Research Article
Special Issue: In honor of David Jerison

Alessio Figalli* and Yash Jhaveri

Regularity properties of monotone measure-preserving maps

https://doi.org/10.1515/ans-2022-0057
received October 31, 2022; accepted March 3, 2023

Abstract: In this note, we extend the regularity theory for monotone measure-preserving maps, also known as optimal transports for the quadratic cost optimal transport problem, to the case when the support of the target measure is an arbitrary convex domain and, on the low-regularity end, between domains carrying certain invariant measures.

Keywords: optimal transport maps, Monge-Ampère equation, doubling measures, unbounded support

MSC 2020: 35J96, 49Q22, 52A41

1 Introduction

Given a pair of atom-free Borel probability measures μ and ν on $\mathbb{R}$, the monotone rearrangement theorem asserts that the function $y = y(x)$ defined implicitly by

$$\int_{-\infty}^{x} d\mu = \int_{-\infty}^{y(x)} d\nu$$

is measure preserving, i.e.,

$$\mu(y^{-1}(\mathcal{E})) = \nu(\mathcal{E}) \quad \text{for any Borel set } \mathcal{E} \subset \mathbb{R}.$$

In particular, y is unique μ almost everywhere and can be made to be monotone, or, equivalently, the derivative of a convex function.

When $\mathbb{R}$ is replaced by $\mathbb{R}^n$, however, before Brenier's discovery [1], a proper generalization of the monotone rearrangement theorem was myth, and only after the work of McCann, in [12], was the myth made real. Precisely, he proved that if μ vanishes on every Lipschitz $(n-1)$-dimensional surface,[1] then a convex potential $u : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ exists whose gradient map $\nabla u = \nabla u(x)$ is unique μ almost everywhere and pushes μ forward to ν, i.e.,

In honor of David Jerison for his 70th birthday.

1 McCann actually assumes that μ vanishes on all (Borel) sets of Hausdorff dimension $n-1$. This guarantees that the set of nondifferentiability points of a convex function are μ-negligible. However, this assumption can be weakened. Since convex functions are differentiable outside of a countable union of Lipschitz hypersurfaces [14], McCann's theorem holds assuming that μ vanishes on every Lipschitz $(n-1)$-dimensional surface.

*** Corresponding author: Alessio Figalli,** Department of Mathematics, ETH Zürich, Rämistrasse 101, Zürich 8092, Switzerland, e-mail: alessio.figalli@math.ethz.ch
Yash Jhaveri: Department of Mathematics & Computer Science, Rutgers University – Newark, Smith Hall, 101 Warren Street, Newark, New Jersey 07102, USA, e-mail: yash.jhaveri@rutgers.edu

$$\mu((\nabla u)^{-1}(\mathcal{E})) = \nu(\mathcal{E}) \quad \text{for any Borel set } \mathcal{E} \subset \mathbb{R}^n.$$

(Brenier's theorem guaranteed the same conclusion as McCann's theorem, but under some restrictive technical conditions on μ and ν.)

The first general regularity result on these Brenier-McCann maps was proved by Caffarelli in [4]: provided that μ and ν are absolutely continuous with respect to n-dimensional Lebesgue measure $\mathscr{L}^n$, their respective densities f and g vanish outside of and are bounded away from zero and infinity on open bounded sets $\mathcal{X}$ and $\mathcal{Y}$, respectively, and $\mathcal{Y}$ is convex, he showed that u is strictly convex in $\mathcal{X}$ (see, e.g., the proof of [7, Theorem 4.6.2]). This result opened the door to the development of a regularity theory for mappings with convex potentials based on the regularity theory for strictly convex solutions to the Monge-Ampère equation [2]; indeed,

$$(\nabla u)_{\#} f = g \text{ is formally, at least, equivalent to } \det \mathcal{D}^2 u = \frac{f}{g(\nabla u)}.$$

Unfortunately, Caffarelli's boundedness assumptions on the domains $\mathcal{X}$ and $\mathcal{Y}$ are restrictive, since many probability densities, especially those found in applications, are supported on all of $\mathbb{R}^n$: Gaussian densities, for example. Motivated by this, in [5], Cordero-Erausquin and Figalli showed that, in several situations of interest, one can ensure the regularity of monotone measure-preserving maps even if the measures under consideration have unbounded supports. However, missing from their collection is the situation where $\mathcal{Y}$ is an *arbitrary* convex domain. Lifting this restriction is a main goal of this article.

1.1 Results

Our main theorem is an extension of Caffarelli's theorem, on the strict convexity of u, in two ways. First, we allow $\mathcal{X}$ and $\mathcal{Y}$ to be unbounded. Second, we permit $\mathcal{X}$ and $\mathcal{Y}$ to carrying certain invariant measures that we call *locally doubling measures* (qualitatively, our notion replaces balls in the classical notion of a doubling measure with ellipsoids, in order to account for the affine invariance of our setting).

Definition 1.1. A nonnegative measure λ is *locally doubling* (*on ellipsoids*) if the following holds: for every ball $\mathcal{B}$, there is a constant $C \geq 1$ such that

$$\lambda(\mathcal{E}) \leq C\lambda\left(\frac{1}{2}\mathcal{E}\right)$$

for all ellipsoids $\mathcal{E} \subset \mathcal{B}$ with center (of mass) in $\mathrm{spt}(\lambda)$. Here, $\frac{1}{2}\mathcal{E}$ is the dilation of $\mathcal{E}$ with respect to its center by 1/2.

This notion of doubling was introduced by Jhaveri and Savin in [11].[2] That said, the first consideration of measures with a "doubling-like" property in the world of solutions to Monge-Ampère equations can be traced back to the work of Jerison [9] and then Caffarelli [3]. In particular, in [3], Caffarelli showed that Alexandrov solutions to

$$\det \mathcal{D}^2 v = \rho,$$

where the measure ρ is doubling on a specific collection of convex sets called sections[3], share the same geometric properties as Alexandrov solutions to Monge-Ampère equations with right-hand sides comparable to Lebesgue measure [2].

We now state our main theorem.

---

2 This family of measures is strictly larger than the family of measures locally comparable to Lebesgue measure on their supports. (See [11] for examples of locally doubling measures not comparable to Lebesgue measure on their supports.)

