven. For a thorough survey of existing numerical methods, we refer the interested reader to the article of Feng, Glowinski, and Neilan [14] and references therein.

Oliker and Prussner [28] and Baldes and Wohlrab [3] initiated the study of discretizations of the 2-dimensional Monge–Ampère equation and obtained a convergence theorem for the Dirichlet problem for the equation $u_{xx}u_{yy} - u_{xy}^2 = f$ on a bounded domain in $\mathbb{R}^2$. This used, among other things, classical constructions of Minkowski [24] and Pogorelov [31].

Benamou and Brenier [4] introduced, on the other hand, a discretization scheme for the dynamic formulation of the optimal transport problem that does not involve the Monge–Ampère equation. This involves solving the system of equations for $\rho : [0,T] \times \mathbb{R}^n \to \mathbb{R}_+$, $v : [0,T] \times \mathbb{R}^n \to \mathbb{R}^n$,

$$\begin{aligned} \partial_t \rho + \operatorname{div}(\rho v) &= 0, \\ \partial_t \phi + |\nabla_x \phi|^2/2 &= 0, \end{aligned} \tag{16}$$

with the constraint $v = \nabla_x \phi$ and the boundary conditions $\rho(0, \cdot\,) = f$, $\rho(T, \cdot\,) = g$. The authors use discretization in space-time that falls within the framework of problems in numerical fluid mechanics. See also the work of Angenent, Haker, and Tannenbaum [1] and Haber, Rehman, and Tannenbaum [18]. More recently, Guittet proved that the Benamou–Brenier scheme converges when the target is convex and the densities are smooth [17].

Recently, Benamou, Froese, and Oberman proposed a convergence proof via a direct discretization of the Monge–Ampère equation [5, 6]. Other recent work includes, e.g., Loeper and Rapetti [21], Sulman, Williams, and Russell [38], Kitagawa [19], and Papadakis, Peyré, and Oudet [29].

Another approach one could pursue is to approximate the measures by empirical measures (sums of Dirac measures). In the simplest case, when the number of Dirac measures is the same for the source and the target, the solution is given by solving the assignment problem that has efficient numerical implementations. We refer the reader to Mérigot and Oudet [23] and [29, p. 213] for relevant references (cf. [10] for an implementation in some simple cases). It is interesting to note that the method presented in this article a fortiori solves an assignment problem, but for target Dirac measures whose location is not a priori known (as explained in section 1.4 above). We also point out semidiscrete methods such as Lévy [20]; cf. Example 1.6 of Gangbo and McCann [16].

We also note recent methods based on the linear programming (Kantorovich) formulation of optimal transport that maintain efficiency by exploiting sparsity, in particular those of Schmitzer [37] and Oberman and Ruan [27].

Lastly, we mention the approach of Cuturi [13] to numerical OT via entropic regularization that aims to solve a regularized problem.

2. The Monge–Ampère optimization problem.

We recall some of the notation from section 1.1. Let Ω and Λ be bounded open sets in $\mathbb{R}^n$ with $0 \in \Omega$ and Λ convex. Let $f \in C^{0,\alpha}(\overline{\Omega})$ and $g \in C^{0,\alpha}(\overline{\Lambda})$ be positive functions bounded away from zero and infinity satisfying (1). Let $\varphi \in C^{2,\alpha}(\Omega)$ be the unique convex solution of (2) with $\varphi(0) = 0$.

Proposition 1.1 is a special case of the following result.

LEMMA 2.1. *With notation and hypotheses as in the above paragraph, φ is the unique solution of the following optimization problem:*

$$\underset{\psi \in \mathrm{Cvx}(\Omega) \cap C^2(\Omega)}{\text{minimize}} \mathcal{F}(\psi) := \int_\Omega h \circ \mathcal{G}(\psi, x) \cdot \rho(x) dx$$

$$\textit{subject to} \quad \nabla\psi(\Omega) \subseteq \Lambda,$$

where

$$\mathcal{G}(\psi, x) := \max\left\{0, -\big(\det\big(\nabla^2\psi(x)\big)\big)^{1/n} + \Big(f(x)/g\left(\nabla\psi(x)\right)\Big)^{1/n}\right\},$$

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.

Before giving the proof we make several remarks.

Remark 2.2. Notice that if $g^{-1/n}$ is convex, 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;
- $\det^{1/n}$ 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 (i.e., the DMAOP) outlined above is a convex problem in the usual sense.

We comment that $g^{-1/n}$ is convex whenever g is log-concave. Indeed, this can be seen by writing $g^{-1/n} = \exp\left(-\frac{1}{n}\log g\right)$ and recalling that exp preserves convexity. Thus the DMAOP is a convex problem for a strictly larger class of target measures.

Remark 2.3. Nevertheless, for the proof of the main theorems we do not require $g^{-1/n}$ to be convex (nor that g be log-concave), though these assumptions ensure that the DMAOP is convex and, thus, feasibly solvable.

Remark 2.4 (intuitive explanation of Lemma 2.1). 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.

Remark 2.5. One could alternatively consider a functional $\widetilde{\mathcal{G}}$ defined by

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

in place of $\mathcal{G}$ in the above. Since $\log\det$ is concave on the set of positive definite matrices, the resulting problem (along with its discretization) is still convex as long as g is log-concave. All of the analysis of this paper, except for section 5, carries over via natural modifications to this alternative problem. More precisely, difficulties arise in section 5 because, unlike the nth root, the logarithm is unbounded for positive numbers near zero. Nevertheless, by making the stronger assumption that the Brenier potential φ is in $C^{2,\alpha}(\overline{\Omega})$ these difficulties are overcome and one obtains a convergence proof (we refer the interested reader to an earlier posted version of this article for details).

Proof. Note that $F(\varphi) = 0$, since φ solves the Monge–Ampère equation, and 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 an 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 strongly convex. Thus $\nabla\psi$ is injective, and we obtain by a change of variables

$$\begin{aligned}\int_{\nabla\psi(\Omega)} g(y)dy &= \int_\Omega g\left(\nabla\psi(x)\right)\det\left(\nabla^2\psi(x)\right)dx \\ &> \int_\Omega f(x)dx.\end{aligned}$$

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 $\int_\Lambda g = \int_\Omega f$ by (1). ∎

3. 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 section 1.5, adding "(k)" in superscripts as necessary.

3.1. The objective function. First, we would like to understand the behavior of the objective function of the DMAOP. The following proposition gives a criterion guaranteeing the optimal cost (i.e., the minimum of the objective function) of the DMAOP converges to zero. The idea is to study the cost associated to the restriction of the solution φ of the Monge–Ampère equation to the kth almost-triangulations, i.e., to estimate the cost

$$d_k := F\Big(\big\{\varphi(x_j^{(k)}), \nabla\varphi(x_j^{(k)})\big\}_{j=1}^{N(k)}\Big) \tag{17}$$

associated to

$$\big\{\varphi(x_j^{(k)}), \nabla\varphi(x_j^{(k)})\big\}_{j=1}^{N(k)} \in (\mathbb{R}\times\mathbb{R}^n)^{N(k)}. \tag{18}$$

A small caveat, of course, is to show first that these data actually satisfy the constraints of the discrete Monge–Ampère optimization problem (DMAOP) and, subsequently, that the kth DMAOP is feasible. This is the content of part (i) of the next proposition.

PROPOSITION 3.1. *Let* $\big\{\{S_i^{(k)}\}_{i=1}^{M(k)}\big\}_{k\in\mathbb{N}}$ *be a sequence of admissible and regular almost-triangulations of* Ω *(recall Definitions* 1.5 *and* 1.8*). Let* φ *be the unique Brenier solution of the Monge–Ampère equation* (2) *with* $\varphi(0) = 0$*, and suppose that* $\varphi \in C^{2,\alpha}(\overline{\Omega})$*. Then*

(i) *the data* (18) *satisfies the constraints* (6)–(8) *for all k sufficiently large;*

(ii) $\lim_k d_k = 0$.

Remark 3.2. Due to the additional regularity assumption on φ, i.e., that $\varphi \in C^{2,\alpha}(\overline{\Omega})$, Proposition 3.1 is not sufficient to establish that Theorem 1.9 follows from Theorem 1.6 (cf. Remark 2.5). We will show later (in section 5) that we can relax the regularity assumption to $\varphi \in C^2(\Omega)$, in particular requiring no regularity up to the boundary. These details are more technical and have been postponed for expository clarity. However, we comment at this point that the regularity condition $\varphi \in C^2(\Omega)$ follows from the assumptions of Theorem 1.9. In fact (recalling our initial assumption that f and g are in $C^{0,\alpha}(\overline{\Omega})$ and $C^{0,\alpha}(\overline{\Lambda})$, respectively), if Ω and Λ are uniformly convex and of class C^2, then we actually do have the regularity $\varphi \in C^{2,\alpha}(\overline{\Omega})$ needed for Proposition 3.1 to apply. For a review of the relevant regularity theory, see [39, Chapter 4].

Denote by

$$\Big\{\psi_j^{(k)}, \eta_j^{(k)}\Big\}_{j=1}^{N} \in (\mathbb{R}\times\mathbb{R}^n)^{N(k)} \tag{19}$$

the solution to the kth DMAOP. The optimal (minimal) cost of the DMAOP associated with the kth almost-triangulation is then

$$c_k := F\Big(\big\{\psi_j^{(k)}, \eta_j^{(k)}\big\}_{i=1}^{N(k)}\Big). \tag{20}$$

Since $c_k \le d_k$, an immediate consequence of Proposition 3.1 is the following.

COROLLARY 3.3. *Under the assumptions of Proposition* 3.1, $\lim_k c_k = 0$.

This statement is analogous to consistency statements for more typical finite difference schemes.

Proof of Proposition 3.1. (i) We claim that the feasibility conditions (6)–(8) are satisfied for $\{\varphi(x_j^{(k)}), \nabla\varphi(x_j^{(k)})\}_{j=1}^{N(k)}$ for all k sufficiently large. First, the convexity of φ implies (6). Second, (7) follows from (2). It remains to check (8). This follows immediately from the strong convexity of $\varphi \in C^{2,\alpha}(\overline{\Omega})$ (recall (2) and the fact that f, g are positive), together with the following lemma. Given a matrix $C = [c_{ij}]$, denote

$$||C|| = \max_{i,j} |c_{ij}|.$$

LEMMA 3.4. *Let $\{\{S_i^{(k)}\}_{i=1}^{M(k)}\}_{k\in\mathbb{N}}$ be a sequence of admissible and regular almost-triangulations of Ω. Then (recall* (4)*)*

$$\lim_k \max_{i\in\{1,\ldots,M(k)\}} \left\| H\left(S_i^{(k)}, \left\{\nabla\varphi(x_{i_0}^{(k)}), \ldots, \nabla\varphi(x_{i_n}^{(k)})\right\}\right) - \nabla^2\varphi(x_{i_0}^{(k)}) \right\| = 0.$$

Proof. First, let

$$h = h(k) := \max_{i\in\{1,\ldots,M(k)\}} \operatorname{diam} S_i^{(k)}.$$

By Definition 1.5,

$$\lim_k h = 0. \tag{21}$$

Fix some $i \in \{1, \ldots, M(k)\}$. Then, with A_i and B_i defined as in Definition 1.3 (though now dependent on k, although we omit that from the notation), notice that the (j,l)th entry of $A_i\nabla^2\varphi(x_{i_0})$ is $\left(x_{i_j} - x_{i_0}\right)^T \left(\nabla\partial_l\varphi\left(x_{i_0}\right)\right)$, which is of course equal to $D_{v_j}\left(\partial_l\varphi\right)\left(x_{i_0}\right)$, where D_v denotes the directional derivative in the direction v and where

$$v_j := x_{i_j} - x_{i_0}, \quad j = 1, \ldots, n.$$

Now $\eta_{i_j} = \nabla\varphi\left(x_{i_j}\right)$, so $\eta_{i_j} - \eta_{i_0} = \nabla\varphi\left(x_{i_j}\right) - \nabla\varphi\left(x_{i_0}\right)$, i.e., the (j,l)th entry of B_i is $\partial_l\varphi\left(x_{i_j}\right) - \partial_l\varphi\left(x_{i_0}\right)$.

Next, set

$$\zeta := \partial_l\varphi, \;\; x := x_{i_0}, \;\; y := x_{i_j}, \;\; \tau_j := \|v_j\| = \|x - y\|, \;\; u_j := \tau_j^{-1}v_j.$$

Note that u_j is of unit length and that $D_{v_j}\left(\partial_l\varphi\right)\left(x_{i_0}\right) = \tau_j D_{u_j}\zeta(x)$. We have

$$\begin{aligned}
\left|[B_i]_{jl} - \left[A_i\nabla^2\varphi\left(x_{i_0}\right)\right]_{jl}\right| &= \left|\zeta(y) - \zeta(x) - \tau_j D_{u_j}\zeta(x)\right| \\
&= \left|\int_0^{\tau_j} D_{u_j}\zeta\left(\frac{(\tau_j - t)x + ty}{\tau_j}\right) dt - \tau_j D_{u_j}\zeta(x)\right| \\
&= \left|\int_0^{\tau_j} \left[D_{u_j}\zeta\left(\frac{(\tau_j - t)x + ty}{\tau_j}\right) - D_{u_j}\zeta(x)\right] dt\right| \\
&\le \int_0^{\tau_j} \left|D_{u_j}\zeta\left(\frac{(\tau_j - t)x + ty}{\tau_j}\right) - D_{u_j}\zeta(x)\right| dt \\
&\le \int_0^{\tau_j} C_1 \left\|\left(\frac{(\tau_j - t)x + ty}{\tau_j}\right) - x\right\|^\alpha dt \\
&= C_1 \int_0^{\tau_j} t^\alpha \, dt = C\tau_j^{\alpha+1},
\end{aligned}$$

where $C_1 = ||\varphi||_{C^{2,\alpha}(\overline{\Omega})}$ and $C = C_1/(1+\alpha)$. Now, write

$$A_i = DU,$$

where

$$D := \operatorname{diag}(\tau_1, \ldots, \tau_n).$$

Thus, the rows of U have unit length. By our last inequality,

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

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 rows of U are unit vectors, $|U_{jl}| \le 1$. Since M_{jl} is a polynomial of $(n-1)!$ terms in the U_{jl}, we have that $|M_{jl}| \le (n-1)!$ for all j,l, and hence $\left|\left[U^{-1}\right]_{jl}\right| \le \frac{(n-1)!}{\det U}$. By Definition 1.8, $\det U$ is bounded below by a constant $R > 0$ (independent of k and i), so we have that $\left|\left[U^{-1}\right]_{jl}\right| \le R'$ for $R' = R^{-1}(n-1)! > 0$ (independent of k and i). Then it follows that

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

Of course, since $A_i = DU$, this means precisely that

$$\max_{i=1,\ldots,M(k)}\left\|A_i^{-1}B_i - \nabla^2\varphi\left(x_{i_0}\right)\right\| \le C'h^{\alpha}$$

for some $C' > 0$ independent of k and i. Since $\nabla^2\varphi\left(x_{i_0}\right)$ is symmetric,

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

Thus,

$$\begin{aligned}\max_i\left\|\frac{1}{2}A_i^{-1}B_i + \frac{1}{2}\left(A_i^{-1}B_i\right)^T - \nabla^2\varphi\left(x_{i_0}\right)\right\| &\le \max_i\left\|\frac{1}{2}A_i^{-1}B_i - \frac{1}{2}\nabla^2\varphi\left(x_{i_0}\right)\right\| \\ &\quad + \max_i\left\|\frac{1}{2}\left(A_i^{-1}B_i\right)^T - \frac{1}{2}\nabla^2\varphi\left(x_{i_0}\right)\right\| \\ &= \max_i\left\|A_i^{-1}B_i - \nabla^2\varphi\left(x_{i_0}\right)\right\| \le C'h^{\alpha},\end{aligned}$$

which, by (21), concludes the proof of Lemma 3.4. ∎

Remark 3.5. It is tempting to rephrase the regularity assumption (Definition 1.8) in terms of eigenvalues instead of determinant; however, the matrices U and $A^{-1}B$ are not symmetric in general, and so the more involved argument we used seems to be necessary to prove Lemma 3.4.

(ii) Given that the feasibility conditions (6)–(8) hold, d_k is well-defined. The rest of the proof is devoted to showing that d_k converges to zero.

Let

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

denote the barycenter of S_i. Since f is uniformly continuous on $\overline{\Omega}$,

$$\max_{i=1,\ldots,M(k)} |f(x_{i_0}^{(k)}) - f(y_i^{(k)})| \to 0. \tag{22}$$

Let $z_i^{(k)}$ denote the barycenter of the simplex formed by the gradients at the vertices of the ith simplex, i.e.,

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

Then similarly, since g is uniformly continuous on $\overline{\Lambda}$ and $\nabla\varphi$ is Lipschitz,

$$\max_{i=1,\ldots,M(k)} |g(\nabla\varphi(x_{i_0}^{(k)})) - g(z_i^{(k)})| \to 0, \tag{23}$$

Since f, g are bounded away from zero and infinity on $\overline{\Omega}$, it follows from (22) and (23) that

$$\max_i \left| \left[f(y_i^{(k)})/g(z_i^{(k)}) \right]^{1/n} - \left[f(x_{i_0}^{(k)})/g(\nabla\varphi(x_{i_0}^{(k)})) \right]^{1/n} \right| \to 0. \tag{24}$$

By (2),

$$\det \nabla^2\varphi(x_{i_0}^{(k)}) = \frac{f(x_{i_0}^{(k)})}{g(\nabla\varphi(x_{i_0}^{(k)}))}. \tag{25}$$

Then, by (17), (9), and (10), we have

$$\begin{aligned}
&(\mathrm{vol}(\Omega))^{-1} d_k \\
&\le \max_i \left| \det{}^{1/n} H\left(S_i^{(k)}, \left\{\nabla\varphi(x_{i_0}^{(k)}), \ldots, \nabla\varphi(x_{i_n}^{(k)})\right\}\right) - \left[f(y_i^{(k)})/g(z_i^{(k)})\right]^{1/n} \right| \\
&\le \max_i \left| \det{}^{1/n} H\left(S_i^{(k)}, \left\{\nabla\varphi(x_{i_0}^{(k)}), \ldots, \nabla\varphi(x_{i_n}^{(k)})\right\}\right) - \left[f(x_{i_0}^{(k)})/g(\nabla\varphi(x_{i_0}^{(k)}))\right]^{1/n} \right| \\
&\quad + \max_i \left| \left[f(y_i^{(k)})/g(z_i^{(k)})\right]^{1/n} - \left[f(x_{i_0}^{(k)})/g(\nabla\varphi(x_{i_0}^{(k)}))\right]^{1/n} \right|.
\end{aligned}$$

In the last expression, the last term tends to zero with k by (24). Meanwhile, the first term tends to zero with k by Lemma 3.4 and (25) (note here that since $\nabla^2\varphi(\overline{\Omega})$ is compact and entirely contained in the set of positive definite matrices,

$$\det H\left(S_i^{(k)}, \left\{\nabla\varphi(x_{i_0}^{(k)}), \ldots, \nabla\varphi(x_{i_n}^{(k)})\right\}\right)$$

is bounded away from zero for all $k \gg 1$ by Lemma 3.4). ∎

4. Proof of the convergence theorem. We now turn to the proof of Theorem 1.6, stating that the potentials $\phi^{(k)}$ (15) converge to the Brenier potential φ. This section is organized as follows. In sections 4.1–4.2 we define the barycentric extension of the gradient of the optimization potentials and show how this relates to the discrete Jacobian on each simplex (Lemma 4.1). This sets the stage for the remainder of the proof, which occupies the rest of this lengthy section. In section 4.3 we describe the strategy for the proof. The proof itself occupies sections 4.4–4.10.