3 These are sets of the form $\{v \leq \ell\}$ for any affine function ℓ.

Theorem 1.2. *Let μ and ν be two locally doubling probability measures on $\mathbb{R}^n$ that vanish on Lipschitz $(n-1)$-dimensional surfaces and are concentrated on two open sets $\mathcal{X}$ and $\mathcal{Y}$, respectively, and suppose that $\mathcal{Y}$ is convex. Then any convex potential u associated to the Brenier-McCann map pushing μ forward to ν is strictly convex in $\mathcal{X}$.*

Remark 1.3. It is well known that $\mathcal{Y}$ needs to be convex. When $\mathcal{Y}$ is not convex, u can fail to be strictly convex and ∇u can behave rather poorly. If we consider Pogorelov's counterexample to the strict convexity of solutions to the Monge-Ampère equation in three dimensions, we see that $\mathcal{Y}$ needs to be convex in order to guarantee the strict convexity of potentials of Brenier-McCann maps. In particular, let $u(x', x_3) = |x'|^{4/3}(1 + x_3^2)$, let $Q_r = \{|x_i| < r/2 \text{ for } i = 1, 2, 3\}$ be the cube with side length $r > 0$ and centered at the origin in $\mathbb{R}^3$, and let $\mathcal{Y} = \nabla u(Q_r)$. If $r \ll 1$, then $f = \det D^2 u$ is analytic and positive in Q_r, and ∇u is the Brenier-McCann map pushing forward $\mu = (f/\|f\|_{L^1(Q_r)})\mathcal{L}^3\llcorner Q_r$ to $\nu = (1/\mathcal{L}^3(\mathcal{Y}))\mathcal{L}^3\llcorner \mathcal{Y}$. The set $\mathcal{Y}$ is open but not convex, and $u = 0$ along $\{x' = 0\}$. Moreover, as demonstrated in, for instance, [4,10], ∇u easily fails to be continuous when $\mathcal{Y}$ is not convex.

Remark 1.4. While the target measure ν need not vanishes on all Lipschitz $(n-1)$-dimensional surfaces in order to invoke McCann's theorem (which asks this only of the source measure μ), it must in order to ensure that our main theorem holds. If $\mu = \mathcal{L}^2\llcorner Q_1$ is the uniform measure on Q_1 the unit cube centered at the origin in $\mathbb{R}^2$ and $\nu = \mathcal{H}^1\llcorner Q_1 \cap \{x_2 = 0\}$ is the 1-dimensional Hausdorff measure restricted to the central horizontal axis of Q_1, then the Brenier-McCann map pushing μ forward to ν is the projection map $(x_1, x_2) \mapsto x_1$. Up to a constant, this map's convex potential is $\frac{1}{2}x_1^2$, which is not strictly convex. With respect our proof of Theorem 1.2, asking this of both μ and ν guarantees the validity of the mass balance formula in Lemma 2.1 (see also Remark 2.2), our main tool.

Remark 1.5. A simple case to which Theorem 1.2 applies, but the corresponding results in [4, 5] do not, is when $\mu = g\mathcal{L}^3\llcorner\{|x_1| < 1\} \times \mathbb{R}^2$ and $\nu = g\mathcal{L}^3\llcorner\mathbb{R}^2 \times \{|x_3| < 1\}$, and g is the standard Gaussian density on $\mathbb{R}^3$ appropriately normalized to make μ and ν probability measures.

With our main theorem in hand, our second and third theorems further extend the known regularity theory for monotone measure-preserving maps, completing the story started by Cordero-Erausquin and Figalli in [5] on monotone transports between unbounded domains.

Theorem 1.6. *Let μ and ν be two locally doubling probability measures on $\mathbb{R}^n$ that vanish on Lipschitz $(n-1)$-dimensional surfaces and are concentrated on two open sets $\mathcal{X}$ and $\mathcal{Y}$, respectively, and suppose that $\mathcal{Y}$ is convex. Then the Brenier-McCann map ∇u pushing μ forward to ν is a homeomorphism from $\mathcal{X}$ onto a full measure subset of $\mathcal{Y}$. Moreover, for every $\mathcal{A} \Subset \mathcal{X}$, a constant $\alpha > 0$ exists such that $\nabla u \in C^{0,\alpha}(\mathcal{A})$. Furthermore, $\nabla u(\mathcal{X}) = \mathcal{Y}$ whenever $\mathcal{X}$ is convex.*

Theorem 1.7. *Let f and g be two functions on $\mathbb{R}^n$ that define locally doubling probability measures concentrated on two open sets $\mathcal{X}$ and $\mathcal{Y}$, respectively, and suppose that $\mathcal{Y}$ is convex. Assume that f and g are bounded away from zero and infinity on compact subsets of $\mathcal{X}$ and $\mathcal{Y}$, respectively. Then for every $\mathcal{E} \Subset \mathcal{X}$, a constant $\varepsilon > 0$ exists such that any convex potential u associated to the Brenier-McCann map pushing f forward to g is $W^{2,1+\varepsilon}(\mathcal{E})$. Also, ∇u is locally a $C^{k+1,\beta}$-diffeomorphism from $\mathcal{X}$ onto its image provided f and g are locally $C^{k,\beta}$ in $\mathcal{X}$ and $\mathcal{Y}$, respectively.*

Remark 1.8. We note that the proof of the Theorem 1.7, given the strict convexity of u (provided by Theorem 1.2), is classical. Indeed, it suffices to localize classical regularity results for the Monge-Ampère equation. We refer the reader to [7, Section 4.6.1] for more details.

1.2 Structure

This remainder of this article is structured as follows.

In Section 2, we prove Theorem 1.2. Our proof is self-contained apart from some facts in convex analysis; we provide explicit references to these used but unproved facts. We remark that our proof is inspired by the proof of the Alexandrov maximum principle in [11] (and, of course, Caffarelli's original proof of the strict convexity of potential functions of optimal transports/solutions to Monge-Ampère equations). If the reader is familiar with [5] or [4], then they might consider directing their attention to Case 2. Case 2b is completely novel. Case 2a illustrates our argument in the setting of [4], which builds on the work of [2] and is the foundation for Case 2b.

Section 3 is dedicated to the proof of Theorem 1.6. Our proof here is similarly self-contained (and an adaptation of Caffarelli's argument of the same result in [2], but, of course, using the line of reasoning developed to prove Theorem 1.2). The Hölder regularity of ∇u is a consequence of appropriately localizing the arguments of [11].

2 Proof of Theorem 1.2

Before we begin, it will be convenient to replace the potential u by the following lower-semicontinuous extension of u outside of $\mathcal{X}$:[4]

$$\underline{u}(x) := \sup_{\substack{z\in\mathcal{X}\\ p\in\partial u(z)}} \{u(z) + p\cdot(x-z)\}.$$

Observe that $\underline{u}|_{\mathcal{X}} = u|_{\mathcal{X}}$. For notational simplicity, we shall not distinguish $\underline{u}$ from u; so when we write u in what follows, we mean $\underline{u}$.

We shall denote the domain of u by $\mathrm{dom}(u)$, namely, $\mathrm{dom}(u) := \{u < +\infty\}$. Note that $\mathrm{dom}(u)$ is convex. We recall that convex functions are locally Lipschitz inside their domain [7, Appendix A.4]. Furthermore, we shall denote the convex hull of a set $\mathcal{A}$ by $\mathrm{conv}(\mathcal{A})$.

Let ℓ define a supporting plane to the graph of u at a point in $\mathcal{X}$. Precisely,

$$\ell(x) = u(z) + p\cdot(x-z) \quad \text{for some } (z,p)\in\mathcal{X}\times\mathbb{R}^n$$

and $\ell \le u$. Note that

$$\Sigma := \{u = \ell\} = \{u \le \ell\}$$

is closed, as u is lower-semicontinuous,

$$\Sigma, \mathcal{X} \subset \mathrm{dom}(u),$$

and, because $\mathcal{Y}$ is convex,

$$\partial u(\mathbb{R}^n) \subset \overline{\partial u(\mathcal{X})} \subset \overline{\mathcal{Y}} \quad \text{and} \quad \mathscr{L}^n(\mathcal{Y}\setminus\partial u(\mathcal{X})) = 0. \tag{2.1}$$

(A proof of (2.1) can be found in [5].)

Recall that an exposed point $\hat{x}$ of $\Sigma\subset\mathbb{R}^n$ is one for which there exists a hyperplane $\Pi\subset\mathbb{R}^n$ tangent to Σ at $\hat{x}$ such that $\Pi\cap\Sigma = \{\hat{x}\}$. Also, remember that optimal/monotone transports balance mass, in the following way.

---

4 Here, $\partial u(z)$ is called the subdifferential of u at z and is defined as follows:

$$\partial u(z) := \{p\in\mathbb{R}^n : u(x)\ge u(z) + p\cdot(x-z) \text{ for all } x\in\mathcal{X}\}.$$

Moreover, for a set $\mathcal{E}\subset\mathbb{R}^n$, we define $\partial u(\mathcal{E}) := \cup_{z\in\mathcal{E}}\partial u(z)$.

Lemma 2.1. (Mass balance formula) *Let $u : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ be convex and such that $(\nabla u)_{\#}\mu = \nu$ where μ and ν are two Borel measures that vanish on all Lipschitz $(n-1)$-dimensional surfaces. Then for all Borel sets $\mathcal{E} \subset \mathbb{R}^n$,*

$$\mu(\mathcal{E}) = \nu(\partial u(\mathcal{E})).$$

Remark 2.2. The mass balance formula was originally proved for measures that are absolutely continuous with respect to Lebesgue measure [13, Lemma 4.6]. However, with respect to absolute continuity, the proof only relies on the measures in question not giving mass to the set of nondifferentiable points of a convex function. As observed in Section 1, such points are contained in a countable union of Lipschitz $(n-1)$-dimensional surfaces. So the set of nondifferentiable points of a convex function is negligible both for μ and ν under our assumption.

Finally, recall that if a nonnegative measure is locally doubling (on ellipsoids), then it is locally doubling on all bounded convex domains [11, Corollary 2.5]. After this preliminary discussion, we can now prove our main theorem.

Proof of Theorem 1.2. Proving that u is strictly convex in $\mathcal{X}$ corresponds to proving that for any supporting plane ℓ to (the graph of) u at a point in $\mathcal{X}$, the set $\Sigma = \{u = \ell\}$ is a singleton. Assuming that Σ is not a singleton, we will show that Σ both has and does not have exposed points, which cannot be; thus, Σ is a singleton, as desired.

Case 1. Σ has no exposed points. If Σ has no exposed points, then $\Sigma \supset \mathbb{R}\mathbf{e}$ for some unit vector $\mathbf{e}$. In turn, $\partial u(\mathbb{R}^n) \subset \mathbf{e}^{\perp}$ [7, Lemma A.25]. But this is impossible given (2.1):

$$0 < \mathscr{L}^n(\partial u(\mathcal{X}) \cap \mathcal{Y}) \leq \mathscr{L}^n(\partial u(\mathbb{R}^n) \cap \mathcal{Y}) \leq \mathscr{L}^n(\mathbf{e}^{\perp} \cap \mathcal{Y}) = 0.$$

Case 2. Σ has an exposed point $\hat{x}$ in $\overline{\mathcal{X}}$. Up to a translation and a rotation, we can assume that

$$\hat{x} = 0 \in \overline{\mathcal{X}}, \quad \Sigma \subset \{x_1 \leq 0\}, \quad \text{and} \quad \Sigma \cap \{x_1 = 0\} = \{0\}.$$

Since $\mathcal{X}$ is open and $\Sigma \cap \mathcal{X}$ is nonempty by construction, there is a point $x_{int} \in \Sigma \cap \mathcal{X}$ and a ball centered at this point completely contained in $\mathcal{X}$. Thus, up to a shearing transformation $x \mapsto x - \eta x_1$ with $\eta \cdot \mathbf{e}_1 = 0$, and a dilation, we may assume that

$$x_{int} = -\mathbf{e}_1 \quad \text{and} \quad \mathcal{B}_d(-\mathbf{e}_1) \Subset \mathcal{X}$$

for some $d > 0$. Finally, up to subtracting ℓ from u, we can assume that

$$\ell \equiv 0.$$

Case 2a: $0 \in \operatorname{int}(\operatorname{dom}(u)) \cap \overline{\mathcal{X}}$. As our exposed point 0 and all of the points in $\overline{\mathcal{B}_d(-\mathbf{e}_1)}$ belong to $\operatorname{int}(\operatorname{dom}(u))$, which is convex (and, by definition, open), the convex hull of the union of $\overline{\mathcal{B}_d(-\mathbf{e}_1)}$ and $\{0\}$ is contained in $\operatorname{int}(\operatorname{dom}(u))$. So there exists an open, bounded set $\mathcal{U} \Subset \operatorname{int}(\operatorname{dom}(u))$ containing $\operatorname{conv}(\overline{\mathcal{B}_d(-\mathbf{e}_1)} \cup \{0\})$. Moreover, we know that

$$\partial u(\mathcal{U}) \subset \operatorname{conv}(\overline{\nabla u(\mathcal{U})}) =: \Upsilon \subset \mathcal{B}_{\mathcal{R}} \cap \overline{\mathcal{Y}}$$

for some $\mathcal{R} > 0$ [7, Lemma A.22]. Let u^* be the Legendre transform of u, namely,

$$u^*(q) := \sup_{x \in \mathbb{R}^n} \{q \cdot x - u(x)\} \tag{2.2}$$

and define

$$\Omega := \partial u^*(\Upsilon) \supset \mathcal{U}.$$

Recalling that ∂u and ∂u^* are inverses of each other [7, Section A.4.2], we deduce that $(\nabla u)_{\#}\rho = \gamma$, where

$$\rho := \mu\llcorner\Omega \quad \text{and} \quad \gamma := \nu\llcorner\Upsilon.$$