Let D be a closed ball such that

$$\operatorname{int} D \supset \overline{\Omega}, \tag{26}$$

where $\operatorname{int} A$ denotes the interior of a set A. By the Arzelà–Ascoli theorem, since $\{\phi^{(k)}\}_k$ is an equicontinuous, uniformly bounded family (recall (11)–(12) and note that $\eta_j^{(k)} \in \overline{\Lambda}$ for all k, j, with Λ bounded, and $\phi^{(k)}(0) = 0$), it has a uniformly converging subsequence. Thus, to prove Theorem 1.6 it suffices to show that every subsequence of $\phi^{(k)}$ that converges uniformly on D converges to φ on $\overline{\Omega}$.

Thus, assume that

$$\phi^{(k)} \to \phi \text{ uniformly on } D \tag{27}$$

for some ϕ, and we need only show that $\phi = \varphi$ on Ω. Notice that ϕ is convex and continuous as a uniform limit of continuous uniformly bounded convex functions.

4.1. Barycentric extension of the gradient of the optimization potentials. 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, and so only $C^{0,1}$ and no better. In order to get a handle on the *second-order convergence* of the $\phi^{(k)}$, we will replace the piecewise constant but discontinuous subdifferentials of $\phi^{(k)}$ with continuous, piecewise-affine functions that interpolate rather than jump, which we may then differentiate once again.

For the remainder of the article, let

$$\psi_1^{(k)}, \dots, \psi_{N(k)}^{(k)} \in \mathbb{R} \quad \text{and} \quad \eta_1^{(k)}, \dots, \eta_{N(k)}^{(k)} \in \overline{\Lambda} \subset \mathbb{R}^n \tag{28}$$

denote the solution of the kth DMAOP (Definition 1.3) associated to the data

$$(\Omega, \Lambda, f, g, \{x_j^{(k)}\}_{j=1}^{N(k)}, \{S_i^{(k)}\}_{i=1}^{M(k)}).$$

Thus, with (14) in mind, we define a vector-valued function $G^{(k)}$ by barycentrically interpolating the values $\{\eta_{i_j}^{(k)}\}_{j=1}^n$ over the ith simplex $S_i^{(k)}$ for all $i = 1, \dots, M(k)$. Namely, for each x in

$$S_i^{(k)} = \operatorname{co}(x_{i_0}^{(k)}, \dots, x_{i_n}^{(k)})$$

write

$$x = \sum_{j=0}^n \sigma_j x_{i_j}^{(k)}, \tag{29}$$

with $\sigma_j \in [0,1]$. Then

$$G^{(k)}(x) := \sum_{j=0}^n \sigma_j \eta_{i_j}^{(k)} \qquad \text{if } \; x \in S_i^{(k)} \tag{30}$$

(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 in the kth almost-triangulation and satisfies $G^{(k)}(x_i^{(k)}) = \eta_i^{(k)}$ for all $i = 1, \dots, N(k)$.

4.2. The motivation for defining the barycentric extension. Next, we explain the main role the functions $G^{(k)}$ play.

Let

$$i^{(k)} : \bigcup_{i=1}^{M(k)} \operatorname{int} S_i^{(k)} \to \{1, \ldots, M(k)\} \tag{31}$$

denote the map assigning to a point the index of the unique simplex in the kth almost-triangulation containing it, i.e., $i^{(k)}\big(\operatorname{int} S_j^{(k)}\big) = j$. Define a (locally constant) matrix-valued function

$$\mathcal{H}^{(k)} : \bigcup_{i=1}^{M(k)} \operatorname{int} S_i^{(k)} \to \operatorname{Sym}^2(\mathbb{R}^n)$$

by

$$\mathcal{H}^{(k)}(x) := H^{(k)}_{i^{(k)}(x)}, \tag{32}$$

where (recall (4))

$$H_j^{(k)} := H\Big(S_j^{(k)}, \Big\{\eta_{j_0}^{(k)}, \ldots, \eta_{j_n}^{(k)}\Big\}\Big).$$

Define also

$$\tau^{(k)}(x) := \text{ the barycenter of the simplex } S^{(k)}_{i^{(k)}(x)} \tag{33}$$

and

$$\gamma^{(k)}(x) := \text{the mean of the } \eta_j^{(k)} \text{ associated to the vertices of simplex } S^{(k)}_{i^{(k)}(x)}. \tag{34}$$

Finally, recalling (9), we define a (locally constant) per-simplex penalty function

$$\mathcal{C}^{(k)}(x) := F^{(k)}_{i^{(k)}(x)}\Big(\big\{\psi_j^{(k)}, \eta_j^{(k)}\big\}_{i=1}^{N(k)}\Big). \tag{35}$$

By (9),

$$\mathcal{C}^{(k)}(x) = \max\Big\{0, -\det{}^{1/n}\Big[\mathcal{H}^{(k)}(x)\Big] + f^{1/n}(\tau^{(k)}(x)) \cdot g^{-1/n}(\gamma^{(k)}(x))\Big\}, \tag{36}$$

By the definition of the optimal cost (20),

$$\begin{aligned} c_k &= \int_{\bigcup_{i=1}^{M(k)} S_i} \mathcal{C}^{(k)}(x)\, dx \\ &= \int_{\bigcup_{i=1}^{M(k)} S_i} \max\Big\{0, -\det{}^{1/n}\Big[\mathcal{H}^{(k)}(x)\Big] + f^{1/n}(\tau^{(k)}(x)) \cdot g^{-1/n}(\gamma^{(k)}(x))\Big\}\, dx. \end{aligned} \tag{37}$$

The following result is the motivation for introducing the functions $G^{(k)}$. When combined with (37), it relates second-order information that we can extract from $\phi^{(k)}$ (via $G^{(k)}$) with the cost c_k, over which we have control by the assumptions of Theorem 1.6. In fact, we have $c_k \to 0$, so we can hope that in some sense, as k becomes large, $\phi^{(k)}$ approaches a subsolution of the Monge-Ampère equation.

LEMMA 4.1. *For* $x \in \bigcup_i \operatorname{int} S_i^{(k)}$, $\mathcal{H}^{(k)}(x) = \nabla G^{(k)}(x) + \left(\nabla G^{(k)}(x)\right)^T$.

Proof. We fix some k and then omit k from our notation in the remainder of the proof. We also fix $i \in \{1, \dots, M(k)\}$ and work within the simplex S_i, i.e., assume that $x \in S_i$, i.e., $i^{(k)}(x) = i$. Now let $v_j = x_{i_j} - x_{i_0}$. We claim that

$$D_{v_j} G(x) = \eta_{i_j} - \eta_{i_0}, \quad x \in \operatorname{int} S_i.$$

Intuitively, this is because g is affine on S_i with $G\left(x_{i_j}\right) = \eta_{i_j}$. For the proof, recall the definition of the functions σ_j from (29). Then, letting $\delta_{st} = 1$ if $s = t$ and zero otherwise,

$$\begin{aligned} D_{v_j} G(x) &= \frac{d}{dt}\Big|_{t=0} G\big(x + t(x_{i_j} - x_{i_0})\big) \\ &= \sum_{s=0}^{n} \frac{d}{dt}\Big|_{t=0} \sigma_s\big(x + t(x_{i_j} - x_{i_0})\big)\eta_{i_s} \\ &= \sum_{s=0}^{n} \frac{d}{dt}\Big|_{t=0} \big(\sigma_s(x) + t(\delta_{js} - \delta_{0s})\big)\eta_{i_s} \\ &= \eta_{i_j} - \eta_{i_0}, \end{aligned}$$

as claimed.

Now $D_{v_j} G = v_j \cdot \nabla G$, so $D_{v_j} G$ is the jth row of $A_i \nabla G$, where ∇G denotes the matrix with jth row $\frac{\partial}{\partial x_j} G$ and where A_i is as in (3). Since $D_{v_j} G = \eta_{i_j} - \eta_{i_0}$ is also the jth row of B_i, we have that $B_i = A_i \nabla G$, i.e., $\nabla G = A_i^{-1} B_i$. The statement now follows from the definition of H_i (4). ∎

4.3. Strategy for the proof. In this subsection we outline the strategy for the proof of Theorem 1.6.

The results of the previous subsection indicate that the optimization potentials should be approximate subsolutions of the Monge–Ampère equation. Since the optimization potentials converge to ϕ, this gives some hope that ϕ itself might be such a subsolution. To make this rigorous we regularize. Let ξ_ε be a standard set of mollifiers (supported on $B_\varepsilon(0)$). Notice that $G^{(k)}$ and $\mathcal{H}^{(k)}$ are only defined on the almost-triangulation of Ω, so we run into trouble near the boundary when convolving with ξ_ε. Thus, we will work with the regions Ω_ϵ given by Definition 1.5.

LEMMA 4.2. *Fix* $\varepsilon > 0$. *As* $k \to \infty$, $\mathcal{H}^{(k)} \star \xi_\varepsilon(x)$ *converges uniformly to* $\nabla^2(\phi \star \xi_\varepsilon)$ *on* Ω_ε.

The proof of Lemma 4.2 takes place in section 4.5. Lemma 4.2 gives us control over the second-order behavior of $\phi \star \xi_\varepsilon$. The proof uses an auxiliary result established in section 4.4 that gives control over the first-order behavior of $\phi \star \xi_\varepsilon$.

The next step of the proof involves taking the limits in k both in the previous lemma and in (37). Thanks to the fact that $c_k \to 0$, this yields the following statement roughly saying that $\phi \star \xi_\varepsilon$ is an approximate subsolution to the Monge–Ampère equation, i.e., that $\nabla(\phi \star \xi_\varepsilon)$ cannot "excessively" shrink volume.

LEMMA 4.3. *Fix* $\varepsilon > 0$. *For* $x \in \Omega_\varepsilon$,

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

The proof of Lemma 4.3 is presented in section 4.6.

The next step in the proof is to take the limit $\varepsilon \to 0$ and show that ϕ must be a weak solution in the sense that $\nabla\phi$ pushes forward μ to ν. The proof of this fact is broken up into several steps in sections 4.7–4.8. First, we define the measures

$$\nu_\varepsilon := (\nabla\phi \star \xi_\varepsilon)_{\#}\mu|_{\Omega_\varepsilon} \tag{38}$$

obtained by pushing forward the restriction of μ to Ω_ε by $\nabla\phi\star\xi_\varepsilon$. Denote the density of these measures by

$$g_\varepsilon\, dx := \nu_\varepsilon. \tag{39}$$

Using Lemma 4.3, we show that a subsequence of these measures (roughly speaking) converges weakly to the target measure ν. Intuitively speaking, Lemma 4.3 says that $\nabla\phi \star \xi_\varepsilon$ does not shrink volume "excessively" at any point. Combining this with the fact that the image of $\nabla\phi \star \xi_\varepsilon$ must lie within $\overline{\Lambda}$ motivates the convergence. The precise result we prove is the following.

PROPOSITION 4.4. *For any sequence $\varepsilon \to 0$, $\mu(\Omega_\varepsilon)^{-1}\nu_\varepsilon$ is a sequence of probability measures converging weakly to $\nu(\Lambda)^{-1}\nu$.*

The proof of Proposition 4.4 is given in section 4.8 based on some auxiliary results proven in section 4.7.

The last step of the proof of Theorem 1.6 is to show that any uniform limit of the optimization potentials coincides with the Brenier potential φ.

LEMMA 4.5. *Let $\phi^{(k)}$ be defined by* (15) *and suppose that $\phi^{(k)}$ converges uniformly to some ϕ. Then $\phi = \varphi$.*

The proof of Lemma 4.5 is presented in section 4.9. It hinges on Proposition 4.4, stability results for optimal transport maps (proved in section 4.10), and all of the previous steps in the proof.

Remark 4.6. 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.

4.4. First-order control on ϕ. We want to show that $G^{(k)}$ approaches $\partial\phi$ in some sense. We make use of the following semicontinuity result of Bagh and Wets [2, Theorem 8.3] (cf. [33, Theorem 24.5]). Recall that f_k epi-converges to f (roughly) if the epigraphs of f_k converge to the epigraph of f; we refer the reader to [34, p. 240] for more precise details.

THEOREM 4.7. *Let f and $\{f_k\}_{k\in\mathbb{N}}$ be lower semicontinuous convex functions with f_k epi-converging to f. Fix $x \in \operatorname{int}\operatorname{dom} f$ and $\epsilon > 0$. Then there exist $\delta > 0$ and $K \in \mathbb{N}$ such that*

$$\partial f_k(y) \subset \partial f(x) + B_\epsilon(0) \quad \forall y \in B_\delta(x), \forall k \geq K.$$

Moreover, if f is differentiable at x, then

$$\lim_{k\to\infty} \partial f_k(x) = \{\nabla f(x)\}. \tag{40}$$

LEMMA 4.8. *As k tends to infinity, $G^{(k)}$ converges to $\nabla\phi$ almost everywhere on Ω.*

Proof. If a sequence of lower semicontinuous finite convex functions converges uniformly on bounded sets to some convex function, then the sequence epi-converges to this function [34, Theorem 7.17]. Thus, we may apply Theorem 4.7 to $\phi^{(k)}$. Fix $x \in \Omega$ and $\epsilon > 0$. There exists a $\delta > 0$ and $K \in \mathbb{N}$ such that

$$\partial\phi^{(k)}(y) \subset \partial\phi(x) + B_\epsilon(0) \quad \forall y \in B_\delta(x), \forall k \geq K.$$

Fix $\epsilon > 0$ and a point $x \in \Omega$ where ϕ is differentiable. 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

$$\partial\phi^{(k)}(x_j^{(k)}) \subset \nabla\phi(x) + B_\varepsilon(0).$$

By (14), $\eta_{i_j}^{(k)} \in \partial\phi^{(k)}(x_{i_j}^{(k)})$. On the other hand, by (30), $G^{(k)}(x)$ is a convex combination of the $\eta_{i_j}^{(k)}$. Thus, $G^{(k)}(x) \in \nabla\phi(x) + B_\varepsilon(0)$. This proves that $G^{(k)} \to \nabla\phi$ almost everywhere since ϕ is differentiable almost everywhere. ∎

4.5. Second-order control on ϕ and a proof of Lemma 4.2. 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, we will work with the regions Ω_ϵ given by Definition 1.5.

The following lemma is the main result of this subsection.

LEMMA 4.9. *Fix $\varepsilon > 0$. On Ω_ε, $\nabla G^{(k)} \star \xi_\varepsilon$ converges uniformly to $\nabla^2(\phi \star \xi_\varepsilon)$ (in each of the n^2 components).*

Lemma 4.9 immediately implies Lemma 4.2, thanks to Lemma 4.1 and symmetrization (noting the symmetry of $\nabla^2(\phi \star \xi_\varepsilon)$).

We start with three auxiliary results. The first states that differentiation and convolution commute when the functions involved are uniformly Lipschitz. We leave the standard proof to the reader. Note that $\nabla\phi \star \xi_\varepsilon := (\nabla\phi) \star \xi_\varepsilon$ is everywhere defined because $\nabla\phi$ exists almost everywhere.

CLAIM 4.10. *For all $\varepsilon > 0$, $\nabla\phi \star \xi_\varepsilon = \nabla(\phi \star \xi_\varepsilon)$ and $\nabla G^{(k)} \star \xi_\varepsilon = \nabla(G^{(k)} \star \xi_\varepsilon)$.*

The second is a mollified version of Lemma 4.8.

CLAIM 4.11. *Fix $\varepsilon > 0$. For $x \in \Omega_\varepsilon$, $G^{(k)} \star \xi_\varepsilon$ converges uniformly to $\nabla\phi \star \xi_\varepsilon$ (in each of the n components).*

Proof. First, we claim pointwise convergence, i.e., that

$$\lim_{k\to\infty} G^{(k)} \star \xi_\varepsilon(x) = \nabla\phi \star \xi_\varepsilon(x) \quad \text{for each } x \in \Omega_\varepsilon. \tag{41}$$

To check that this is true, note that for $x \in \Omega_\varepsilon$,

$$\begin{aligned}\Big|G^{(k)} \star \xi_\varepsilon(x) - \nabla\phi \star \xi_\varepsilon(x)\Big| &= \Big|\int \big(G^{(k)}(y) - \nabla\phi(y)\big)\xi_\varepsilon(y-x)\,dy\Big| \\ &\le \int \big|G^{(k)}(y) - \nabla\phi(y)\big|\xi_\varepsilon(y-x)\,dy\end{aligned}$$

(once again note that these integrals make sense since $\nabla\phi$ exists almost everywhere). Since $|G^{(k)}(y) - \nabla\phi(y)| \to 0$ almost everywhere by Lemma 4.8, while $|\xi_\varepsilon| \le C$ and the $G^{(k)}$ are uniformly bounded, equation (41) follows from bounded convergence (note that ε is constant in this limit).