In particular, if we let $\phi : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ be defined by

$$\phi(x) := \sup_{\substack{z\in\Omega \\ p\in\partial u(z)}} \{u(z) + p\cdot(x-z)\},$$

then, by construction, ϕ and u agree on Ω, ϕ is (globally) Lipschitz,

$$\partial\phi(\mathbb{R}^n) = \partial\phi(\Omega) = \Upsilon,$$

and $0 \in \Omega$ is an exposed point for $\{\phi = 0\} = \{\phi \le 0\}$.

Now, let $\phi_\varepsilon(x) := \phi(x) - \varepsilon(x_1 + 1)$ and define

$$\mathcal{S}_0 := \{\phi = 0\} \cap \{x_1 \ge -1\} \quad \text{and} \quad \mathcal{S}_\varepsilon := \{\phi_\varepsilon \le 0\}.$$

Also, let γ_ε be defined by

$$\gamma_\varepsilon := (\mathrm{Id} - \varepsilon\mathbf{e}_1)_\#\gamma.$$

Notice that, by construction, $\mathcal{S}_0$ is compact, $\mathcal{S}_0 \subset \{x_1 \le 0\}$, 0, $-\mathbf{e}_1 \in \mathcal{S}_0$, and $\mathcal{S}_\varepsilon \to \mathcal{S}_0$ in the Hausdorff sense as $\varepsilon \to 0$; in particular, there exists $\mathcal{D} > 0$ such that

$$\mathcal{S}_\varepsilon \subset \mathcal{B}_\mathcal{D} \quad \text{for all } \varepsilon \ll 1.$$

Also, if $a_\varepsilon > 0$ is such that $\Pi_\varepsilon := \{x_1 = a_\varepsilon\}$ is a supporting plane to $\mathcal{S}_\varepsilon$, we see that

$$a_\varepsilon \to 0 \quad \text{as } \varepsilon \to 0$$

and

$$\varepsilon = |\phi_\varepsilon(0)| \le \max_{\mathcal{S}_\varepsilon} |\phi_\varepsilon| \le (1 + a_\varepsilon)\varepsilon.$$

Let $\mathcal{A}_\varepsilon$ be the John transformation (affine map) that normalizes $\mathcal{S}_\varepsilon$ [6]:

$$\mathcal{A}_\varepsilon x := \mathcal{L}_\varepsilon(x - x_\varepsilon),$$

where x_ε is the center of mass of $\mathcal{S}_\varepsilon$ and $\mathcal{L}_\varepsilon : \mathbb{R}^n \to \mathbb{R}^n$ is a symmetric and positive definite linear transformation. Set

$$\tilde{\phi}_\varepsilon(x) := \frac{\phi_\varepsilon(\mathcal{A}_\varepsilon^{-1}x)}{\varepsilon} \quad \text{and} \quad \tilde{\mathcal{S}}_\varepsilon := \mathcal{A}_\varepsilon(\mathcal{S}_\varepsilon).$$

Then

$$\mathcal{B}_1 \subset \tilde{\mathcal{S}}_\varepsilon \subset \mathcal{B}_{n^{3/2}}$$

and

$$1 = |\tilde{\phi}_\varepsilon(\tilde{0}_\varepsilon)| \le \max_{\tilde{\mathcal{S}}_\varepsilon} |\tilde{\phi}_\varepsilon| \le 1 + a_\varepsilon \quad \text{with } \tilde{0}_\varepsilon := \mathcal{A}_\varepsilon(0).$$

Recall that affine transformations preserve the ratio of the distances between parallel planes; therefore, letting $\Pi_{-1} := \{x_1 = -1\}$, $\Pi_0 := \{x_1 = 0\}$, and $\tilde{\Pi}_i := \mathcal{A}_\varepsilon(\Pi_i)$ for $i = -1, 0, \varepsilon$, we have that

$$\frac{\mathrm{dist}(\tilde{\Pi}_0, \tilde{\Pi}_\varepsilon)}{\mathrm{dist}(\tilde{\Pi}_{-1}, \tilde{\Pi}_\varepsilon)} = \frac{\mathrm{dist}(\Pi_0, \Pi_\varepsilon)}{\mathrm{dist}(\Pi_{-1}, \Pi_\varepsilon)} = \frac{a_\varepsilon}{1 + a_\varepsilon}.$$

In turn,

$$\mathrm{dist}(\tilde{0}_\varepsilon, \partial\tilde{\mathcal{S}}_\varepsilon) \le \mathrm{dist}(\tilde{\Pi}_0, \tilde{\Pi}_\varepsilon) \le \mathrm{dist}(\tilde{\Pi}_{-1}, \tilde{\Pi}_\varepsilon)\frac{a_\varepsilon}{1 + a_\varepsilon} \le \mathrm{diam}(\tilde{\mathcal{S}}_\varepsilon)a_\varepsilon \le 2n^{3/2}a_\varepsilon,$$

and considering the cone generated by $\partial\tilde{\mathcal{S}}_\varepsilon$ over $(\tilde{0}_\varepsilon, \tilde{\phi}_\varepsilon(\tilde{0}_\varepsilon))$, we find that

$$\mathcal{K}_\varepsilon := \mathrm{conv}\left(\mathcal{B}_{r_n} \cup \left\{\frac{r_n}{a_\varepsilon}\mathbf{e}_1\right\}\right) \subset \partial\tilde{\phi}_\varepsilon(\tilde{\mathcal{S}}_\varepsilon) \quad \text{with } r_n := \frac{1}{2n^{3/2}}.$$

(For more details on this inclusion, see, e.g., [7, Theorem 2.8].) So if we let

$$\tilde{\rho}_\varepsilon := (\mathcal{A}_\varepsilon)_\# \rho \quad \text{and} \quad \tilde{\gamma}_\varepsilon := (\varepsilon^{-1}\mathcal{L}_\varepsilon^{-1})_\# \gamma_\varepsilon,$$

then $(\nabla\tilde{\phi}_\varepsilon)_\#\tilde{\rho}_\varepsilon = \tilde{\gamma}_\varepsilon$, and, by the mass balance formula,

$$\tilde{\gamma}_\varepsilon(\mathcal{K}_\varepsilon) \le \tilde{\gamma}_\varepsilon(\partial\tilde{\phi}_\varepsilon(\tilde{\mathcal{S}}_\varepsilon)) = \tilde{\rho}_\varepsilon(\tilde{\mathcal{S}}_\varepsilon) \le \tilde{\rho}_\varepsilon\Big(\mathcal{B}_{n^{3/2}}\Big). \tag{2.3}$$

On the other hand, since $\mathcal{S}_\varepsilon \subset \mathcal{B}_{\mathcal{D}}$ and $\tilde{\mathcal{S}}_\varepsilon \supset \mathcal{B}_1$ for $\varepsilon \ll 1$, we see that

$$|\mathcal{A}_\varepsilon(w) - \mathcal{A}_\varepsilon(z)| \ge \frac{1}{\mathcal{D}}|w - z| \quad \text{for all } w, z \in \mathbb{R}^n.$$

In turn, for $\varepsilon \ll 1$,

$$\tilde{\Omega}_\varepsilon := \mathcal{A}_\varepsilon(\Omega) \supset \mathcal{A}_\varepsilon(\mathcal{B}_d(-\mathbf{e}_1)) \supset \mathcal{B}_{\frac{d}{\mathcal{D}}}(\mathcal{A}_\varepsilon(-\mathbf{e}_1)).$$

Therefore, if we define

$$\tilde{\mathcal{S}}_{\varepsilon,d} := \left\{\operatorname{dist}(\cdot,\partial\tilde{\mathcal{S}}_\varepsilon) \ge \frac{d}{2\mathcal{D}}\right\},$$

then there exists a dimensional constant $C_n > 0$ and a point $\tilde{z}_d$ such that

$$\mathcal{A}_\varepsilon(\mathcal{B}_d(-\mathbf{e}_1)) \cap \tilde{\mathcal{S}}_{\varepsilon,d} \supset \mathcal{B}_{\frac{d}{C_n\mathcal{D}}}(\tilde{z}_d).$$

Also, by, for example, [7, Corollary A.23],

$$\partial\tilde{\phi}_\varepsilon(\tilde{\mathcal{S}}_{\varepsilon,d}) \subset \mathcal{B}_{\frac{6\mathcal{D}}{d}}.$$

Thus, for all $\varepsilon \ll 1$,

$$\tilde{\rho}_\varepsilon\Big(\mathcal{B}_{n^{3/2}}\Big) \le \tilde{\mu}_\varepsilon\Big(\mathcal{B}_{2n^{3/2}}(\tilde{z}_d)\Big) \le C_\mu^k\tilde{\mu}_\varepsilon\Big(\mathcal{B}_{\frac{d}{C_n\mathcal{D}}}(\tilde{z}_d)\Big) = C_\mu^k\tilde{\rho}_\varepsilon\Big(\mathcal{B}_{\frac{d}{C_n\mathcal{D}}}(\tilde{z}_d)\Big), \tag{2.4}$$

where $\tilde{\mu}_\varepsilon := (\mathcal{A}_\varepsilon)_\#\mu$, the number $k \in \mathbb{N}$ is such that $2n^{3/2} \le 2^k\frac{d}{C_n\mathcal{D}}$, and C_μ is the doubling constant for μ in $\mathcal{B}_{4\mathcal{D}n^{3/2}}$. (The last equality holds since $\tilde{\mu}_\varepsilon$ and $\tilde{\rho}_\varepsilon$ agree on $\tilde{\Omega}_\varepsilon$.) Moreover, by using the mass balance formula again, we deduce that

$$\tilde{\rho}_\varepsilon\Big(\mathcal{B}_{\frac{d}{C_n\mathcal{D}}}(\tilde{z}_d)\Big) \le \tilde{\rho}_\varepsilon(\mathcal{A}_\varepsilon(\mathcal{B}_d(-\mathbf{e}_1)) \cap \tilde{\mathcal{S}}_{\varepsilon,d}) = \tilde{\gamma}_\varepsilon(\partial\tilde{\phi}_\varepsilon(\mathcal{A}_\varepsilon(\mathcal{B}_d(-\mathbf{e}_1)) \cap \tilde{\mathcal{S}}_{\varepsilon,d})) \le \tilde{\gamma}_\varepsilon\Big(\mathcal{B}_{\frac{6\mathcal{D}}{d}}\Big) \tag{2.5}$$

for all $\varepsilon \ll 1$. Consequently, combining the three chains of inequalities (2.3), (2.4), and (2.5), we have that

$$\tilde{\gamma}_\varepsilon(\mathcal{K}_\varepsilon) \le C_\mu^k\tilde{\gamma}_\varepsilon\Big(\mathcal{B}_{\frac{6\mathcal{D}}{d}}\Big). \tag{2.6}$$

Now, let $t_m\mathbf{e}_1 \in \mathcal{K}_\varepsilon$ for $m = 1, \dots, \mathcal{M}$ be a sequence of points chosen[5] so that

$$\frac{1}{2}\mathcal{K}_m \subset \mathcal{K}_m\backslash\mathcal{K}_{m-1} \quad \text{with } \mathcal{K}_m := \operatorname{conv}\Big(\mathcal{B}_{r_n} \cup \{t_m\mathbf{e}_1\}\Big) \quad \text{and} \quad \mathcal{K}_0 := \mathcal{B}_{r_n}.$$

By construction, $\left\{\frac{1}{2}\mathcal{K}_m\right\}_{m=1}^{\mathcal{M}}$ is a disjoint family, and

$$\mathcal{M} = \mathcal{M}(a_\varepsilon) \to \infty \quad \text{as } a_\varepsilon \to 0. \tag{2.7}$$

Hence, since $\varepsilon\mathcal{L}_\varepsilon(\mathcal{K}_\varepsilon) \subset \Upsilon$, we find that