Next, notice that the $G^{(k)}$ are uniformly bounded independently of k, in fact (since we may assume, without loss of generality, that $0 \in \Lambda$),

$$|G^{(k)}| \le \operatorname{diam}\Lambda \tag{42}$$

by (30). Write

$$G^{(k)} = (G_1^{(k)}, \ldots, G_n^{(k)}). \tag{43}$$

Thus, $G_i^{(k)} \star \xi_\varepsilon$, $i = 1, \ldots, n$, are uniformly bounded independently of k,

$$\begin{aligned}|\nabla(G_i^{(k)} \star \xi_\varepsilon)| &= |G_i^{(k)} \star \nabla\xi_\varepsilon(x)| \\ &= \Big|\int G_i^{(k)}(y)\nabla\xi_\varepsilon(x-y)\,dy\Big| \\ &\le \operatorname{diam}\Lambda\Big|\int \nabla\xi_\varepsilon(x-y)\,dy\Big| \le C,\end{aligned}$$

with $C = C(\Lambda, \varepsilon)$. Thus, the $G^{(k)} \star \xi_\varepsilon$ have uniformly bounded derivatives (in each component). The statement of the claim now follows from Remark 4.12 below. ∎

Remark 4.12. We will use the following fact more than once. If a uniformly bounded sequence of differentiable functions $\{f_k\}_{k\in\mathbb{N}}$ with uniformly bounded (first) derivatives on compact sets satisfies $f_k \to f$ pointwise, then $f_k \to f$ uniformly on compact sets. (This can be established easily using the Arzelà–Ascoli theorem.)

The third auxiliary result is a one-variable interpolation-type result.

CLAIM 4.13. *Let $I \subset \mathbb{R}$ be a closed interval, and let $\{f_k\}_{k\in\mathbb{N}}$, $f : I \to \mathbb{R}$ be smooth functions such that* (i) *$f_k \to f$ uniformly,* (ii) *the f_k'' are uniformly bounded independently of k, and* (iii) *f'' is bounded. Then $f_k' \to f'$ uniformly.*

Proof. We make use of the Landau–Kolmogorov inequality

$$\|g'\|_\infty \le C\|g\|_\infty^{1/2}\|g''\|_\infty^{1/2} \tag{44}$$

for $g \in C^2(I)$ (see, e.g., [12]). Apply (44) to $g = f_k - f$ to obtain $\|f_k' - f'\|_\infty \le C\|f_k - f\|_\infty^{1/2}\|f_k'' - f''\|_\infty^{1/2}$. Since f'' is bounded and the f_k'' are uniformly bounded, we have that $\|f_k'' - f''\|_\infty^{\frac{1}{2}}$ is uniformly bounded in k. Of course, $\|f_k - f\|_\infty \to 0$ by uniform convergence. Therefore, $\|f_k' - f'\|_\infty \to 0$, as desired. ∎

Proof of Lemma 4.9. The functions (recall (43))

$$\partial_j^2(G_i^{(k)} \star \xi_\varepsilon) = G_i^{(k)} \star \partial_j^2 \xi_\varepsilon, \quad i, j \in \{1, \ldots, n\},$$

are uniformly bounded independently of k by a constant depending on ε (by the uniform boundedness of $G_i^{(k)}$—recall (42)). By Claim 4.10, $\nabla\phi \star \xi_\varepsilon = \nabla(\phi \star \xi_\varepsilon)$, so also $\nabla(\nabla\phi \star \xi_\varepsilon) = \nabla^2(\phi \star \xi_\varepsilon)$. Thus, since $\phi \star \xi_\varepsilon$ is smooth, $\nabla(\nabla\phi \star \xi_\varepsilon)$ is bounded in all of its n^2 components.

Let $x \in \Omega_\varepsilon$. Then fix i, j and let $\delta > 0$ be small enough such that $I := \{x + te_j : t \in [-\delta, \delta]\} \subset \Omega_\varepsilon$. By Claim 4.11, $G_i^{(k)} \star \xi_\varepsilon \to \partial_i \phi \star \xi_\varepsilon$ uniformly. Restricting ourselves to the jth variable and applying Claim 4.13,

$$\partial_j(G_i^{(k)} \star \xi_\varepsilon) \to \partial_j(\partial_i \phi \star \xi_\varepsilon)$$

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 now follows from Remark 4.12. Finally, invoking Claim 4.10 implies the statement of Lemma 4.9. ∎

4.6. Obtaining a density inequality for $\phi \star \xi_\varepsilon$ and a proof of Lemma 4.3. In this subsection we prove Lemma 4.3.

First, we prove a mollified version of (37).

LEMMA 4.14. *Fix $\varepsilon > 0$. For k sufficiently large,*

$$\int_{\Omega_\varepsilon} \max\left\{0, -\det{}^{1/n}\big(\mathcal{H}^{(k)} \star \xi_\varepsilon\big) + \left(\frac{f \circ \tau^{(k)}}{g \circ \gamma^{(k)}}\right)^{1/n} \star \xi_\varepsilon\right\} \leq c_k,$$

where c_k is the optimal cost defined in (20)*, and $\tau^{(k)}$ and $\gamma^{(k)}$ are defined in* (33)–(34)*.*

Proof. Note that $\max(0, \cdot)$ is convex, so applying Jensen's inequality to (36) yields

$$\begin{aligned}
\mathcal{C}^{(k)} \star \xi_\varepsilon &= \max\left\{0, -\det{}^{1/n} \mathcal{H}^{(k)} + (f^{1/n} \circ \tau^{(k)}) \cdot (g^{-1/n} \circ \gamma^{(k)})\right\} \star \xi_\varepsilon \\
&\geq \max\left\{0, \Big[-\det{}^{1/n} \mathcal{H}^{(k)} + (f^{1/n} \circ \tau^{(k)}) \cdot (g^{-1/n} \circ \gamma^{(k)})\Big] \star \xi_\varepsilon\right\} \\
&= \max\left\{0, \big[-\det{}^{1/n} \mathcal{H}^{(k)}\big] \star \xi_\varepsilon + \left(\frac{f \circ \tau^{(k)}}{g \circ \gamma^{(k)}}\right)^{1/n} \star \xi_\varepsilon\right\}.
\end{aligned}$$

Now by the convexity of $-\det{}^{1/n}(\cdot)$ on the set of positive semidefinite matrices and Jensen's inequality once again, we have

$$\big[-\det{}^{1/n} \mathcal{H}^{(k)}\big] \star \xi_\varepsilon \geq -\det{}^{1/n}\big(\mathcal{H}^{(k)} \star \xi_\varepsilon\big),$$

and combining the last two inequalities yields

$$\mathcal{C}^{(k)} \star \xi_\varepsilon \geq \max\left\{0, -\det{}^{1/n}\big(\mathcal{H}^{(k)} \star \xi_\varepsilon\big) + \left(\frac{f \circ \tau^{(k)}}{g \circ \gamma^{(k)}}\right)^{1/n} \star \xi_\varepsilon\right\}. \tag{45}$$

Recall from Definition 1.5 that

$$\Omega_\varepsilon + B_\epsilon(0) \subset \bigcup_{i=1}^{M(k)} S_i^{(k)} \quad \forall k \gg 1.$$

Thus for such k sufficiently large, noting that $\mathcal{C}^{(k)} \geq 0$, we have

$$\begin{aligned}\int_{\Omega_\varepsilon} \mathcal{C}^{(k)} \star \xi_\varepsilon &= \int_{\Omega_\varepsilon} \int_{B_\varepsilon(0)} \mathcal{C}^{(k)}(x-y)\xi_\varepsilon(y)\,dy\,dx \\ &= \int_{B_\varepsilon(0)} \xi_\varepsilon(y) \int_{\Omega_\varepsilon} \mathcal{C}^{(k)}(x-y)\,dx\,dy \\ &\leq \int_{B_\varepsilon(0)} \xi_\varepsilon(y) \int_{\bigcup_{i=1}^{M(k)} S_i^{(k)}} \mathcal{C}^{(k)}(x)\,dx\,dy \\ &= c_k,\end{aligned}$$

where the last line follows from (37). Combining with (45) completes the proof. ∎

At least intuitively, in order to prove Lemma 4.3 we need to "take the limit as $k \to \infty$" in Lemma 4.14, so that we can employ Lemma 4.2. The main technical obstacle in taking the limit in k is controlling the behavior of $\tau^{(k)}$ and $\gamma^{(k)}$; the proof relies on tools from convex analysis.

Remark 4.15. Notice that in the case that f and g are uniform densities on Ω and Λ, respectively, the proof of Lemma 4.3 is almost trivial. Even in the case that only g is uniform, the proof is still considerably 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 of Lemma 4.3. Let $\alpha, \beta > 0$ and fix $x \in \Omega_\varepsilon$. Using Theorem 4.7, for every $z \in \overline{B}_\alpha(x)$, there exists $\delta(z) > 0$ and $N_\beta(z)$ such that

$$\partial\phi^{(k)}(y) \subset \partial\phi(z) + B_\beta(0) \quad \forall y \in B_{\delta(z)}(z),\ \forall k \geq N_\beta(z). \tag{46}$$

By compactness, there exist $z_1, \ldots, z_p \in \overline{B}_\alpha(x)$ such that the $B_{\delta(z_i)}$ cover $\overline{B}_\alpha(x)$. Setting

$$N'_{x,\alpha,\beta} := \max_{i\in\{1,\ldots,p\}} N_\beta(z_i),$$

we thus have that

$$\partial\phi^{(k)}(y) \subset \bigcup_{z\in\overline{B}_\alpha(x)} \partial\phi(z) + B_\beta(0) \quad \forall k \geq N'_{x,\alpha,\beta},\ \forall y \in \overline{B}_\alpha(x). \tag{47}$$

For k sufficiently large, i.e.,

$$k \geq N_\alpha$$

for some N_α depending only on α, we have that

$$\text{the simplex } i^{(k)}(x) \text{ (recall (31)) containing } x \text{ is contained in } \overline{B}_\alpha(x) \tag{48}$$

by the admissibility of our sequence of almost-triangulations (Definition 1.5); in particular

$$\tau^{(k)}(x) \in \overline{B}_\alpha(x).$$

Statement (48) also implies, by (14), that $\gamma^{(k)}(x)$ is a convex combination of $n+1$ elements of $\bigcup_{y\in\overline{B}_\alpha(x)} \partial\phi^{(k)}(y)$, so by (47),

$$\gamma^{(k)}(x) \in \mathrm{co}\Bigg(\bigcup_{z\in\overline{B}_\alpha(x)} \partial\phi(z) + B_\beta(0)\Bigg) \quad \forall k \geq N_{x,\alpha,\beta},$$

where

$$N_{x,\alpha,\beta} := \max\{N_\alpha, N'_{x,\alpha,\beta}\}.$$

Thus for $k \geq N_{x,\alpha,\beta}$,

$$\frac{f^{1/n}(\tau^{(k)}(x))}{g^{1/n}(\gamma^{(k)}(x))} \geq \frac{\min\{f^{1/n}(y) : y \in \overline{B}_\alpha(x)\}}{\max\left\{g^{1/n}(z) : z \in \mathrm{co}\left(\bigcup_{z\in\overline{B}_\alpha(x)} \partial\phi(z) + B_\beta(0)\right)\right\}}. \tag{49}$$

For almost every $x \in \Omega_\varepsilon$ we have that for any $\gamma > 0$ there exists

$$C(x,\gamma) > 0$$

such that [33, Corollary 24.5.1]

$$\partial\phi(z) \subset \nabla\phi(x) + B_\gamma(0) = B_\gamma(\nabla\phi(x)) \quad \forall z \in \overline{B}_\alpha(x),\ \forall \alpha \in (0, C).$$

Hence, for $\alpha \in (0, C)$,

$$\bigcup_{z\in\overline{B}_\alpha(x)} \partial\phi(z) + B_\beta(0) \subset B_{\gamma+\beta}(\nabla\phi(x)),$$

implying that

$$\mathrm{co}\Bigg(\bigcup_{z\in\overline{B}_\alpha(x)} \partial\phi(z) + B_\beta(0)\Bigg) \subset B_{\gamma+\beta}(\nabla\phi(x)). \tag{50}$$

Therefore for a.e. x and any $\alpha, \beta, \gamma > 0$ with $\alpha \in (0, C(x,\gamma))$ we have by (49) and (50) that

$$\frac{f^{1/n}(\tau^{(k)}(x))}{g^{1/n}(\gamma^{(k)}(x))} \geq \frac{\min\{f^{1/n}(y) : y \in \overline{B}_\alpha(x)\}}{\max\left\{g^{1/n}(z) : z \in \overline{B}_{\gamma+\beta}(\nabla\phi(x))\right\}} \text{ for } k \geq N_{x,\alpha,\beta}.$$

It follows that

$$\liminf_{k\to\infty} \frac{f^{1/n}(\tau^{(k)}(x))}{g^{1/n}(\gamma^{(k)}(x))} \geq \frac{\min\{f^{1/n}(y) : y \in \overline{B}_\alpha(x)\}}{\max\left\{g^{1/n}(z) : z \in \overline{B}_{\gamma+\beta}(\nabla\phi(x))\right\}} \tag{51}$$

for a.e. x and any $\alpha, \beta, \gamma > 0$ with $\alpha \in (0, C(x,\gamma))$. By the continuity of f and g,

$$\lim_{\alpha\to 0} \min\{f^{1/n}(y) : y \in \overline{B}_\alpha(x)\} = f^{1/n}(x), \tag{52}$$

and for a.e. x also

$$\lim_{(\beta,\gamma)\to 0} \max\left\{g^{1/n}(z) : z \in B_{\gamma+\beta}(\nabla\phi(x))\right\} = g^{1/n}(\nabla\phi(x)). \tag{53}$$

Taking limits in (51) (first $\alpha \to 0$, followed by $(\beta,\gamma) \to 0$) and applying (52) and (53),

$$\liminf_{k\to\infty} \left(\frac{f(\tau^{(k)}(x))}{g(\gamma^{(k)}(x))}\right)^{1/n} \geq \left(\frac{f(x)}{g(\nabla\phi(x))}\right)^{1/n} \qquad \text{for a.e. } x. \tag{54}$$

Next we observe that mollification preserves this inequality in the following sense:

$$\begin{aligned}\liminf_{k\to\infty}\left[\left(\frac{f\circ\tau^{(k)}}{g\circ\gamma^{(k)}}\right)^{1/n}\star\xi_\varepsilon\right](x) &= \liminf_{k\to\infty}\int\xi_\varepsilon(x-y)\cdot\left(\frac{f(\tau^{(k)}(y))}{g(\gamma^{(k)}(y))}\right)^{1/n}dy\\ &\geq \int\xi_\varepsilon(x-y)\cdot\liminf_{k\to\infty}\left(\frac{f(\tau^{(k)}(y))}{g(\gamma^{(k)}(y))}\right)^{1/n}dy\\ &\geq \int\xi_\varepsilon(x-y)\cdot\left(\frac{f(y)}{g(\nabla\phi(y))}\right)^{1/n}dy\\ &= \left[\left(\frac{f}{g\circ\nabla\phi}\right)^{1/n}\star\xi_\varepsilon\right](x), \end{aligned}\tag{55}$$

where we have used the Fatou–Lebesgue theorem (applicable since the sequence of integrands is dominated by an integrable function; indeed, the domain is bounded and, as f and g are bounded away from zero, the integrands are uniformly bounded) to pass the lim inf within the integral and (54) in the penultimate step. Note that (55) makes sense since $\nabla\phi$ exists almost everywhere.

Now, take a lim inf in Lemma 4.14 and use the fact that $c_k\to 0$, along with (55), to see that

$$\begin{aligned}0=\liminf_{k\to\infty}c_k &\geq \liminf_{k\to\infty}\int_{\Omega_\varepsilon}\max\left\{0,-\det{}^{1/n}\big(\mathcal{H}^{(k)}\star\xi_\varepsilon\big)+\left(\frac{f\circ\tau^{(k)}}{g\circ\gamma^{(k)}}\right)^{1/n}\star\xi_\varepsilon\right\}\\ &\geq \int_{\Omega_\varepsilon}\liminf_{k\to\infty}\max\left\{0,-\det{}^{1/n}\big(\mathcal{H}^{(k)}\star\xi_\varepsilon\big)+\left(\frac{f\circ\tau^{(k)}}{g\circ\gamma^{(k)}}\right)^{1/n}\star\xi_\varepsilon\right\}\\ &\geq \int_{\Omega_\varepsilon}\max\left\{0,-\limsup_{k\to\infty}\det{}^{1/n}\big(\mathcal{H}^{(k)}\star\xi_\varepsilon\big)+\left(\frac{f}{g\circ\nabla\phi}\right)^{1/n}\star\xi_\varepsilon\right\}.\end{aligned}\tag{56}$$

Note that we have passed the lim inf inside of the integral using Fatou's lemma (applicable since the integrands are nonnegative), and we have used the fact that

$$\liminf_k \max(0,a_k)=\max(0,\liminf_k a_k),$$

which follows from the monotonicity and continuity of $\max(0,\cdot)$.

Now it follows from (56) that

$$\limsup_{k\to\infty}\det{}^{1/n}\big(\mathcal{H}^{(k)}\star\xi_\varepsilon\big)\geq\left(\frac{f}{g\circ\nabla\phi}\right)^{1/n}\star\xi_\varepsilon\tag{57}$$

almost everywhere on Ω_ε. But Lemma 4.2 implies that in fact

$$\det\big(\mathcal{H}^{(k)}\star\xi_\varepsilon\big)$$

is pointwise convergent on Ω_ε as $k\to\infty$. Then we conclude that the sequence on the left-hand side of (57) is actually convergent almost everywhere in Ω_ε and

$$\left[\lim_{k\to\infty}\det\big(\mathcal{H}^{(k)}\star\xi_\varepsilon\big)\right]^{1/n}\geq\left(\frac{f}{g\circ\nabla\phi}\right)^{1/n}\star\xi_\varepsilon$$

almost everywhere in Ω_ε. Of course, by Lemma 4.2 we then have

$$\det{}^{1/n} \nabla^2(\phi \star \xi_\varepsilon) \geq \left(\frac{f}{g \circ \nabla\phi}\right)^{1/n} \star \xi_\varepsilon$$

almost everywhere in Ω_ε and in fact, by the continuity of both sides of the inequality, everywhere in Ω_ε. This implies (since ξ_ε is supported on $B_\varepsilon(0)$)

$$\det{}^{1/n} \nabla^2(\phi \star \xi_\varepsilon)(x) \geq \left(\frac{\inf\{f(y) : y \in B_\varepsilon(x)\}}{\sup\{g\left(\nabla\phi(y)\right) : y \in B_\varepsilon(x), \nabla\phi(y) \text{ exists}\}}\right)^{1/n}$$

for $x \in \Omega_\varepsilon$, completing the proof of Lemma 4.3. □

4.7. Passing to the limit in ε—part I. In this subsection we prove Proposition 4.4.

Let

$$\Lambda_\varepsilon := \nabla\phi_\varepsilon(\Omega_\varepsilon). \tag{58}$$

Because Λ is convex, $\Lambda_\varepsilon \subset \overline{\Lambda}$.

To ease the notation in the following, we set

$$\phi_\varepsilon := \phi \star \xi_\varepsilon.$$

CLAIM 4.16. *Fix $\varepsilon > 0$. Let g_ε be defined as* (39). *For $y \in \Lambda_\varepsilon$,*

$$g_\varepsilon(y) = f\big((\nabla\phi_\varepsilon)^{-1}(y)\big) / \det \nabla^2\phi_\varepsilon\big((\nabla\phi_\varepsilon)^{-1}(y)\big).$$

Proof. Since convolution with a nonnegative kernel preserves convexity, ϕ_ε is a smooth convex function. Moreover, Lemma 4.3 in fact implies that ϕ_ε is uniformly convex on any compact subset of Ω_ε, so $\nabla\phi_\varepsilon$ is invertible on Ω_ε. This is evident if Ω_ε is convex, but is also true in general. 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 of ϕ_ε in the direction of a unit vector parallel to this line must be nonnegative along this line and strictly positive near both x and y (recall that for ε small, ϕ_ε is defined, convex, and finite on a ball containing Ω, in particular on the convex hull of Ω, and uniformly convex when restricted to any compact subset of Ω_ε). In conclusion, $\nabla\phi_\varepsilon(x)$ and $\nabla\phi_\varepsilon(y)$ cannot agree. Thus,

$$(\nabla\phi_\varepsilon)^{-1} \text{ exists on } \Omega_\varepsilon \tag{59}$$

as claimed. The standard formula for the push-forward of a measure and the definitions (38)–(39) then imply the statement. □

Let

$$\overline{g}_\varepsilon(y) := \begin{cases} f\big((\nabla\phi_\varepsilon)^{-1}(y)\big)\dfrac{\sup\left\{g(\nabla\phi(x)) : x \in B_\varepsilon\big((\nabla\phi_\varepsilon)^{-1}(y)\big), \nabla\phi(x) \text{ exists}\right\}}{\inf\left\{f(x) : x \in B_\varepsilon\big((\nabla\phi_\varepsilon)^{-1}(y)\big)\right\}}, & y \in \Lambda_\varepsilon, \\ g, & y \in \overline{\Lambda}\backslash\Lambda_\varepsilon. \end{cases} \tag{60}$$

CLAIM 4.17. *On $\overline{\Lambda}$,*

$$g_\varepsilon \leq \overline{g}_\varepsilon. \tag{61}$$

Proof. The inequality is trivial on $\overline{\Lambda}\backslash\Lambda_\varepsilon$ since by the definitions (38), (39), and (58),

$$g_\varepsilon = 0 \text{ outside of } \Lambda_\varepsilon.$$

On the other hand, Lemma 4.3 and Claim 4.16 precisely imply (61) on Λ_ε. ∎

For any ψ convex defined on a convex set C, let

$$\mathrm{dom}(\psi)$$

denote the set of points in C at which ψ is finite, and let

$$\mathrm{sing}(\psi) \tag{62}$$

denote the set of points in $\mathrm{int}\,\mathrm{dom}(\psi)$ where ψ is not differentiable. Similarly, denote by

$$\Delta(\psi) \tag{63}$$

the complement of $\mathrm{sing}(\psi)$ in $\mathrm{int}\,\mathrm{dom}(\psi)$. Note that $\Delta(\psi)$ has full measure in $\mathrm{int}\,\mathrm{dom}(\psi)$ because convex functions are locally Lipschitz.

LEMMA 4.18. *As $\varepsilon \to 0$, $\overline{g}_\varepsilon \to g$ almost everywhere on $\overline{\Lambda}\backslash\partial\phi\,(\mathrm{sing}(\phi))$.*

Remark 4.19. As in Remark 4.15, the proof of this lemma becomes considerably easier in the case in which g is a uniform density and trivial when both f and g are uniform densities.

Before proving Lemma 4.18 we make several technical remarks. Recall from the beginning of the proof (see (26)) that ϕ is taken to be defined on a ball D containing $\overline{\Omega}$ in its interior, and in fact ϕ is the uniform limit of the $\phi^{(k)}$ on D. Accordingly, ϕ is convex and continuous on D, and $\mathrm{dom}(\phi) = D$. Likewise $\mathrm{dom}(\phi_\varepsilon) = D_\varepsilon$, where D_ε is the closed ball (concentric with D) of radius ε less than that of D.

Thus far, we have only studied the behavior of ϕ inside of Ω as there has been no need to consider its behavior elsewhere. However, since Ω may not be convex, it is important to consider ϕ as being defined on a (larger) convex set in order to employ the language and results of convex analysis.

Then the convex conjugate of ϕ,

$$\phi^*(y) := \sup_{x\in D}[\langle x, y\rangle - \phi(x)],$$

is finite on all of $\mathbb{R}^n$, i.e., $\mathrm{dom}(\phi^*) = \mathbb{R}^n$.

The next claim collects basic properties concerning the Legendre dual that we will need later.

CLAIM 4.20. (i) $\nabla\phi^*_\varepsilon = (\nabla\phi_\varepsilon)^{-1}$ *on* Λ_ε,
(ii) $\nabla\phi^*_\varepsilon \to \nabla\phi^*$ *pointwise on* $\Delta(\phi^*)$.

Proof. (i) This follows from (59) and the standard formula for the gradient of the Legendre dual of a smooth strongly convex function [33, Theorem 26.5], noting that, by Lemma 4.3, ϕ_ε is indeed strongly convex on compact subsets of Ω_ε.

(ii) Note that $\phi_\varepsilon \to \phi$ uniformly on compact subsets of $\mathrm{int}\,\mathrm{dom}(\phi) = \mathrm{int}\,D$. Then by [34, Theorem 7.17], the ϕ_ε epi-converge to ϕ (we again recall that f_k epi-converges to f (roughly) if the epigraphs of f_k converge to the epigraph of f; see [34, p. 240] for more precise details). Then by [34, Theorem 11.34], we have that the ϕ^*_ε epi-converge to ϕ^*. Again using Theorem 4.7, we have that $\partial\phi^*_\varepsilon(x) \to \nabla\phi^*(x)$ for all $x \in \mathrm{int}\,\mathrm{dom}\,\phi^* = \mathbb{R}^n$ such that $\nabla\phi^*(x)$ exists. ∎

Proof of Lemma 4.18. It suffices to assume that $y \in \Delta(\phi^*) \cap \left(\overline{\Lambda} \backslash \partial\phi\left(\mathrm{sing}(\phi)\right)\right)$, since otherwise y is contained in a measure zero set. Thus, by Claim 4.20(ii),

$$\lim_{\varepsilon\to 0} \nabla\phi^*_\varepsilon(y) = \nabla\phi^*(y).$$

Let

$$E := \{\varepsilon > 0 \,:\, y \in \Lambda_\varepsilon\}.$$

Since $\overline{g}_\varepsilon(y) = g(y)$ whenever $\varepsilon \notin E$ (recall (60)), it suffices to show that $\overline{g}_{\varepsilon_j}(y) \to g(y)$ for all sequences $\varepsilon_j \in E$ that tend to zero. Let ε_j be such a sequence. Notice that since $y \in \Lambda_{\varepsilon_j}$ for all j, by Claim 4.20(i), $\nabla\phi^*_{\varepsilon_j}(y) = (\nabla\phi_{\varepsilon_j})^{-1}(y)$ for all j, so

$$\lim_{j\to\infty} \left(\nabla\phi_{\varepsilon_j}\right)^{-1}(y) = \nabla\phi^*(y). \tag{64}$$

Let $\gamma > 0$. Then by (64) there exists $N \in \mathbb{N}$ so that

$$\left|\left(\nabla\phi_{\varepsilon_j}\right)^{-1}(y) - \nabla\phi^*(y)\right| < \gamma/2 \quad \forall j \geq N.$$

Take N large enough so that $\varepsilon_j < \gamma/2$ for all $j \geq N$, so then

$$B_{\varepsilon_j}\left(\left(\nabla\phi_{\varepsilon_j}\right)^{-1}(y)\right) \subset B_\gamma\left(\nabla\phi^*(y)\right) \cap \Omega \quad \forall j \geq N.$$

Thus,

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

Then taking the lim inf as $j \to \infty$ followed by the limit as $\gamma \to 0$, and using the continuity of f,

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

Also,

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

(where the limit follows by the continuity of f), so

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

Thus,

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

It follows that

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

Thus, to conclude the proof of the lemma it remains only to show that

$$\lim_{j\to\infty} \sup\left\{g(\nabla\phi(x)) : x \in B_{\varepsilon_j}\left(\left(\nabla\phi_{\varepsilon_j}\right)^{-1}(y)\right), \nabla\phi(x)\,\text{exists}\right\} = g(y). \tag{65}$$

By the same arguments as above, we have that

$$\begin{aligned}&\limsup_{j\to\infty} \sup\left\{g(\nabla\phi(x)) : x \in B_{\varepsilon_j}\left(\left(\nabla\phi_{\varepsilon_j}\right)^{-1}(y)\right), \nabla\phi(x)\,\text{exists}\right\}\\ &\le \sup\left\{g(\nabla\phi(x)) : x \in B_\gamma\left(\nabla\phi^*(y)\right)\cap\Omega, \nabla\phi(x)\,\text{exists}\right\}\end{aligned} \tag{66}$$

for any $\gamma > 0$.

We claim that ϕ is differentiable at $\nabla\phi^*(y)$. Indeed, $z \in \partial\phi(\nabla\phi^*(y))$ if and only if $\nabla\phi^*(y) \in \partial\phi^*(z)$ [33, Corollary 23.5.1]. Thus plugging in y for z, we see that $y \in \partial\phi(\nabla\phi^*(y))$. By assumption, $y \notin \partial\phi\left(\text{sing}(\phi)\right)$, so it must be that $\nabla\phi^*(y) \notin \text{sing}(\phi)$, as claimed, and, moreover,

$$\nabla\phi\left(\nabla\phi^*(y)\right) = y. \tag{67}$$

Next, for any $\alpha > 0$ there exists $\gamma > 0$ such that [33, Corollary 24.5.1]

$$\partial\phi(\nabla\phi^*(y) + v) \subset \nabla\phi(\nabla\phi^*(y)) + B_\alpha(0) = B_\alpha(y) \quad \forall v \in \overline{B}_\gamma(0). \tag{68}$$

Together with the continuity of g, this implies that the right-hand side of (66) converges to $g(y)$ as $\gamma \to 0$, so we have

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

Of course, we also have

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

where v_j is a vector with length less than ε_j chosen so that ϕ is differentiable at $\left(\nabla\phi_{\varepsilon_j}\right)^{-1}(y) + v_j$. Then, following (68), the continuity of g, (64), and (67), we have that the right-hand side tends to $g(y)$ as $j \to \infty$. Thus, (65) holds, and the proof of Lemma 4.18 is complete. ∎

4.8. Passing to the limit in ε—part II: Proof of Proposition 4.4. Define a measure supported on Λ,

$$\tilde{\nu} := (\nabla\phi)_\# \mu. \tag{69}$$

Claim 4.21. *For any sequence $\varepsilon \to 0$, $\nu_\varepsilon = (\nabla\phi_\varepsilon)_\# \mu|_{\Omega_\varepsilon}$ converges weakly to $\tilde{\nu}$.*

Proof. Let $\varepsilon \to 0$, and let ζ be a bounded continuous function on $\mathbb{R}^n$. Recalling the definition of ν_ε (38), the change-of-variables formula for the push-forward measure gives

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

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

$$\begin{aligned}\lim_{\varepsilon\to 0} \mu(\Omega_\varepsilon)^{-1} \int \zeta \, d\nu_\varepsilon &= \mu(\Omega)^{-1} \int_\Omega \zeta \circ \nabla\phi \, d\mu \\ &= \mu(\Omega)^{-1} \int_\Omega \zeta \, d\left((\nabla\phi)_\# \mu\right).\end{aligned}$$

This proves that $\mu(\Omega)\mu(\Omega_\varepsilon)^{-1}\nu_\varepsilon$ converges weakly to $\tilde{\nu} := (\nabla\phi)_\#\mu$. Since

$$\lim_{\varepsilon\to 0} \mu(\Omega_\varepsilon) = \mu(\Omega),$$

we are done. ∎

Observe that $\tilde{\nu}$ must be absolutely continuous because the densities g_ε of ν_ε are bounded above uniformly in ε (see (60)),

$$g_\varepsilon \le \sup f \sup g / \inf f < C, \tag{70}$$

and supported on the compact set $\overline{\Lambda}$. Hence, $\tilde{\nu}$ has a density that we denote by

$$\tilde{g} \, dx := \tilde{\nu}. \tag{71}$$

Proposition 4.4 follows from Claim 4.21 and the next two results.

LEMMA 4.22. *$\tilde{g} \le g$ almost everywhere.*

COROLLARY 4.23. *$\tilde{g} = g$ almost everywhere, i.e., $\tilde{\nu} = \nu$.*

Proof of Corollary 4.23. By Lemma 4.22 $g \ge \tilde{g}$ almost everywhere; thus,

$$\begin{aligned}\int_\Lambda |g - \tilde{g}| &= \int_\Lambda (g - \tilde{g}) \\ &= \int_\Lambda g - \int_\Lambda \tilde{g} = \nu(\Lambda) - \tilde{\nu}(\Lambda).\end{aligned} \tag{72}$$

As noted in the previous paragraph $\tilde{\nu}$ is absolutely continuous (with respect to the Lebesgue measure). Hence, as $\partial\Lambda$ is a Lebesgue null set, $\tilde{\nu}(\partial\Lambda) = 0$, i.e., Λ is a continuity set of $\tilde{\nu}$. Therefore, by Claim 4.21,

$$\tilde{\nu}(\Lambda) = \tilde{\nu}(\overline{\Lambda}) = \lim_{\varepsilon\to 0} \nu_\varepsilon(\overline{\Lambda}) = \lim_{\varepsilon\to 0} \mu|_{\Omega_\varepsilon}\left((\nabla\phi_\varepsilon)^{-1} (\overline{\Lambda})\right).$$

Now $(\nabla\phi_{\varepsilon_n})^{-1}(\overline{\Lambda}) \supset \Omega_{\varepsilon_n}$, so $\mu|_{\Omega_{\varepsilon_n}}\left((\nabla\phi_{\varepsilon_n})^{-1}(\overline{\Lambda})\right) = \mu(\Omega_{\varepsilon_n})$, and $\tilde{\nu}(\Lambda) = \mu(\Omega)$. Of course, by (1), $\mu(\Omega) = \nu(\Lambda)$, so by (72) we have that $\int_\Lambda |g - \tilde{g}| = 0$. This implies that $g = \tilde{g}$ almost everywhere, so $\nu = \tilde{\nu}$. ∎

Proof of Lemma 4.22. Define (recall (62))

$$S := \partial\phi\left(\mathrm{sing}(\phi)\right). \tag{73}$$

CLAIM 4.24. *$\Delta(\phi^*) \cap S$ is a $\tilde{\nu}$-null set.*

Proof. First we claim that on the set where ϕ^* is differentiable, S can be written as

$$\Delta(\phi^*) \cap S = \Delta(\phi^*) \cap (\nabla\phi^*)^{-1}\left(\mathrm{sing}(\phi)\right). \tag{74}$$

Indeed, suppose that $y \in \Delta(\phi^*) \cap (\nabla\phi^*)^{-1}(\mathrm{sing}(\phi))$. Then $\nabla\phi^*(y) \in \mathrm{sing}(\phi)$. Now by duality [33, Corollary 23.5.1], $y \in \partial\phi(\nabla\phi^*(y))$, and so $y \in S$.

Then suppose that $y \in \Delta(\phi^*) \cap S$. Then $y \in \partial\phi(x)$ for some $x \in \mathrm{sing}(\phi)$, implying, again by duality [33, Corollary 23.5.1], that $x \in \partial\phi^*(y)$, i.e., $x = \nabla\phi^*(y)$, and $y \in (\nabla\phi^*)^{-1}(\mathrm{sing}(\phi))$. This gives the claimed set equality.

Next, we claim that

$$\Delta(\phi^*) \cap S \text{ is Borel.} \tag{75}$$

For the proof, define an auxiliary vector-valued function $\Phi : \mathbb{R}^n \to \overline{\Omega}$ by

$$\Phi(y) := \begin{cases} \nabla\phi^*(y) & \text{on } \Delta(\phi^*), \\ z_0 & \text{on } \mathrm{sing}(\phi^*), \end{cases} \tag{76}$$

where $z_0 \in \Delta(\phi) \cap \Omega$ (any such (fixed) z_0 will do).

Claim 4.25. *Φ is a Borel-measurable function.*

Proof. Recall that the set of points of differentiability of a continuous function (e.g., ϕ and ϕ^*) is Borel (this is an elementary fact, though see, e.g., [41]), and hence $\mathrm{sing}(\phi)$, $\mathrm{sing}(\phi^*)$ are also Borel.

Let U be open. Recall that $\nabla\phi^*$ is continuous on $\Delta(\phi^*)$ [33, Corollary 24.5.1]. Thus, if $z_0 \notin U$, then $\Phi^{-1}(U)$ is open in $\Delta(\phi^*)$, i.e., $\Phi^{-1}(U) = O \cap \Delta(\phi^*)$ for some open O. If $z_0 \in U$, then $\Phi^{-1}(U)$ is the union of $\mathrm{sing}(\phi^*)$ with some set open in $\Delta(\phi^*)$. In either case, $\Phi^{-1}(U)$ is Borel, i.e., Φ is a Borel-measurable function. ∎

Therefore, $\Phi^{-1}(\mathrm{sing}(\phi))$ is Borel. Notice, since by construction $z_0 \notin \mathrm{sing}(\phi)$, that

$$\Phi^{-1}(\mathrm{sing}(\phi)) = (\nabla\phi^*)^{-1}(\mathrm{sing}(\phi)) \cap \Delta(\phi^*),$$

which together with (74) proves (75).