---

5 A possible way to construct such a sequence is to choose $t_m = 5^m$. To ensure that $t_m\mathbf{e}_1 \in \mathcal{K}_\varepsilon$ for any $m = 1, \dots, \mathcal{M}$, one needs $\mathcal{M} \le \frac{\log r_n - \log a_\varepsilon}{\log 5}$.

$$\mathcal{M}\tilde{\gamma}_\varepsilon\big(\mathcal{B}_{r_n}\big) \le \sum_{m=1}^{\mathcal{M}} \tilde{\gamma}_\varepsilon(\mathcal{K}_m) \le C_\gamma \sum_{m=1}^{\mathcal{M}} \tilde{\gamma}_\varepsilon\Big(\frac{1}{2}\mathcal{K}_m\Big) \le C_\gamma \tilde{\gamma}_\varepsilon(\mathcal{K}_\varepsilon),$$

with C_γ denoting the doubling constant for γ in $\mathcal{B}_{2\mathcal{R}}$, which is the same as the doubling constant for ν in $\mathcal{B}_{2\mathcal{R}}$; since Υ is convex, γ inherits its doubling property from ν. All in all, considering the aforementioned chain of inequalities and (2.6) and denoting by $j \in \mathbb{N}$ the smallest number such that $\frac{6\mathcal{D}}{d} \le 2^j r_n$, we see that

$$0 < \mathcal{M}\tilde{\gamma}_\varepsilon\big(\mathcal{B}_{r_n}\big) \le C_\mu^k C_\gamma \tilde{\gamma}_\varepsilon\Big(\mathcal{B}_{\frac{6\mathcal{D}}{d}}\Big) \le C_\mu^k C_\gamma^{j+1} \tilde{\gamma}_\varepsilon\big(\mathcal{B}_{r_n}\big),$$

or, equivalently,

$$\mathcal{M} \le C_\mu^k C_\gamma^{j+1}.$$

But this is impossible for small ε, concluding the proof.

Case 2b: $0 \in \partial(\mathrm{dom}(u)) \cap \overline{\mathcal{X}}$. In this subcase, let $u_\varepsilon(x) \coloneqq u(x) - \varepsilon(x_1 + 1)$ and define

$$\mathcal{S}_0 \coloneqq \Sigma \cap \{x_1 \ge -1\} \quad \text{and} \quad \mathcal{S}_\varepsilon \coloneqq \{u_\varepsilon \le 0\}.$$

Like before, for all $\varepsilon \ll 1$,

$$\mathcal{S}_\varepsilon \subset \mathcal{B}_{\mathcal{D}}$$

for some $\mathcal{D} > 0$. Here, however, as $0 \in \partial(\mathrm{dom}(u))$, we have that

$$\partial\tilde{u}_\varepsilon(\tilde{\mathcal{S}}_\varepsilon) \supset \mathrm{conv}\big(\mathcal{B}_{r_n} \cup \mathbb{R}^+\mathbf{e}_1\big) \quad \text{with } r_n \coloneqq \frac{1}{2n^{3/2}}.$$

The function $\tilde{u}_\varepsilon$ is defined in an analogous fashion to how $\tilde{\phi}_\varepsilon$ was defined in Case 2a (but replacing ϕ by u) and, again, $\tilde{\mathcal{S}}_\varepsilon \coloneqq \mathcal{A}_\varepsilon(\mathcal{S}_\varepsilon)$ with $\mathcal{A}_\varepsilon$ denoting the John map associated to $\mathcal{S}_\varepsilon$ whose linear part is $\mathcal{L}_\varepsilon$. In turn, arguing as we did in Case 2a, where again $k \in \mathbb{N}$ is such that $2n^{3/2} \le 2^k \frac{d}{C_n\mathcal{D}}$ and C_μ is the doubling constant for μ in $\mathcal{B}_{4\mathcal{D}n^{3/2}}$, but in the original variables, we deduce that

$$\nu_\varepsilon(\varepsilon\mathcal{L}_\varepsilon(\mathcal{K}_\varepsilon)) \le C_\mu^k \nu_\varepsilon\Big(\varepsilon\mathcal{L}_\varepsilon\Big(\mathcal{B}_{\frac{6\mathcal{D}}{d}}\Big)\Big)$$

for all $\varepsilon \ll 1$ (cf. (2.6)). Here, instead, $\nu_\varepsilon \coloneqq (\mathrm{Id} - \varepsilon\mathbf{e}_1)_\# \nu$ and

$$\mathcal{K}_\varepsilon \coloneqq \mathrm{conv}\Big(\mathcal{B}_{r_n} \cup \Big\{\frac{r_n}{\varepsilon}\mathbf{e}_1\Big\}\Big).$$

Now notice that

$$|\mathcal{L}_\varepsilon(\mathbf{e}_1)| = |\mathcal{A}_\varepsilon(0) - \mathcal{A}_\varepsilon(-\mathbf{e}_1)| \le 2n^{3/2}.$$

Moreover, we claim there exists an $\mathcal{N} \gg 2n^{3/2} > 0$ such that

$$\|\varepsilon\mathcal{L}_\varepsilon\| \le \mathcal{N} \quad \text{for all } \varepsilon \ll 1.$$

Indeed, if not, then we can find a sequence of points $z_\varepsilon \in \mathcal{S}_\varepsilon$ and slopes $p_\varepsilon \in \partial u_\varepsilon(z_\varepsilon) \cap \mathrm{span}(\mathcal{S}_0)^\perp$ such that $|p_\varepsilon| \to \infty$. In particular, in the limit, we find a point $z_0 \in \mathcal{S}_0$ such that $\partial u(z_0) \cap \mathrm{span}(\mathcal{S}_0)^\perp$ contains a sequence of slopes $\{p_j\}_{j\in\mathbb{N}}$ with $|p_j| = j$. But as $p_j \in \mathrm{span}(\mathcal{S}_0)^\perp$, we see that $p_j \cdot (x - z_0) = p_j \cdot x = p_j \cdot (x - z)$ for any $z \in \mathcal{S}_0$. Hence, $p_j \in \partial u(z)$ for all $z \in \mathcal{S}_0$ and $j \in \mathbb{N}$. However, this is impossible; $\mathcal{S}_0 \cap \mathrm{int}(\mathrm{dom}(u))$ is nonempty, and on this set, u is locally Lipschitz, proving the claim.