Thus, $\Delta(\phi^*) \cap S$ is $\tilde{\nu}$-measurable since $\tilde{\nu}$ is absolutely continuous. Compute

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

This integral vanishes since the integrand is only nonzero on the Lebesgue null set $\mathrm{sing}(\phi)$, while μ is absolutely continuous. The proof of Claim 4.24 is complete. ∎

Define a set $E \subset \mathrm{sing}(\phi^*)$ by

$$(\Delta(\phi^*) \cap S)^c = S^c \cup \mathrm{sing}(\phi^*) =: S^c \cup E \tag{77}$$

and by requiring that the union in the last expression be disjoint. 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. Claim 4.24 implies that

$$\tilde{\nu} = \tilde{\nu}|_{(\Delta(\phi^*)\cap S)^c} = \tilde{\nu}|_{S^c \cup \mathrm{sing}(\phi^*)} = \tilde{\nu}|_{S^c \cup E}. \tag{78}$$

By the absolute continuity of $\tilde{\nu}$ (recall (71)),

$$\tilde{g}(x) = \lim_{r\to 0} \frac{\tilde{\nu}(B_r(x))}{\mathrm{vol}(B_r(x))} \quad \text{for a.e. } x \in \Lambda. \tag{79}$$

Let $\alpha > 0$ and let U be an open set containing $S^c \cup E$ with

$$m\big(U \backslash (S^c \cup E)\big) < \alpha \tag{80}$$

(where m denotes the Lebesgue measure). This is possible because $S^c \cup E$ is Borel thanks to (75) and (77). Then, using (78),

$$\tilde{\nu}(B_r(x)) = \tilde{\nu}\big(B_r(x) \cap (S^c \cup E)\big) \leq \tilde{\nu}(B_r(x) \cap U). \tag{81}$$

Now since $B_r(x) \cap U$ is open, by Claim 4.21,

$$\tilde{\nu}(B_r(x) \cap U) \leq \liminf_{\varepsilon \to 0} \nu_\varepsilon(B_r(x) \cap U) \quad \forall r > 0. \tag{82}$$

Observe that

$$\begin{aligned} B_r(x) \cap U &= \big(B_r(x) \cap (S^c \cup E)\big) \cup \big(B_r(x) \cap (U \backslash (S^c \cup E))\big) \\ &\subset \big(B_r(x) \cap (S^c \cup E)\big) \cup \big(U \backslash (S^c \cup E)\big), \end{aligned}$$

so then by (39), (70), (80), and Claim 4.17,

$$\begin{aligned} \nu_\varepsilon(B_r(x) \cap U) &\leq \nu_\varepsilon\big(B_r(x) \cap (S^c \cup E)\big) + \nu_\varepsilon\big(U \backslash (S^c \cup E)\big) \\ &\leq \int_{B_r(x) \cap (S^c \cup E)} g_\varepsilon + Cm\big(U \backslash (S^c \cup E)\big) \\ &\leq \int_{B_r(x) \cap S^c} \overline{g}_\varepsilon + C\alpha, \end{aligned}$$

where in the last step we used the fact that E has Lebesgue measure zero. Now by Lemma 4.18, $\overline{g}_\varepsilon \to g$ almost everywhere on S^c, so by bounded convergence (since the $\overline{g}_\varepsilon$ are uniformly bounded by definition (60) by the same bound as in (70)) the last expression is convergent and

$$\liminf_{\varepsilon \to 0} \nu_\varepsilon(B_r(x) \cap U) \leq \int_{B_r(x) \cap S^c} g + C\alpha. \tag{83}$$

Then by (81), (82), and (83), 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.,

$$\tilde{\nu}(B_r(x)) \leq \int_{B_r(x)} g. \tag{84}$$

Then by (79), (84), and continuity of g,

$$\tilde{g}(x) \leq \liminf_{r \to 0} \frac{1}{\mathrm{vol}(B_r(x))} \int_{B_r(x)} g = g(x)$$

for a.e. x, concluding the proof of Lemma 4.22. ∎

4.9. Concluding the proof via the stability of optimal transport. We are at last in a position to complete the proof of the main theorem. As explained in section 4.3 it remains only to establish Lemma 4.5, whose proof hinges on two claims. The proof of these claims will follow the proof of the lemma.

Recall that $\nabla\varphi$ is the unique optimal transport map from μ to ν and $\varphi(0) = 0$.

Claim 4.26. *Fix $\delta > 0$. As ε tends to zero, $\nabla\phi_\varepsilon$ converges to $\nabla\varphi$ in probability with respect to $\mu|_{\Omega_\delta}/\mu(\Omega_\delta)$.*

Claim 4.27. *As ε tends to zero, $\nabla\phi_\varepsilon$ converges to $\nabla\phi$ almost everywhere on Ω.*

Proof of Lemma 4.5. Let $\beta > 0$. A consequence of Claim 4.26 is that there exists a sequence $\varepsilon_j \to 0$ such that $\nabla\phi_{\varepsilon_j} \to \nabla\varphi$ μ-almost everywhere on Ω_β, hence almost everywhere on Ω_β (because f is bounded away from zero). But, by Claim 4.27, $\nabla\phi_{\varepsilon_j} \to \nabla\phi$ almost everywhere on Ω_β. Hence, $\nabla\phi = \nabla\varphi$ on Ω_β. Since $\bigcup_{\beta>0}\Omega_\beta = \Omega$, we have that $T = \nabla\phi$ on Ω, i.e., $\nabla\varphi = \nabla\phi$ almost everywhere. Since $\varphi, \phi \in C^{0,1}(\overline{\Omega})$, both are absolutely continuous, and since $\phi(0) = 0 = \varphi(0)$, we have that $\phi = \varphi$ on $\overline{\Omega}$. ∎

Proof of Claim 4.26. The stability theorem for optimal transport maps states that whenever the push-forward of a given probability measure α under a sequence of optimal transport maps $\{T_j\}_{j\in\mathbb{N}}$ converges weakly to β, i.e.,

$$(T_j)_\#\alpha \to \beta \text{ weakly as } j \to \infty,$$

then T_j converges in probability to the unique optimal transport map pushing forward α to β, assuming such a unique map exists [40, Corollary 5.23]. We need a slight extension of this result where instead of a fixed measure α we have a sequence of measures α_j converging weakly to α, and

$$(T_j)_\#\alpha_j \to \beta \text{ weakly as } j \to \infty.$$

The result we need is stated in Proposition 4.29 below. Its proof is given in section 4.10.

Now, Proposition 4.29 may be applied to

$$\alpha_j := \mu(\Omega_{\varepsilon(j)})^{-1}\mu_{\varepsilon(j)}, \quad T_j := \nabla\phi_{\varepsilon(j)}, \quad \alpha := \mu(\Omega)^{-1}\mu, \quad \beta := \mu(\Omega)^{-1}\nu,$$

where $\{\epsilon(j)\}_{j\in\mathbb{N}}$ is any sequence of positive numbers converging to 0. Indeed, Brenier's theorem [8] gives that T_j is an optimal transport map pushing forward α_j to $\mu(\Omega_{\varepsilon(j)})^{-1}\nu_{\varepsilon(j)}$, and these latter measures converge weakly to β by Proposition 4.4, while α_j evidently weakly converges to α. Thus,

$$\lim_{\varepsilon\to 0}\mu|_{\Omega_\varepsilon}\Big(\big\{x \in \Omega : d\big(\nabla\phi_\varepsilon(x), \nabla\varphi(x)\big) \geq \gamma\big)\big\}\Big) \to 0 \quad \forall\gamma > 0. \tag{85}$$

Let $\delta, \gamma > 0$. Then for all ε sufficiently close to zero, $\Omega_\delta \subset \Omega_\varepsilon \subset \Omega$, so

$$\begin{aligned}\mu|_{\Omega_\delta}\big(\{x \in \Omega_\delta : d\left(\nabla\phi_\varepsilon(x), \nabla\varphi(x)\right) \geq \gamma\}\big) &= \mu\big(\Omega_\delta \cap \{x \in \Omega_\delta : d\left(\nabla\phi_\varepsilon(x), \nabla\varphi(x)\right) \geq \gamma\}\big)\\ &\leq \mu\big(\Omega_\varepsilon \cap \{x \in \Omega : d\left(\nabla\phi_\varepsilon(x), \nabla\varphi(x)\right) \geq \gamma\}\big)\\ &= \mu|_{\Omega_\varepsilon}\big(\{x \in \Omega : d\left(\nabla\phi_\varepsilon(x), \nabla\varphi(x)\right) \geq \gamma\}\big),\end{aligned}$$

and the last expression approaches zero as $\varepsilon \to 0$ by (85), concluding the proof of Claim 4.26. ∎

Proof of Claim 4.27. Because ϕ is convex and continuous, $\phi_\varepsilon \to \phi$ pointwise. Semicontinuity of the subdifferential map [33, Theorem 24.5] gives that for any $x \in \Omega$ and $\alpha > 0$,

$$\partial\phi_\varepsilon(x) \subset \partial\phi(x) + B_\alpha(0) \quad \forall\varepsilon \text{ sufficiently small}$$

(this is a low-brow version of Theorem 4.7). Thus at a point x such that $\partial\phi$ is a singleton,

$$|\nabla\phi(x) - \nabla\phi_\varepsilon(x)| < \alpha$$

for all ε sufficiently small, i.e., $\nabla\phi_\varepsilon(x) \to \nabla\phi(x)$. Since $\partial\phi$ is a singleton almost everywhere, the claim follows. ◻

4.10. A stability result. Let $\Pi(\alpha, \beta)$ denote the set of probability measures on $X \times Y$ whose marginals are α on X and β on Y, i.e., for every $\mu \in \Pi(\alpha, \beta)$, $(\pi_1)_{\#}\mu = \alpha, (\pi_2)_{\#}\mu = \beta$, where $\pi_1 : X \times Y \to X, \pi_2 : X \times Y \to Y$ are the natural projections. Elements of $\Pi(\alpha, \beta)$ are called transference plans. Given a function $c : X \times Y \to \mathbb{R}$, define the cost associated to $\mu \in \Pi(\alpha, \beta)$ by

$$\int_{X\times Y} c\, d\mu.$$

A transference plan is called optimal if it realizes the infimum of the cost over $\Pi(\alpha, \beta)$. Optimal transference plans satisfy the following standard stability result [40, Theorem 5.20].

THEOREM 4.28. *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 α_j and β_j be sequences of probability measures on X and Y, respectively, such that α_j converges weakly to α and β_j converges weakly to β. For each j, let π_j be an optimal transference plan between α_j and β_j. Assume that*

$$\int c\, d\pi_j < \infty\ \forall j, \quad \liminf_j \int c\, d\pi_j < \infty.$$

Then there exists a subsequence $\{j_l\}_{l\in\mathbb{N}}$ such that π_{j_l} converges weakly to an optimal transference plan.

The following result, and its proof, are a slight modification of [40, Corollary 5.23].

PROPOSITION 4.29. *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 α_j and β_j be sequences of probability measures on X and Y, respectively, such that $\alpha_j \le C\alpha$ for all j, and α_j converges weakly to α and β_j converges weakly to β. For each j, let π_j be an optimal transference plan between α_j and β_j. Assume that*

$$\int c\, d\pi_j < \infty\ \forall j, \quad \liminf_j \int c\, d\pi_j < \infty.$$

Suppose that there exist measurable maps $T_j, T : X \to Y$ such that $\pi_j = (\mathrm{id} \otimes T_j)_{\#}\, \alpha_j$ and $\pi = (\mathrm{id} \otimes T)_{\#}\, \alpha$. Assume additionally that π is the unique optimal transference plan in $\Pi(\alpha, \beta)$. Then

$$\lim_{j\to\infty} \alpha_j \left[\{x \in X \,:\, |T_j(x) - T(x)| > \varepsilon\}\right] = 0 \quad \forall\varepsilon > 0.$$

Proof. First, note that by Theorem 4.28 and the uniqueness of π, we have that $\pi_j \to \pi$ weakly (and there is no need to take a subsequence). Now, let $\varepsilon > 0$ and $\delta > 0$. By Lusin's theorem, there exists a compact set $K \subset X$ with $\alpha(X\backslash K) < C^{-1}\delta$ (so $\alpha_j(X\backslash K) < \delta$) 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 $\pi = (\mathrm{id} \otimes T)_{\#}\, \alpha$, 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,

$$\begin{aligned} 0 = \pi(A_\varepsilon) &\geq \limsup_{j\to\infty} \pi_j(A_\varepsilon) \\ &= \limsup_{j\to\infty} \pi_j\left(\{(x,y) \in K \times Y \,:\, |T(x) - y| \geq \varepsilon\}\right) \\ &= \limsup_{j\to\infty} \alpha_j\left(\{x \in K \,:\, |T(x) - T_j(x)| \geq \varepsilon\}\right) \\ &\geq \limsup_{j\to\infty} \alpha_j\left(\{x \in X \,:\, |T(x) - T_j(x)| \geq \varepsilon\}\right) - \alpha_j(X\backslash K) \\ &\geq \limsup_{j\to\infty} \alpha_j\left(\{x \in X \,:\, |T(x) - T_j(x)| \geq \varepsilon\}\right) - \delta, \end{aligned}$$

and the desired result follows by letting δ tend to zero. ∎

5. Weakening the regularity assumption. As promised in Remark 3.2, we now set out to prove a stronger version of Proposition 3.1, i.e., a version that does not require regularity of the Brenier potential φ up to the boundary. In the following we only assume $\varphi \in C^2(\Omega)$. As discussed in Remark 3.2, this regularity assumption is satisfied automatically under the hypotheses of Theorem 1.9, and consequently Corollary 5.5 below establishes that Theorem 1.9 follows from Theorem 1.6.

Recall (9), which we restate here for convenience:

(86)
$$F_i\Big(\big\{\psi_j^{(k)}, \eta_j^{(k)}\big\}_{j=1}^{N(k)}\Big) := \max\left\{0, -(\det H_i)^{1/n} + \Big(f\Big(\tfrac{\sum_{j=0}^n x_{i_j}}{n+1}\Big)\big/g\Big(\tfrac{\sum_{j=0}^n \eta_{i_j}}{n+1}\Big)\Big)^{1/n}\right\}.$$

Furthermore, recall (10), also restated here:

$$F\Big(\big\{\psi_j^{(k)}, \eta_j^{(k)}\big\}_{j=1}^{N(k)}\Big) := \sum_{i=1}^{M(k)} V_i \cdot F_i\Big(\big\{\psi_j^{(k)}, \eta_j^{(k)}\big\}_{j=1}^{N(k)}\Big), \tag{87}$$

so again F_i is a per-simplex penalty, and F is the objective function of the DMAOP.

To prove the strengthened version of Proposition 3.1, it will not suffice as it did before to simply plug the discrete data (18) associated to the Brenier potential φ into the DMAOP and hope that the corresponding cost goes to zero as $k \to \infty$. The reason is that we no longer have that the Hessian of φ is bounded away from zero on $\overline{\Omega}$. Instead, we will define functions that are strongly convex on $\mathbb{R}^n$ and that agree with φ on subsets that exhaust $\overline{\Omega}$. Since the functions that we construct may not be differentiable, we will also need to mollify slightly before plugging the associated data into the DMAOP.

Let U be open and compactly contained in Ω. In turn, let V be open such that $U \subset\subset V$ and $V \subset\subset \Omega$. Then there exists $\alpha > 0$ such that $\nabla^2\varphi(x) \geq \alpha I$ for all $x \in \overline{V}$, and $\delta := \mathrm{dist}(\overline{U}, V^c)$ is strictly positive.

Also, $\operatorname{dist}(\nabla\varphi(x), \Lambda^c)$ is continuous in x over the compact set $\overline{U}$ and (since $\nabla\varphi(\overline{U}) \subset \Lambda$) strictly positive. Hence the function attains a positive minimum γ over $\overline{U}$, i.e., $\operatorname{dist}(\nabla\varphi(x), \Lambda^c) \geq \gamma > 0$ for all $x \in \overline{U}$.

Let $R = \operatorname{diam}(\Omega)$, and define

$$\beta := \min\left\{\alpha, \frac{\alpha\delta^2}{R^2}, \frac{\gamma}{2R}\right\}. \tag{88}$$

For every point $y \in \overline{U}$, define a quadratic polynomial Q_y on all of $\mathbb{R}^n$ by

$$Q_y(x) := \varphi(y) + \langle\nabla\varphi(y), x-y\rangle + \frac{1}{2}\beta|x-y|^2. \tag{89}$$

LEMMA 5.1. *For any fixed $y \in \overline{U}$, $Q_y \leq \varphi$ on $\overline{U}$ and $\nabla Q_y(\Omega_+) \subset \Lambda$ for some open set Ω_+ such that $\Omega \subset\subset \Omega_+$.*

Proof. To simplify the proof of this claim, we fix $y \in \overline{U}$ and consider

$$\psi(x) := \varphi(x) - \varphi(y) - \langle\nabla\varphi(y), x-y\rangle. \tag{90}$$

For the first statement it suffices to show that

$$\psi(x) \geq \frac{1}{2}\beta|x-y|^2 \quad \forall x \in \overline{U}. \tag{91}$$

Note that ψ is convex on $\mathbb{R}^n$ (since the Brenier potential may be taken to be defined on $\mathbb{R}^n$) with $\psi(0) = 0$, $\nabla\psi(0) = 0$, and $\nabla^2\psi(x) \geq \alpha I$ for all $x \in \overline{V}$. Evidently $\psi \geq 0$ everywhere. By integration along rays, for $x \in \overline{B_\delta(y)} \subset \overline{V}$,

$$\psi(x) \geq \frac{1}{2}\alpha|x-y|^2. \tag{92}$$

In particular, since $\alpha \geq \beta$, inequality (91) follows but only for $x \in \overline{B_\delta(y)}$.

Now let $x \in \overline{U} \setminus \overline{B_\delta(y)}$. In order to deal with the possible nonconvexity of the domain U, let $z \in \partial B_\delta(y)$ (so $|x - z| = \delta$) such that x, y, z are collinear. From (90) we have that $\psi(z) \geq \frac{1}{2}\alpha\delta^2$. From this fact, together with (88), we obtain

$$\frac{1}{2}\beta|x-y|^2 \leq \frac{1}{2}\beta R^2 \leq \frac{1}{2}\alpha\delta^2 \leq \psi(z).$$

Now by our choice of z, we can write $z = tx + (1-t)y$ for some $t \in [0, 1]$. Then by convexity (and the nonnegativity of ψ),

$$\psi(z) \leq t\psi(x) + (1-t)\psi(y) = t\psi(x) \leq \psi(x).$$

Thus we have established that $\frac{1}{2}\beta|x-y|^2 \leq \psi(x)$, and the first statement of the lemma is proved.

Take Ω_+ to be an open set with $\operatorname{diam}(\Omega_+) \leq \frac{3}{2}R$ such that $\Omega \subset\subset \Omega_+$. It remains to show that $\nabla Q_y(\Omega_+) \subset \Lambda$. Let $x \in \Omega_+$, and note that

$$\nabla Q_y(x) = \nabla\varphi(y) + \beta(x-y),$$

so (recalling (88))

$$|\nabla Q_y(x) - \nabla\varphi(y)| \leq \frac{3}{2}\beta R \leq \frac{3}{4}\gamma.$$

But since $\operatorname{dist}(\nabla\varphi(x), \Lambda^c) \geq \gamma$, it follows that $\nabla Q_y(x) \in \Lambda$. This completes the proof. ∎

In summary, we have shown that for any U open and compactly contained in Ω, there exists $\beta > 0$ such that the quadratic polynomial Q_y as defined in (89) satisfies $Q_y \leq \varphi$ on $\overline{U}$ and $\nabla Q_y(\Omega_+) \subset \Lambda$ for all y, for some open Ω_+ such that $\Omega \subset\subset \Omega_+$. (Note that β does not depend on y.) Then define φ_U on all of $\mathbb{R}^n$ via

$$\varphi_U(x) := \sup_{y\in\overline{U}} Q_y(x). \tag{93}$$

Evidently $\varphi_U = \varphi$ on $\overline{U}$ and $\nabla\varphi_U(\Omega_+) \subset \Lambda$ (see, e.g., [32, Proposition 2.7]). Since the pointwise supremum of β-strongly convex functions is β-strongly convex, φ_U is β-strongly convex.

Although φ_U is not necessarily differentiable, we can substitute φ_U with a smooth approximation via the following lemma.

LEMMA 5.2. *Let $\varepsilon > 0$, and consider an open set $U \subset\subset \Omega$. There exists a smooth convex function $\tilde{\varphi} : \mathbb{R}^n \to \mathbb{R}$ with $\nabla^2\tilde{\varphi} \geq \beta I$, $\nabla\tilde{\varphi}(\overline{\Omega}) \subset \Lambda$, and $\|\tilde{\varphi} - \varphi\|_{C^2(\overline{U})} < \varepsilon$.*

Proof. Let V be open such that $U \subset\subset V \subset\subset \Omega$. By the preceding arguments, we can take φ_V to be β-strongly convex and agreeing with φ on V such that $\nabla\varphi_V(\Omega_+) \subset \Lambda$. Let ξ_δ denote, as before, a standard mollifier supported on $B_\delta(0)$. For δ sufficiently small, a δ-neighborhood of U is contained in V, so in fact, for small δ, we have $\varphi_V \star \xi_\delta = \varphi \star \xi_\delta$ on $\overline{U}$. It follows that $\|\varphi_V \star \xi_\delta - \varphi\|_{C^2(\overline{U})} < \varepsilon$ for small enough δ. Furthermore, for δ sufficiently small, a δ-neighborhood of $\overline{\Omega}$ is contained in Ω_+, so by the convexity of Λ, we have that $\nabla(\varphi_V \star \xi_\delta)(\overline{\Omega}) \subset \Lambda$. Noticing that mollification preserves β-strong convexity, the proof is completed by taking $\tilde{\varphi} = \varphi_V \star \xi_\delta$ for some δ small enough. ∎

Remark 5.3. One can avoid using the convexity of Λ in the proof of the preceding lemma by a more complicated argument. However, since we have assumed this fact elsewhere in this article, we make use of it here to keep the proof as simple as possible.

As before, let Ω_ε be as in Definition 1.5, and define $U_\varepsilon := \Omega_\varepsilon + B_{\varepsilon/2}(0) \subset \Omega$. By the preceding lemma, we can let φ_ε be smooth and convex such that $\nabla^2\varphi_\varepsilon \geq \beta I$, $\nabla\varphi_\varepsilon(\overline{\Omega}) \subset \Lambda$, and $\|\varphi_\varepsilon - \varphi\|_{C^2(\overline{U_\varepsilon})} < \varepsilon$. (Note that here φ_ε is *not* the same as $\varphi \star \xi_\varepsilon$.)

By analogy with (17), we consider the cost

$$d_\varepsilon^{(k)} := F\Big(\big\{\varphi_\varepsilon(x_j^{(k)}), \nabla\varphi_\varepsilon(x_j^{(k)})\big\}_{j=1}^{N(k)}\Big) \tag{94}$$

associated to the data

$$\big\{\varphi_\varepsilon(x_j^{(k)}), \nabla\varphi_\varepsilon(x_j^{(k)})\big\}_{j=1}^{N(k)} \in (\mathbb{R}\times\mathbb{R}^n)^{N(k)} \tag{95}$$

extracted from our modified Brenier potential φ_ε.

We now state and prove our improvement of Proposition 3.1.

PROPOSITION 5.4. *Let $\big\{\{S_i^{(k)}\}_{i=1}^{M(k)}\big\}_{k\in\mathbb{N}}$ be a sequence of admissible and regular almost-triangulations of Ω (recall Definitions* 1.5 *and* 1.8*). Let φ be the unique Brenier solution of the Monge–Ampère equation* (2) *with $\varphi(0) = 0$, and suppose that $\varphi \in C^2(\Omega)$. Then*

(i) *the data* (95) *satisfies the constraints* (6)–(8) *for all k sufficiently large;*

(ii) $\limsup_k d_\varepsilon^{(k)} = o(\varepsilon)$*, where $d_\varepsilon^{(k)}$ is defined as in* (94).

Let c_k be the optimal cost of the kth DMAOP. Let $\varepsilon > 0$. If the data (95) associated with φ_ε is feasible (which is true by part (i) of Proposition 5.4 for k sufficiently large), then $c_k \leq d_\varepsilon^{(k)}$. Thus by part (ii) of Proposition 5.4, $\limsup_k c_k = o(\varepsilon)$. This yields the following analogue of Corollary 3.3.

COROLLARY 5.5. *Under the assumptions of Proposition* 5.4*,* $\lim_k c_k = 0$.

Proof of Proposition 5.4. First let

$$I_k = \left\{ i = 1, \ldots, M(k) \, : \, S_i^{(k)} \subset \overline{U_\varepsilon} \right\}, \tag{96}$$

and let

$$J_k = \{1, \ldots, M(k)\} \setminus I_k. \tag{97}$$

Also, recall that given a matrix $A = [a_{ij}]$, we define

$$||A|| = \max_{i,j} |a_{ij}|.$$

Since φ_ε is smooth (and so in particular in $C^{2,\alpha}(\overline{\Omega})$) and strongly convex on $\mathbb{R}^n$, we have by the same reasoning as in the proof of Lemma 3.4 that (recalling (4))

$$\lim_k \max_{i \in \{1,\ldots,M(k)\}} \left\| H\left(S_i^{(k)}, \left\{\nabla\varphi_\varepsilon(x_{i_0}^{(k)}), \ldots, \nabla\varphi_\varepsilon(x_{i_n}^{(k)})\right\}\right) - \nabla^2\varphi_\varepsilon(x_{i_0}^{(k)}) \right\| = 0.$$

From this it follows (as in the proof of Proposition 3.1) that the data (95) satisfies the constraint (8). That the data satisfies the constraints (6) and (7) is evident from the construction of φ_ε.

Furthermore, since $\|\varphi_\varepsilon - \varphi\|_{C^2(\overline{U_\varepsilon})} < \varepsilon$ by construction, we have that

$$\limsup_k \max_{i \in I_k} \left\| H\left(S_i^{(k)}, \left\{\nabla\varphi_\varepsilon(x_{i_0}^{(k)}), \ldots, \nabla\varphi_\varepsilon(x_{i_n}^{(k)})\right\}\right) - \nabla^2\varphi(x_{i_0}^{(k)}) \right\| \leq \varepsilon.$$

From this it follows (as in the proof of Proposition 3.1) that

$$\limsup_k \max_{i \in I_k} F_i \left(\left\{ \varphi_\varepsilon(x_j^{(k)}), \nabla\varphi_\varepsilon(x_j^{(k)}) \right\}_{j=1}^{N(k)} \right) = o(\varepsilon).$$

Henceforth we will abbreviate $F_i := F_i\big(\{\varphi_\varepsilon(x_j^{(k)}), \nabla\varphi_\varepsilon(x_j^{(k)})\}_{j=1}^{N(k)}\big)$. Then the preceding implies that

$$\limsup_k \sum_{i \in I_k} V_i \cdot F_i = o(\varepsilon).$$

Thus to establish that $\limsup_k d_\varepsilon^{(k)} = o(\varepsilon)$, it will suffice to show that

$$\limsup_k \sum_{i \in J_k} V_i \cdot F_i = o(\varepsilon).$$

However, since f is bounded above, g is bounded away from zero, and $(\cdot)^{1/n}$ is bounded below by zero over the nonnegative numbers, it follows from (86) that $F_i \leq C$ for some constant C (depending only on f, g). Thus

$$\limsup_k \sum_{i \in J_k} V_i \cdot F_i \leq C \limsup_k \sum_{i \in J_k} V_i = C \limsup_k \operatorname{vol}\left(\bigcup_{i \in J_k} S_i \right).$$

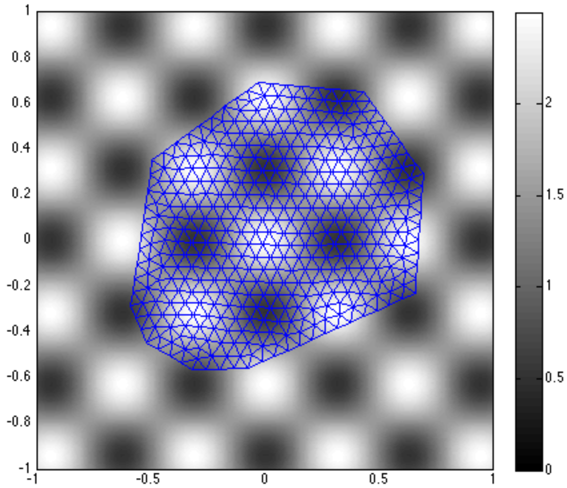

FIG. 1. *Source measure supported on a convex polygon* Ω *(shown triangulated in the above). The shading in the background represents the density of* f*, though we understand that* $f \equiv 0$ *outside of* Ω*. There are* 405 *points in this triangulation.*

Now suppose $i \in J_k$, and assume k is large enough such that the maximal simplex diameter is at most $\varepsilon/4$. Then S_i contains a point x that is not in $\overline{U_\varepsilon}$. Recall that $U_\varepsilon = \Omega_\varepsilon + B_{\varepsilon/2}(0)$, so it follows that $\mathrm{dist}(x, \Omega_\varepsilon) \geq \varepsilon/2$. Since $\mathrm{diam}(S_i) \leq \varepsilon/4$, we see that $S_i \subset \overline{\Omega} \setminus \Omega_\varepsilon$. Thus

$$\mathrm{vol}\left(\bigcup_{i \in J_k} S_i\right) \leq \mathrm{vol}\left(\overline{\Omega} \setminus \Omega_\varepsilon\right) = o(\varepsilon),$$

and this completes the proof. ∎

6. Numerical experiments for the DMAOP.

6.1. Implementation details. Only two details of the implementation bear mentioning. First, we used DistMesh for the triangulation of Ω [30]. Second, we solved each convex optimization problem using MOSEK [25], called via the modeling language YALMIP [22].

6.2. 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 nonuniform 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.

For our first example, we consider a source measure (see Figure 1) supported on a convex polygon Ω with an oscillatory density f bounded away from zero.

For the triangulation (consisting of 405 vertices) pictured in Figure 1, the DMAOP took 49.3 seconds to solve on a 2011 MacBook Pro with a 2.2 GHz Intel Core i7 processor.

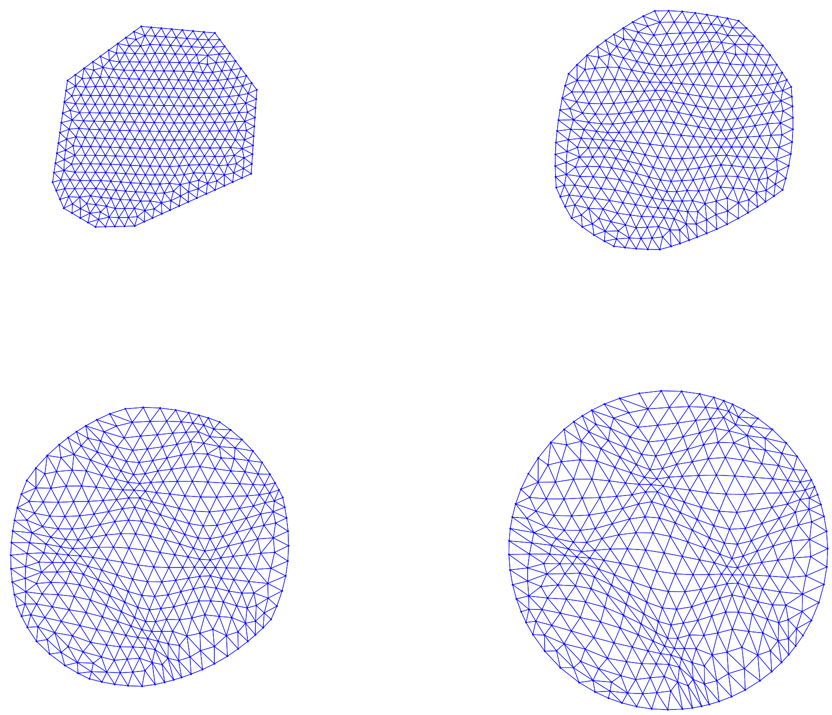

FIG. 2. *Visualization of numerical solution of the dynamical optimal transport problem associated with the example of Figures* 1 *and* 4. *Times* $t = 0, \frac{1}{3}, \frac{2}{3}, 1$ *are depicted at upper left, upper right, lower left, and lower right, respectively. (The target measure is uniform on the unit disc.)*

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. See Figure 2.

Next we consider a source measure with uniform density on a nonconvex support. See Figure 3 for a visualization of the domain, its triangulation, and the numerical solution to the dynamical optimal transport problem. Our triangulation uses 340 points, and solving the DMAOP took 51.2 seconds. Note that a detailed theoretical study of a similar example is given in [11]. In Figure 4 we visualize the computed convex potential.

Lastly we consider an example in which the source measure has highly irregular support (again with uniform density on its support). See Figure 5 for details. There are 359 points in our 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 succeeds at highlighting the discontinuity sets of these inverse maps.

6.3. Run time and convergence analysis. We now fix an example problem and analyze the performance of our algorithm on discretizations of varying coarseness

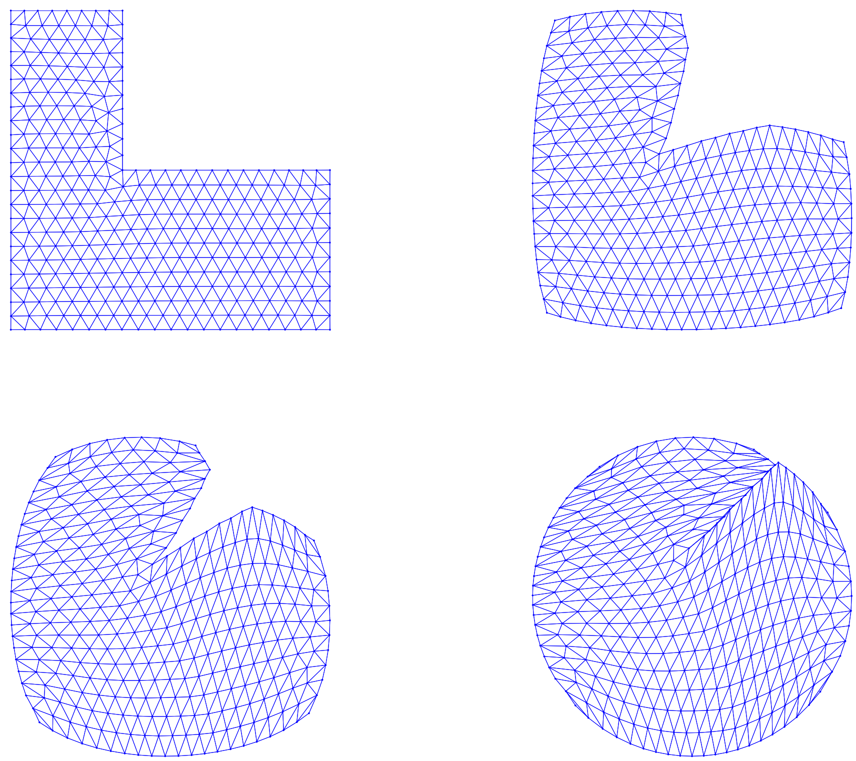

FIG. 3. *Visualization of numerical solution of the dynamical optimal transport problem with source measure taken to be uniform and supported on the domain in the upper left and target measure taken to be uniform on the unit disc. Times* $t = 0, \frac{1}{3}, \frac{2}{3}, 1$ *are depicted at upper left, upper right, lower left, and lower right, respectively.*

to gain some idea of the performance of our method, as well as its fundamental weaknesses. Specifically, we analyze the problem in the second example considered above (depicted in Figure 3).

See Figure 6 for the dependence of run time on problem size. (All numerical computations were performed on a 2011 MacBook Pro with a 2.2 GHz Intel Core i7 processor.) The asymptotic behavior of the total run time (which includes mesh generation as well as setting up the convex problem in the modeling language YALMIP) is, empirically, no worse than quadratic. The time spent by the convex solver (MOSEK) on the actual optimization problem (arguably a more fundamental quantity) is also empirically quadratic. It is reasonable that this would be the case, since the number of constraints of the DMAOP grows quadratically in the number of discretization points.

Next we examine the dependence of the cost (as in Definition 1.3) of our numerical solution on problem size. Figure 7 indicates that the cost decays as $N^{-\frac{1}{2}}$ (where N is the number of discretization points). Since we are in dimension two, we expect that the mesh scale h decays as $N^{-\frac{1}{2}}$, so in fact the cost decays like h. Note that the proof of Lemma 3.4 guarantees that the optimal cost is $O(h^{\alpha})$ whenever the Brenier potential satisfies $\varphi \in C^{2,\alpha}(\overline{\Omega})$, and generally one may take $\alpha = 1$ if f and g are

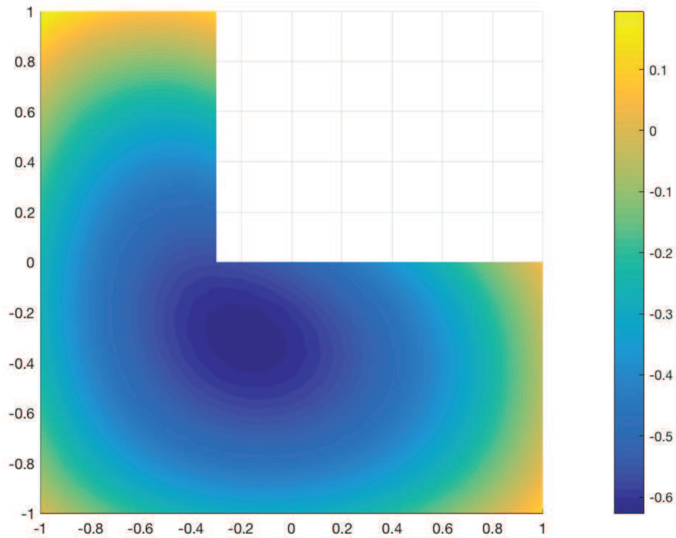

FIG. 4. *Visualization of the convex potential retrieved by solving the DMAOP associated to the source measure and triangulation in Figure 3, with shading corresponding to the value of the potential. (The target measure is the uniform on the unit disc.)*

sufficiently regular.