Therefore,

$$\varepsilon\mathcal{L}_\varepsilon(\mathcal{K}_\varepsilon) \subset \mathcal{B}_{\mathcal{N}} \quad \text{and} \quad \varepsilon\mathcal{L}_\varepsilon\Big(\mathcal{B}_{\frac{6\mathcal{D}}{d}}\Big) \subset \mathcal{B}_{\frac{6\mathcal{D}\mathcal{N}}{d}}.$$

And so, arguing exactly like we did in Case 2a, we find that

$$\mathcal{M} \le C_\mu^k C_\nu^{j+1},$$

where C_ν is the doubling constant for ν in $\mathcal{B}_{6\mathcal{D}\mathcal{N}/d}$ and $\mathcal{M} = \mathcal{M}(\varepsilon) \to \infty$ as $\varepsilon \to 0$ is the analogous count for this case's $\mathcal{K}_\varepsilon$ (cf. (2.7)). But, again, this is impossible.

Case 3. Σ has an exposed point $\hat{x}$ in $\mathbb{R}^n\backslash\overline{\mathcal{X}}$. In this case, up to a translation, a dilation, a rotation, and subtracting ℓ from u, we can assume that

$$\hat{x} = 0, \quad \Sigma \subset \{x_1 \le 0\}, \quad \ell \equiv 0, \quad \text{and} \quad \mathcal{S}_0 \coloneqq \Sigma \cap \{x_1 \ge -1\} \Subset \mathbb{R}^n\backslash\overline{\mathcal{X}}.$$

Like before, let $u_\varepsilon(x) \coloneqq u(x) - \varepsilon(x_1 + 1)$ and define

$$\mathcal{S}_\varepsilon \coloneqq \{u_\varepsilon \le 0\} \quad \text{and} \quad \nu_\varepsilon \coloneqq (\mathrm{Id} - \varepsilon\mathbf{e}_1)_\# \nu.$$

Again, $\mathcal{S}_\varepsilon \to \mathcal{S}_0$ as $\varepsilon \to 0$, so

$$\mathrm{diam}(\mathcal{S}_\varepsilon) \le 2\mathrm{diam}(\mathcal{S}_0) \quad \text{and} \quad \mathcal{S}_\varepsilon \Subset \mathbb{R}^n\backslash\overline{\mathcal{X}}$$

for all $\varepsilon \ll 1$. For these small positive ε, then,

$$0 = \mu(\mathcal{S}_\varepsilon) = \nu_\varepsilon(\partial u_\varepsilon(\mathcal{S}_\varepsilon)),$$

where the second equality follows from the mass balance formula. (Recall that μ vanishes on $\mathbb{R}^n\backslash\overline{\mathcal{X}}$.) Moreover, as $\mathcal{Y}$ is convex,

$$\partial u_\varepsilon(\mathcal{S}_\varepsilon) \subset \mathrm{spt}(\nu_\varepsilon) = \overline{\mathcal{Y}} - \varepsilon\mathbf{e}_1$$

(cf. (2.1)). Thus, any open subset of $\partial u_\varepsilon(\mathcal{S}_\varepsilon)$ must be in the interior of the support of ν_ε. In turn, considering the cone generated by $\partial\mathcal{S}_\varepsilon$ over $(0, u_\varepsilon(0))$, for $0 < \varepsilon \ll 1$, we find that

$$0 = \nu_\varepsilon(\partial u_\varepsilon(\mathcal{S}_\varepsilon)) \ge \nu_\varepsilon\big(\mathcal{B}_{r_\varepsilon}\big) > 0 \quad \text{with } r_\varepsilon \coloneqq \frac{|u_\varepsilon(0)|}{2\mathrm{diam}(\mathcal{S}_0)}.$$

(Again, for more details on this inclusion, see, e.g., [7, Theorem 2.8].) This is a contradiction and concludes the proof. □

3 Proof of Theorem 1.6

Again, we replace u by its lower-semicontinuous extension outside of $\mathcal{X}$, exactly as we did at the beginning of Section 2. We split the proof into three parts.

Part 1. u is continuously differentiable inside $\mathcal{X}$. We follow the argument used to prove [2, Corollary 1]. Assume for the sake of a contradiction that the result is false. Up to a translation, let $0 \in \mathcal{X}$ be a point at which u has two distinct supporting planes. After a rotation, dilation, and subtracting off an affine function from u, we may assume that

$$u(x) \ge \max\{x_1, 0\}, \quad u(0) = 0, \quad \text{and} \quad \frac{u(-s\mathbf{e}_1)}{s} \to 0 \quad \text{as } s \to 0.$$

Now consider the function u_σ defined by

$$u_\sigma(x) \coloneqq u(x) - \tau(x_1 + 2\sigma) \quad \text{with } \tau \coloneqq \frac{u(-\sigma\mathbf{e}_1)}{\sigma}.$$

Note that $\tau \to 0$ as $\sigma \to 0$. If

$$\mathcal{S}_\sigma \coloneqq \{u_\sigma \le 0\},$$

then, by the strict convexity of u provided by Theorem 1.2, we see that

$$\mathcal{S}_\sigma \subset \mathcal{B}_\mathcal{D} \Subset \mathcal{X} \cap \mathrm{int}(\mathrm{dom}(u))$$

for some $\mathcal{D} > 0$ and for all $\sigma \ll 1$; also, for these small positive σ,

$$\partial u_\sigma(\mathcal{S}_\sigma) \subset \mathcal{B}_{\mathcal{R}} \cap \mathcal{Y}$$

for some $\mathcal{R} > 0$. Moreover, if $\Pi_{-a} := \{x_1 = -a\}$ and $\Pi_b := \{x_1 = b\}$ denote the two parallel planes that tangentially sandwich $\mathcal{S}_\sigma$, we see that

$$a > \sigma \quad \text{and} \quad b < \frac{2\tau\sigma}{1-\tau},$$

provided $\sigma > 0$ is small enough to guarantee that $\tau < 1$. Furthermore,

$$\max_{\mathcal{S}_\sigma} |u_\sigma| = |u_\sigma(0)| = 2\tau\sigma,$$

and

$$\frac{\operatorname{dist}(\Pi_b, \Pi_0)}{\operatorname{dist}(\Pi_{-a}, \Pi_b)} = \frac{b}{a+b} \le \frac{b}{a} \le \frac{2\tau}{1-\tau} \to 0 \quad \text{as } \sigma \to 0,$$

where $\Pi_0 := \{x_1 = 0\}$.