We also study the decay of a two-sided cost (not explicitly optimized in the DMAOP) that penalizes both excessive contraction and excessive expansion. With a view toward Definition 1.3, consider the quantity

$$\tilde{c} := \sum_{i=1}^{M} V_i \cdot \left| -(\det H_i)^{1/n} + \left(f\left(\frac{\sum_{j=0}^{n} x_{i_j}}{n+1}\right) / g\left(\frac{\sum_{j=0}^{n} \eta_{i_j}}{n+1}\right)\right)^{1/n} \right|,$$

where the η_j are the η_j of our solution of the DMAOP, and the H_i are defined with respect to the η_j as in (4). $\tilde{c}$ can be thought of as the average over the simplices of a two-sided penalty on area distortion. The dependence of $\tilde{c}$ on N is depicted in Figure 7 and does not differ qualitatively from the dependence of the DMAOP cost on N.

6.4. Discussion. We include here some remarks on features and drawbacks of our numerical method.

First, we comment that the method can be used to compute discontinuous optimal maps by inverting optimal maps from nonconvex to convex domains, and the discontinuity sets can be resolved sharply (see section 6.2 for examples). These examples are in practice no more computationally expensive than convex-to-convex examples.

Also, although we have not taken advantage of this feature in the examples of section 6.2, we remark that the method suggests an adaptive approach in which resolution is added where it is wanted or needed. Indeed, we may simply solve the DMAOP for a triangulation with a greater density of vertices in desired areas.

A significant limitation of our implementation is the requirement that the target measure have density g for which $g^{-1/n}$ is convex. While the DMAOP still admits

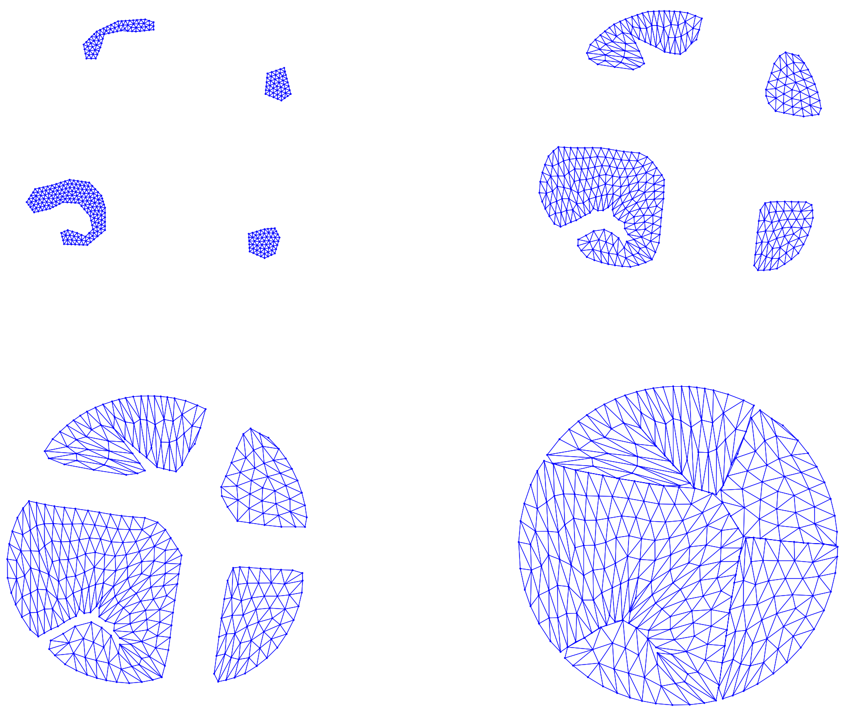

FIG. 5. *Visualization of numerical solution of the dynamical optimal transport problem with source measure taken to be uniform and supported on the domain in the upper left and target measure taken to be uniform on the unit disc. Times* $t = 0, \frac{1}{3}, \frac{2}{3}, 1$ *are depicted at upper left, upper right, lower left, and lower right, respectively.*

a minimizer for general target measures, it is only clear a priori that the DMAOP can be practically solved when it is convex. We propose a remedy for this issue in section 8.

Another limitation is that the method is only first-order accurate. This is confirmed empirically in section 6.3, but it is also to be expected due to the use of first-order finite difference quotients in the definition (4) of the H_i in terms of the subgradients η_j. One might hope to replace these with higher-order difference quotients, and indeed we will do this in section 7.

Lastly, the quadratic growth of run time in N is a major drawback. This growth is due to to the quadratic increase in the number of constraints of the DMAOP. The modification introduced in section 7 will also address this issue.

7. Improving the efficiency and accuracy of the DMAOP. As remarked above, the quadratic growth of the problem size of the DMAOP in the number of discretization points all but disqualifies it from competitive use in application. The quadratic growth of the problem size derives from our global enforcement of the convexity of the potential. One might hope that this issue could be remedied by instead relying on a constraint that enforces convexity while relying only on local (second-derivative) information. Moreover, one might hope that the first-order accuracy of the scheme could be improved.

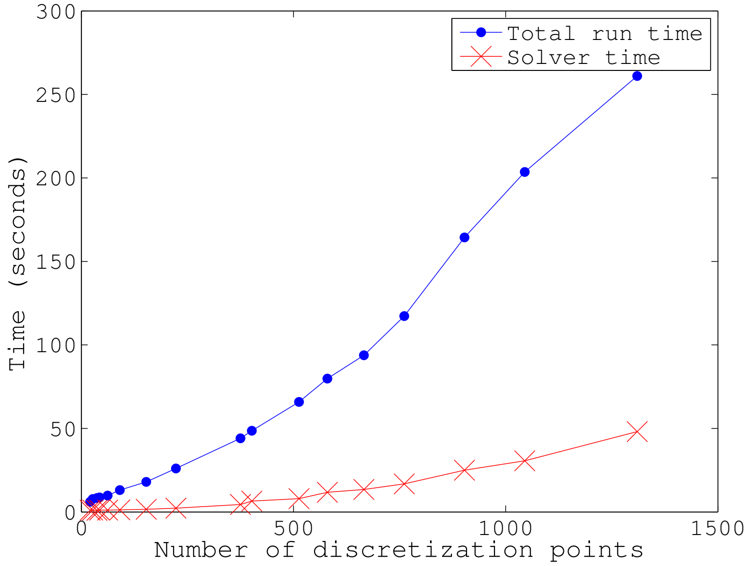

FIG. 6. *Plot of total run time of the algorithm and run time of the convex solver against the number of discretization points.*

In this section, we introduce a modification of the DMAOP that addresses both of these concerns. The basic idea of the modification is to consider only the values of the convex potential as optimization variables. Instead of considering subgradients as separate variables, as we did in the DMAOP, and enforcing convexity *globally* via subgradient inequalities (which previously resulted in the imposition of a quadratically growing number of constraints), we define gradients directly in terms of the values of the potential via finite difference quotients and enforce convexity *locally* by enforcing the positive-semidefiniteness of finite difference Hessians of the potential. Moreover, we do all of this on a rectangular grid using standard second-order accurate finite difference quotients.

We will restrict our attention to the case of $\overline{\Omega} = [-1,1]^n =: K$, and we assume that Λ is convex and g is constant, equal to $1/\mathrm{vol}(\Lambda)$. (We shall see in the next section that general g can be tackled by solving several problems with constant target density.) For numerical experiments we shall consider $n = 2$, $\overline{\Lambda} = K$. We do not prove the convergence of this method, though we make some remarks on this matter below.

For $h > 0$ (the mesh size), let $\mathbb{Z}_h = \{hk \,:\, k \in \mathbb{Z}\}$ and $\mathbb{Z}_h^n = (\mathbb{Z}_h)^n$. For any set $\mathcal{O}$ define $\mathcal{O}_h := \mathcal{O} \cap \mathbb{Z}_h$ and $\partial\mathcal{O}_h := \partial\mathcal{O} \cap \mathbb{Z}_h$. We only consider h such that $1/h$ is an integer. We will treat discretizations with respect to the grid $K_h := [0,1]_h^n$, i.e., solutions of our discrete problems will be functions $u : K_h \to \mathbb{R}$. We will refer to such functions, i.e., functions $K_h \to \mathbb{R}$, as *grid functions*. The space of grid functions will be denoted $\mathbb{R}^{K_h}$.

At every point $x \in K_h$, we define finite difference operators $D_{h,x} = (D_{h,x,i})_{i=1}^n : \mathbb{R}^{K_h} \to \mathbb{R}^n$ and $D^2_{h,x} = (D^2_{h,x,i,j})_{i,j=1}^n : \mathbb{R}^{K_h} \to \mathbb{R}^{n\times n}$, which are second-order accurate finite difference approximations of the gradient and Hessian, respectively. At interior points $x \in K_h \setminus \partial K_h$, we take these to be given, respectively, by the usual centered

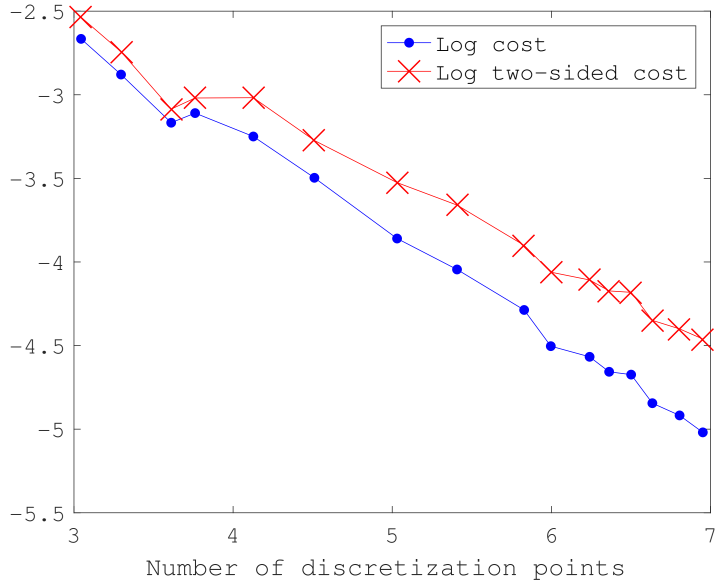

FIG. 7. *Log-log plot of the DMAOP cost of the numerical solution and a two-sided cost ($\tilde{c}$, introduced in section* 6.3*) against the number of discretization points.*

finite difference approximations of first derivatives and the usual 3^n-point stencil for the Hessian. More precisely, for an interior point x and $u \in \mathbb{R}^{K_h}$, we take

$$D_{h,x,i}u = \frac{1}{2h}\left[u(x+he_i) - u(x-he_i)\right],$$

$$D^2_{h,x,i,i}u = \frac{1}{h^2}\left[u(x+he_i) - 2u(x) + u(x-he_i)\right],$$

and (for $i \neq j$)

$$D^2_{h,x,i,j}u = \frac{1}{4h^2}\sum_{\sigma_1,\sigma_2\in\{-1,1\}} \sigma_1\sigma_2\, u(x + h\sigma_1 e_i + h\sigma_2 e_j).$$

For points $x \in \partial K_h$, we replace the relevant components of these finite difference operators with appropriate second-order accurate forward or backward finite differences as needed.

Then we solve the following problem.

DEFINITION 7.1. *Let $h > 0$ and $f \in \mathbb{R}^{K_h}$. The revised discrete Monge–Ampère optimization problem (RDMAOP) associated to the data (f, h) is*

$$\begin{aligned} &\underset{\psi\in\mathbb{R}^{K_h}}{\text{minimize}} \sum_{x\in K_h} \max\left\{0, -\left(\det D^2_{h,x}\psi\right)^{1/n} + \left(\mathrm{vol}(\Lambda) f(x)\right)^{1/n}\right\} \\ &\textit{subject to } D_{h,x}\psi \in K,\ x \in K_h, \\ &\qquad\qquad D^2_{h,x}\psi \geq 0,\ x \in K_h. \end{aligned}$$

Compared to Definition 1.3 (DMAOP), we see that instead of including the gradients (in Definition 1.3, the η_j) as optimization variables and then defining finite difference Hessians in terms of these variables, the RDMAOP instead defines both gradients and Hessians directly in terms of the grid function potential ψ, the only optimization variable. Moreover, the convexity of ψ is enforced locally. We remark that convexity could alternatively be enforced via more sophisticated methods, in particular via the wide stencil finite differences of Oberman [26], but this does not seem to be needed in the examples we consider (since the forward optimal transport maps are C^1).

7.1. Theoretical remarks on convergence. Although (as we will confirm below) this method is more efficient and more accurate than the DMAOP, the DMAOP lends itself more naturally to a convergence proof because it allows for the simple construction of optimization potentials that are bona fide convex functions on $\mathbb{R}^n$. By contrast, it is not clear how to extract a convex function defined on $\mathbb{R}^n$ from a grid function with positive semidefinite discrete Hessians (or even from a grid function that is convex in the wide stencil sense of [26]). Thus the major missing piece in a convergence proof for the RDMAOP is a way of extracting convex potentials from the grid functions retrieved by the optimization. We conjecture that this can be done, but our working proof involves new ideas that take us too far afield from the analysis in this paper and which are applicable more widely in the numerical analysis for nonlinear elliptic PDEs. For these reasons, a full theoretical investigation of the convergence of the RDMAOP will be the subject of future work.

7.2. Numerical experiments. We remark that the RDMAOP is a second-order cone program. We call MOSEK directly to solve it numerically. The numerical results of this section were performed on a 2015 MacBook Pro with a 2.5 GHz Intel Core i7 processor.

First we consider a nontrivial example (see [5]) for which an explicit solution is available. To maintain consistency with [5], we will consider $\Omega = \Lambda = (-0.5, 0.5)^2$, though of course our discussion can be transferred by scaling to the domain $(-1,1)^2$ in order to fit into the framework described above.

Define

$$q(z) = \left(-\frac{1}{8\pi}z^2 + \frac{1}{256\pi^3} + \frac{1}{32\pi}\right)\cos(8\pi z) + \frac{1}{32\pi^2} z\sin(8\pi z),$$

and in turn define the source density

$$\begin{aligned} f(x_1,x_2) = 1 &+ 4(q''(x_1)q(x_2) + q(x_1)q''(x_2)) \\ &+ 16(q(x_1)q(x_2)q''(x_1)q''(x_2) - q'(x_1)^2 q'(x_2)^2) \end{aligned} \tag{98}$$

on $(-0.5, 0.5)^2$. The corresponding OT problem (with target density uniform on $(-0.5,0.5)^2$) admits the explicit solution

$$\varphi_{x_1}(x_1,x_2) = x_1 + 4q'(x_1)q(x_2), \quad \varphi_{x_2}(x_1,x_2) = x_2 + 4q(x_1)q'(x_2).$$

The source density and the image of the grid K_h under the computed transport map with $h = 1/64$ are shown in Figure 8.

Detailed numerical results (for the rescaled problem) are recorded in Table 1. We observe error on the order of h^2, consistent with the second-order accuracy of the scheme, as well as linear growth in computational time with respect to the number of discretization points. (Note that the number of discretization points is on the order of h^{-2} since we are working in dimension 2.) We record separately the time taken by an

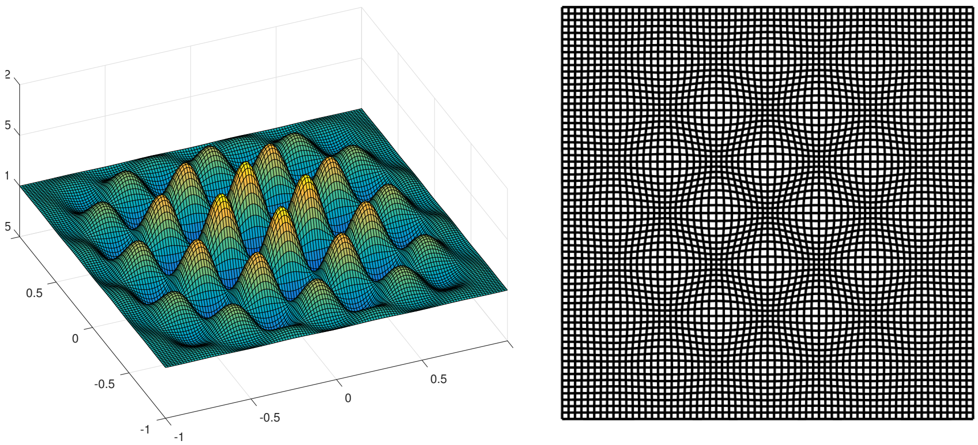

FIG. 8. *The source density defined by* (98)*, after rescaling to* $(-1,1)^2$*, and the image of* K_h *under the computed transport map* ($h = 1/64$)*.*

TABLE 1
RDMAOP results for the OT problem defined by the source density of (98)*. Errors refer to error of the transport (i.e., gradient) map.* N_x *denotes the number of discretization points along a single dimension. Thus the number of optimization variables is* N_x^2*. "On. time" and "Off. time" refer to the times taken by the online and offline parts of the computation, respectively.*

h	N_x	L^∞ error	L^2 error	On. time (s)	Off. time (s)
2^{-3}	17	0.0324	0.0124	1.53	0.72
2^{-4}	33	0.0022	0.0010	1.38	0.15
2^{-5}	65	0.0010	0.0005	2.15	0.11
2^{-6}	129	0.0003	0.0002	8.20	1.06
2^{-7}	257	0.0002	0.0001	38.58	3.18

offline computational step, general to all RDMAOPs on a given grid (i.e., independent of source density), in which finite difference operators are constructed for the grid.

The reader can compare this example directly to [5]. (We remark that scaling has been considered so that direct comparison of performance via Table 1 is indeed appropriate.) The run times for given problem sizes are quite similar, and the error of our method is smaller, likely due to the higher order of accuracy.

We next consider the OT problem with source density

$$f(x) = e^{-\frac{1}{2(1.4)^2}|x|^2}(3 + \sin(8\pi x_1)\sin(6\pi x_2)), \tag{99}$$

normalized to define a probability measure. The source density and the image of the grid K_h under the computed transport map with $h = 1/64$ are shown in Figure 9.

We record numerical results in Table 2. We do not have an exact solution against which to compare our numerical solution, so we measure the error of our numerical solution for given $h = 2^{-k}$ by comparing it against the numerical solution for $h = 2^{-(k+1)}$, subsampled on the grid of size 2^{-k}. Again we observe error of order h^2 and linear growth of computational time with respect to the number of discretization points.