Now, set

$$\tilde{u}_\sigma(x) := \frac{u_\sigma(\mathcal{A}_\sigma^{-1}x)}{2\tau\sigma} \quad \text{and} \quad \tilde{\mathcal{S}}_\sigma := \mathcal{A}_\sigma(\mathcal{S}_\sigma),$$

where $\mathcal{A}_\sigma$ is the John map that normalizes $\mathcal{S}_\sigma$. Arguing as we did in the proof of Theorem 1.2, we find the same contradiction as we did in Case 2 when τ sufficiently small; the only difference is that we consider a slightly different chain of inequalities:

$$\tilde{\nu}_\sigma(\mathcal{K}_\sigma) \le \tilde{\nu}_\sigma(\partial\tilde{u}_\sigma(\tilde{\mathcal{S}}_\sigma)) = \tilde{\mu}_\sigma(\tilde{\mathcal{S}}_\sigma) \le C_\mu \tilde{\mu}_\sigma\left(\frac{1}{2}\tilde{\mathcal{S}}_\sigma\right) = C_\mu \tilde{\nu}_\sigma\left(\partial\tilde{u}_\sigma\left(\frac{1}{2}\tilde{\mathcal{S}}_\sigma\right)\right) \le C_\mu \tilde{\nu}_\sigma\left(\mathcal{B}_{\frac{1}{r_n}}\right)$$

(cf. (2.6)), where $\tilde{\mu}_\sigma$ and $\tilde{\nu}_\sigma$ are defined so that $(\nabla\tilde{u}_\sigma)_\#\tilde{\mu}_\sigma = \tilde{\nu}_\sigma$ and

$$\mathcal{K}_\sigma := \operatorname{conv}\left(\mathcal{B}_{r_n} \cup \left\{\frac{r_n + n(1-\tau)}{2\tau}\mathbf{e}_1\right\}\right) \quad \text{with } r_n := \frac{1}{2n^{3/2}}.$$

This proves that u is differentiable.

By [7, Lemma A.24], for example, we know that differentiable convex functions are continuously differentiable. So we conclude that u is continuously differentiable in $\mathcal{X}$.

Part 2. $\nabla u(\mathcal{X}) = \mathcal{Y}$ when $\mathcal{X}$ is convex. Because ∇u is continuous in $\mathcal{X}$, its image $\mathcal{Y}' := \nabla u(\mathcal{X})$ is an open set of full ν-measure contained inside $\mathcal{Y}$. Also, as the assumptions on μ and ν are symmetric, the optimal transport map ∇v from ν to μ is continuous, and $\mathcal{X}' := \nabla v(\mathcal{Y})$ is an open set of full μ-measure contained inside $\mathcal{X}$. Hence, by recalling that ∇u and ∇v are inverses of each other [8, Corollary 2.5.13], we conclude that $\mathcal{X}' = \mathcal{X}$ and $\mathcal{Y}' = \mathcal{Y}$, as desired.

Part 3. ∇u is locally Hölder continuous inside $\mathcal{X}$. Thanks to the strict convexity and C^1 regularity of u, we can localize the arguments of the proof of [11, Theorem 1.1] to obtain the local Hölder continuity of u inside $\mathcal{X}$.

More precisely, if u^* denotes the Legendre transform of u (2.2), as in [11], one can show that u^* satisfies a weak form of Alexandrov's maximum principle [11, Lemma 3.2], from which one deduces the engulfing property for the sections of u^* [11, Lemma 3.3]. Iteratively applying this engulfing property, one obtains a polynomial strict convexity bound for u^*. This bound implies the local Hölder continuity of u inside $\mathcal{X}$ [11, Proof of Theorem 1.1]. We leave the details of this adaptation to the interested reader.

Funding information: AF acknowledges the support of the ERC (Grant No. 721675 "Regularity and Stability in Partial Differential Equations (RSPDE)") and the Lagrange Mathematics and Computation Research Center. YJ was supported in part by the NSF (Grant No. DMS-1954363).

Conflict of interest: The authors state no conflict of interest.

References

[1] Y. Brenier, *Polar factorization and monotone rearrangement of vector-valued functions*, Comm. Pure Appl. Math. **44** (1991), no. 4, 365–417.
[2] L.A. Caffarelli, *A localization property of viscosity solutions to the Monge-Ampère equation and their strict convexity*, Ann. Math. **131** (1990), 129–134.
[3] L.A. Caffarelli, *Some regularity properties of solutions of Monge-Ampère equation*, Comm. Pure Appl. Math. **44** (1991), no. 8–9, 965–969.
[4] L.A. Caffarelli, *The regularity of mappings with a convex potential*, J. Amer. Math. Soc. **5** (1992), no. 1, 99–104.
[5] D. Cordero-Erausquin and A. Figalli, *Regularity of monotone transport maps between unbounded domains*, Discrete Contin. Dyn. Syst. **39** (2019), no. 12, 7101–7112.
[6] C. E. Gutiérrez, *The Monge-Ampère equation*, Progr. Nonlinear Differential Equations Appl. vol. 44, Birkhäuser Boston, Boston, MA, 2001.
[7] A. Figalli, *The Monge-Ampère equation and its applications*, Zurich Lectures in Advanced Mathematics, European Mathematical Society (EMS), Zürich, 2017, p. x+200.
[8] A. Figalli and F. Glaudo, *An invitation to optimal transport, Wasserstein distances, and gradient flows*, EMS Textbooks in Mathematics, EMS Press, Berlin, 2021, vi+136 pp.
[9] D. Jerison, *Harmonic measure in convex domains*, Bull. Amer. Math. Soc. (N.S.) **21** (1989), no. 2, 255–260.
[10] Y. Jhaveri, *On the (in)stability of the identity map in optimal transportation*, Calc. Var. Partial Differential Equations **58** (2019), no. 3, Art. 96, 25 pp.
[11] Y. Jhaveri and O. Savin, *On the regularity of optimal transports between degenerate densities*, Arch. Ration. Mech. Anal. **245** (2022), no. 2, 819–861.
[12] R. J. McCann, *Existence and uniqueness of monotone measure-preserving maps*, Duke Math. J. **80** (1995), 309–323.
[13] C. Villani, *Topics in optimal transportation*, Graduate Studies in Mathematics, vol. 58, American Mathematical Society, Providence, RI, 2003, xvi+370 pp.
[14] L. Zajíček, *On the points of multiplicity of monotone operators*, Comment. Math. Univ. Carolinae **19** (1978), no. 1, 179–189.