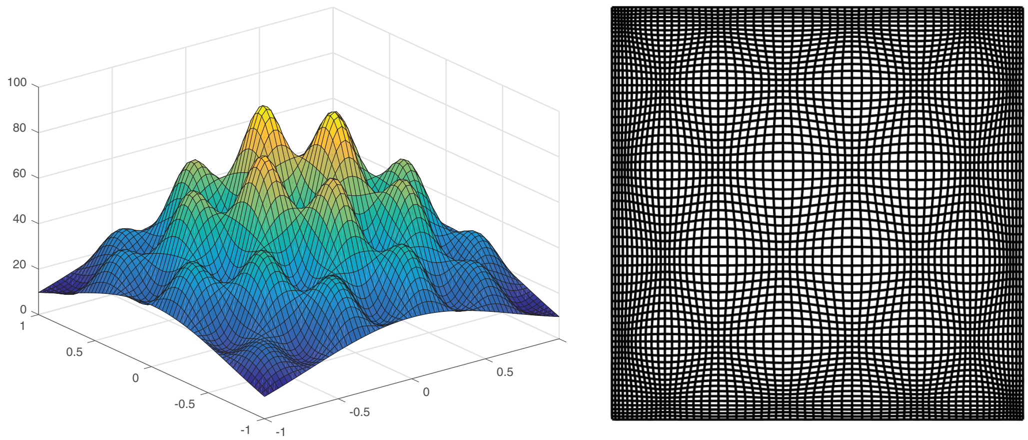

Fig. 9. *The source density defined by* (99) *and the image of* K_h *under the computed transport map* ($h = 1/64$)*.*

Table 2

RDMAOP results for the OT problem defined by the source density of (99)*. Errors refer to error of the transport (i.e., gradient) map. Error for a given* $h = 2^{-k}$ *is measured by comparing against the computed solution for* $h = 2^{-(k+1)}$*.* N_x *denotes the number of discretization points along a single dimension. Thus the number of optimization variables is* N_x^2*.*

h	N_x	L^∞ error	L^2 error	On. time (s)	Off. time (s)
2^{-3}	17	0.0375	0.0179	1.27	0.58
2^{-4}	33	0.0119	0.0049	1.51	0.11
2^{-5}	65	0.0034	0.0012	1.90	0.10
2^{-6}	129	0.0013	0.0007	9.51	0.40
2^{-7}	257	–	–	42.31	3.13

8. Removing the restriction on the target measure. We propose a method for solving a given OT problem with general target density by solving a sequence of OT problems with constant target densities. This approach amounts to a fixed point iteration and is inspired by the Ricci iteration in differential geometry [35, 36], where a similar approach for solving complex Monge–Ampère equations is introduced. We make no attempt at rigor in the following argument.

Suppose that Λ is convex as above, and let the target measure be given by $g\,dx$ on its support Λ. Let $V = \mathrm{vol}(\Lambda)$, so the constant density $1/V$ defines the uniform probability measure on Λ.

Recall that we want to solve (in some sense)

$$\begin{aligned} \det\left(\nabla^2\varphi(x)\right) &= \frac{f(x)}{g\left(\nabla\varphi(x)\right)}, \quad x \in \Omega, \\ \nabla\varphi(\Omega) &= \Lambda. \end{aligned} \tag{100}$$

Note that the solution φ of this problem simultaneously solves an OT problem with source density

$$\frac{f/(g\circ\nabla\varphi)}{\int_\Omega f/(g\circ\nabla\varphi)\,dx} = \frac{f/(g\circ\nabla\varphi)}{\int_\Omega \det\left(\nabla^2\varphi(x)\right)\,dx} = \frac{f/(g\circ\nabla\varphi)}{V} \tag{101}$$

and constant target density $1/V$.

Keeping this fact in mind, we turn to defining our iterative procedure. Let $f^{(0)} = f$, and let $\varphi^{(0)}$ be the solution of

$$\begin{aligned}\det\left(\nabla^2\varphi(x)\right) &= Vf^{(0)}(x), \quad x\in\Omega,\\ \nabla\varphi(\Omega) &= \Lambda,\end{aligned}$$

i.e., the original OT problem with the target measure replaced by the uniform measure on Λ. We then define $\tilde f^{(1)}(x) := f(x)/g(\nabla\varphi^{(0)}(x))$, set $f^{(1)} := \tilde f^{(1)}/\big(\int_\Omega \tilde f^{(1)}\,dx\big)$, and solve another OT problem with source density $f^{(1)}$ and uniform target measure. This iteration can be repeated many times.

More precisely, we define $\varphi^{(i+1)}$ to be the solution of

$$\begin{aligned}\det\left(\nabla^2\varphi(x)\right) &= Vf^{(i+1)}(x) \quad x\in\Omega,\\ \nabla\varphi(\Omega) &= \Lambda,\end{aligned}$$

where $f^{(i+1)} := \tilde f^{(i+1)}/\big(\int_\Omega \tilde f^{(i+1)}\,dx\big)$ for $\tilde f^{(i+1)} := f/(g\circ\nabla\varphi^{(i)})$.

Note from our argument above (in particular, (101)) that the solution of (100) is a fixed point of this iteration.

Suppose that $\varphi^{(i)}$ converges to some $\tilde\varphi$ as $i\to\infty$ in some suitably strong sense. Then we expect that $\tilde\varphi$ is a fixed point and hence solves (100). We do not prove such convergence (and remark that it is not at all obvious to us how to do so), but we observe below that rapid convergence occurs in practice, at least with densities bounded away from zero.

8.1. Numerical implementation of the iteration. We consider the same setting as in section 7 but now allow g to be nonconstant. We fix some mesh size $h > 0$ for the entirety of the iterative procedure. We inductively define $\psi^{(i)} \in \mathbb{R}^{K_h}$ to be a grid function solving the RDMAOP associated to the data $(f^{(i)}, h)$ (see Definition 7.1), where $f^{(i)} := \tilde f^{(i)}/I\big[\tilde f^{(i)}\big]$ for $\tilde f^{(i)}(x) := f(x)/g(D_{h,x}\psi^{(i-1)})$ and for I an operator on grid functions that approximates the integral over K. In particular, for u a grid function, we take $I[u]$ to return the true integral of the function that is constant on each grid cube, with its value on a given grid cube equal to the average of u over the vertices of the cube. We terminate iteration when the L^2 distance between two consecutive solutions first drops below $h^2/2$.

8.2. Numerical experiments. Consider once again the source density of (98) and (unnormalized) target density

$$g(x) := 3 + \sin(3\pi x_1)\sin(2\pi x_2). \tag{102}$$

The source density and the image of the grid K_h under the computed transport map with $h = 1/64$ are shown in Figure 10.

Detailed numerical results are recorded in Table 3. Once again we do not have an exact solution against which to compare our numerical solution, so we measure the error for $h = 2^{-k}$ by comparing against the result for $h = 2^{-(k+1)}$. Once again we observe error of order h^2 and linear growth of computational time with respect to

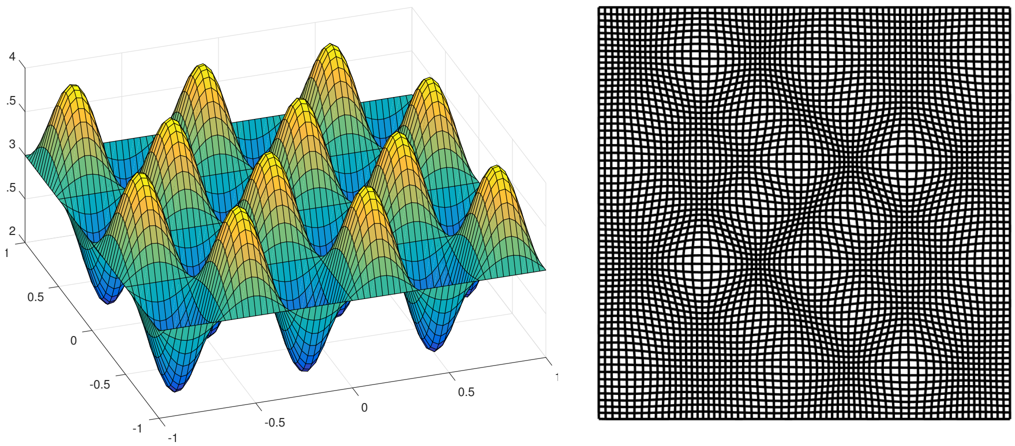

FIG. 10. *The target density defined by* (102) *and the image of K_h under the computed transport map ($h = 1/64$), with source density as in Figure* 8.

TABLE 3
RDMAOP results for the OT problem defined by the source density of (98) *and the target density of* (102). *Errors refer to error of the transport (i.e., gradient) map. Error for a given $h = 2^{-k}$ is measured by comparing against the computed solution for $h = 2^{-(k+1)}$.*

h	L^∞ error	L^2 error	Iterations	On. time (s)	Off. time (s)
2^{-3}	0.0814	0.0317	3	3.36	0.59
2^{-4}	0.0175	0.0095	4	5.59	0.09
2^{-5}	0.0056	0.0031	4	8.97	0.09
2^{-6}	0.0017	0.0008	5	45.16	0.40
2^{-7}	–	–	5	207.02	3.08

the number of discretization points. The number of iterations needed to reach a fixed point does not appear to depend strongly on h.

Next we consider an example in which the iterative method does not perform as well as one might hope. Define (unnormalized) source and target densities by

$$f = 1.5 - \mathbf{1}_{B(0,0.5)} \tag{103}$$

and

$$g(x) = 6 + 25 \sum_{x_1', x_2' \in \{-1,1\}} e^{-\frac{1}{2(0.2)^2}|x-(x_1',x_2')|^2}. \tag{104}$$

The target density and the image of the grid K_h under the computed transport map with $h = 1/64$ are shown in Figure 11.

Results are recorded in Table 4. This time we observe observe error of order h, though still linear growth of computational time. This outcome seems to be typical for the iterative procedure when the target density is close to zero on a large portion of the target domain. Indeed, the iterative procedure completely fails to converge numerically in many situations with highly degenerate target densities. It is not clear

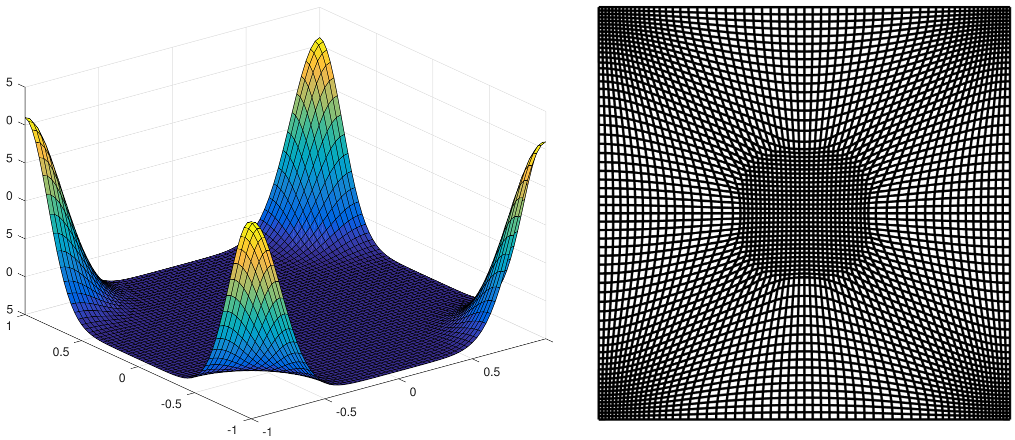

FIG. 11. *The target density defined by* (104) *and the image of K_h under the computed transport map ($h = 1/64$), with source density as in Figure* 103.

TABLE 4
RDMAOP results for the OT problem defined by the source density of (103) *and the target density of* (104). *Errors refer to error of the transport (i.e., gradient) map. Error for a given $h = 2^{-k}$ is measured by comparing against the computed solution for $h = 2^{-(k+1)}$.*

h	L^∞ error	L^2 error	Iterations	On. time (s)	Off. time (s)
2^{-3}	0.0312	0.0192	6	7.42	0.63
2^{-4}	0.0156	0.0069	7	10.53	0.10
2^{-5}	0.0091	0.0041	7	14.11	0.12
2^{-6}	0.0045	0.0025	9	78.90	0.38
2^{-7}	–	–	10	381.73	3.11

to the authors whether this shortcoming is inherent in the method or whether it can be overcome with more stable fixed point iteration techniques.

9. Future directions.

First, it is an interesting open question whether our convergence proof can be upgraded to yield error bounds.

Second, we plan to follow up on the remarks of section 7.1 regarding the essential difficulties of proving convergence for the RDMAOP.

Third, the theoretical and numerical convergence properties of the fixed point iteration of section 7 remain open, as does the possibility of a better numerical approach to the fixed point iteration.

Acknowledgments. The authors are grateful to O. Chodosh, V. Jain, and L. Panchev for many stimulating discussions and initial collaboration on this project and on [11]. We also thank the anonymous referees for many helpful comments.

REFERENCES

[1] S. ANGENENT, S. HAKER, AND A. TANNENBAUM, *Minimizing flows for the Monge–Kantorovich problem*, SIAM J. Math. Anal., 35 (2003), pp. 61–97, https://doi.org/10.1137/S0036141002410927.

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

[3] A. Baldes and 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. 17, Springer, New York, 1991, pp. 19–30.

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

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

[6] J.-D. Benamou, B. D. Froese, and A. M. Oberman, *A Viscosity Solution Approach to the Monge–Ampère Formulation of the Optimal Transportation Problem*, preprint, https://arxiv.org/abs/1208.4873, 2012.

[7] S. Boyd and L. Vandenberghe, *Convex Optimization*, Cambridge University Press, Cambridge, uk, 2004.

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

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

[10] O. Chodosh, V. Jain, M. Lindsey, L. Panchev, and Y. A. Rubinstein, *Visualizing optimal transportation maps*, abstracts of papers presented to the American Mathematical Society, Vol. 34, No. 1086-28-496, 2013.

[11] O. Chodosh, V. Jain, M. Lindsey, L. Panchev, and Y. A. Rubinstein, *On discontinuity of planar optimal transport maps*, J. Topology Anal., 7 (2015), pp. 239–260.

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

[13] M. Cuturi, *Sinkhorn distances: Lightspeed computation of optimal transportation distances*, Neural Information Processing Systems Conference, 2013.

[14] X.-B. Feng, R. Glowinski, and M. Neilan, *Recent developments in numerical methods for fully nonlinear second order partial differential equations*, SIAM Rev., 55 (2013), pp. 205–267, https://doi.org/10.1137/110825960.

[15] B. D. Froese and A. M. Oberman, *Convergent filtered schemes for the Monge–Ampère partial differential equation*, SIAM J. Numer. Anal., 51 (2013), pp. 423–444, https://doi.org/10.1137/120875065.

[16] W. Gangbo and R. McCann, *The geometry of optimal transportation*, Acta Math., 177 (1996), pp. 113–161.

[17] K. Guittet, *On the time-continuous mass transport problem and its approximation by augmented Lagrangian techniques*, SIAM J. Numer. Anal., 41 (2003), pp. 382–399, https://doi.org/10.1137/S0036142901386069.

[18] E. Haber, T. Rehman, and A. Tannenbaum, *An efficient numerical method for the solution of the L_2 optimal mass transfer problem*, SIAM J. Sci. Comput., 32 (2010), pp. 197–211, https://doi.org/10.1137/080730238.

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

[20] B. Lévy, *A numerical algorithm for L_2 semi-discrete optimal transport in 3D*, ESAIM Math. Model. Numer. Anal., 49 (2015), pp. 1693–1715.

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

[22] J. Löfberg, *YALMIP: A toolbox for modeling and optimization in MATLAB*, in Proceedings of the 2004 IEEE International Symposium on Computer Aided Control Systems Design, IEEE, 2004, https://doi.org/10.1109/CACSD.2004.1393890.

[23] Q. Mérigot and É. Oudet, *Discrete optimal transport: Complexity, geometry and applications*, Discrete Comput. Geom., 55 (2016), pp. 263–283.

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

[25] MOSEK ApS, *The MOSEK Optimization Toolbox for MATLAB Manual*, Version 7.1 (Revision 28), http://docs.mosek.com/7.1/toolbox/index.html, 2015.

[26] A. Oberman, *Wide stencil finite difference schemes for the elliptic Monge-Ampère equation and functions of the eigenvalues of the hessian*, Discrete Contin. Dyn. Syst. Ser. B, 10 (2008), pp. 221–238.

[27] A. Oberman and Y. Ruan, *An Efficient Linear Programming Method for Optimal Transportation*, preprint, https://arxiv.org/abs/1509.03668, 2015.

[28] V. I. Oliker and 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), pp. 271–293.
[29] N. Papadakis, G. Peyré, and E. Oudet, *Optimal transport with proximal splitting*, SIAM J. Imaging Sci., 7 (2014), pp. 212–238, https://doi.org/10.1137/130920058.
[30] P.-O. Persson and G. Strang, *A simple mesh generator in MATLAB*, SIAM Rev., 46 (2004), pp. 329–345, https://doi.org/10.1137/S0036144503429121.
[31] A. V. Pogorelov, *The Minkowski Multidimensional Problem*, Wiley, New York, 1978.
[32] J. Rauch and B. A. Taylor, *The Dirichlet problem for the multidimensional Monge–Ampère equation*, Rocky Mountain J. Math., 7 (1977), pp. 345–364.
[33] R. T. Rockafellar, *Convex Analysis*, Princeton University Press, Princeton, NJ, 1970.
[34] R. T. Rockafellar and R. J.-B. Wets, *Variational Analysis*, Springer, New York, 1997.
[35] Y. A. Rubinstein, *The Ricci iteration and its applications*, C. R. Math. Acad. Sci. Paris, 345 (2007), pp. 445–448.
[36] Y. A. Rubinstein, *Some discretizations of geometric evolution equations and the Ricci iteration on the space of Kähler metrics*, Adv. Math., 218 (2008), pp. 1526–1565.
[37] B. Schmitzer, *A sparse multiscale algorithm for dense optimal transport*, J. Math. Imaging Vision, 56 (2016), pp. 238–259.
[38] M. M. Sulman, J. F. Williams, and R. D. Russell, *An efficient approach for the numerical solution of the Monge–Ampère equation*, Appl. Numer. Math., 61 (2011), pp. 298–307.
[39] C. Villani, *Topics in Optimal Transportation*, AMS, Providence, RI, 2003.
[40] C. Villani, *Optimal Transport: Old and New*, Springer, Berlin, 2009.
[41] Z. Zahorski, *Sur l'ensemble des points de non-dérivabilité d'une fonction continue*, Bull. Soc. Math. France, 74 (1946), pp. 147–178.