

A perturbation argument for a Monge-Ampère type equation arising in optimal transportation

Luis Caffarelli*
Univ. Texas Austin

María del Mar González†
Univ. Politècnica de Catalunya

Truyen Nguyen‡
Univ. of Akron

Abstract

We prove some asymptotic interior regularity results for potential functions of optimal transportation problems with power costs. We show that our problems are equivalent to optimal transportation problems whose cost functions are sufficiently small perturbations of the quadratic cost but they do not satisfy the well known condition (A.3) guaranteeing regularity. The proof consists in a perturbation argument from the standard Monge-Ampère equation in order to obtain interior Hölder estimates for second derivatives of potentials, and a careful understanding of why we might fail to have an Alexandroff weak solution when restricted to subdomains. In particular, we provide some quantitative estimates along the way on how the equation degenerates near the boundary.

1 Introduction

Let Ω, Ω' be two bounded domains in \mathbb{R}^n , $n \geq 1$. Let f, g be nonnegative mass distributions in Ω and Ω' , respectively, that satisfy the compatibility condition

$$\int_{\Omega} f = \int_{\Omega'} g.$$

Consider the optimal transport problem between Ω and Ω' : given a continuous cost function $c : \mathbb{R}^n \rightarrow [0, \infty)$, find an admissible transport map $\Theta : \Omega \rightarrow \Omega'$ that minimizes the total cost functional

$$\int_{\Omega} c(x - \Theta(x))f(x) dx.$$

We say that a Borel map $\Theta : \Omega \rightarrow \Omega'$ is admissible if it is measure preserving, that is, $\Theta_{\#}f dx = g dy$. This is equivalent to

$$\int_{\Omega} h(\Theta(x))f(x) dx = \int_{\Omega'} h(y)g(y) dy$$

for every continuous function h . It is well known (c.f. [9], [20]) that the solution of the transport problem exists for any strictly convex cost, and it is given in terms of a potential function $u \in \mathcal{C}(\bar{\Omega})$, that satisfies a Monge-Ampère type equation

$$\begin{cases} g(x - \nabla c^*(-\nabla u)) \det(I + D^2 c^*(-\nabla u) D^2 u) = f(x), \\ \partial_c u(\Omega) = \Omega', \end{cases} \quad (1.1)$$

*Partially supported by National Science Foundation Grant DMS-0654267.

†Supported by Spain Government projects MTM2008-06349-C03-01, MTM2011-27739-C04-01, GenCat 2009SGR345 and NSF grant DMS-0635607.

‡Supported by NSF grant DMS-0901449.

where I is the $n \times n$ identity matrix, $\partial_c u(\Omega) := \cup_{x \in \Omega} \partial_c u(x)$ and the c -subdifferential $\partial_c u(x)$ is defined by (2.3).

In this paper we concentrate on the convex cost $c(z) = \frac{1}{p}|z|^p$ ($p > 1$) and investigate the regularity of its corresponding optimal map. First we study the asymptotic behavior of the equation when the two domains Ω and Ω' are sufficiently far away. Second, we consider the perturbation problem $p > 2$, $p \rightarrow 2$ but Ω and Ω' do not need to be far apart. In both cases, our aim is to prove interior $\mathcal{C}^{2,\alpha}$ estimates for the potential u . This is achieved by showing that our problems are equivalent to optimal transportation problems whose cost functions are close to the quadratic cost. We then perform a perturbation argument around the standard Monge-Ampère equation to derive the regularity of u .

Before stating precisely our main result, we recall that a bounded domain $U \subset \mathbb{R}^n$ is called strongly convex if U has a \mathcal{C}^2 defining function ρ satisfying $\langle D^2 \rho(x) \xi, \xi \rangle > 0$ for all $x \in \partial U$ and all vectors ξ in the tangent space to ∂U at x . It is also known that if U is strongly convex then there exist a constant $C > 0$ and a defining function $\tilde{\rho}$ for U such that

$$\langle D^2 \tilde{\rho}(x) w, w \rangle \geq C |w|^2, \quad \forall x \in \partial U \text{ and } \forall w \in \mathbb{R}^n. \quad (1.2)$$

For any $\delta > 0$, denote

$$\Omega_\delta := \{x \in \Omega : \text{dist}(x, \partial\Omega) > \delta\}.$$

Theorem 1.1. *Suppose Ω and Ω' are two strongly convex bounded domains in \mathbb{R}^n . Assume that $f \in \mathcal{C}_{loc}^\alpha(\Omega)$ and $g \in \mathcal{C}_{loc}^\alpha(\Omega')$ are two nonnegative mass distributions satisfying*

$$\int_\Omega f(x) dx = \int_{\Omega'} g(y) dy, \quad \Lambda_1 \leq f(x) \leq \Lambda_2 \text{ in } \Omega \quad \text{and} \quad \Lambda_1 \leq g(y) \leq \Lambda_2 \text{ in } \Omega', \quad (1.3)$$

where $\Lambda_1 < \Lambda_2$ are positive constants. Let $u \in \mathcal{C}(\bar{\Omega})$ be the potential from the optimal transport problem (1.1) between Ω and Ω' with cost $c(z) = \frac{1}{p}|z|^p$, $p > 1$. Given any $\delta_0 > 0$, we have:

- There exists $\lambda_0 > 0$ depending on $\delta_0, n, p, \alpha, \Lambda_1, \Lambda_2, \Omega$ and Ω' such that if $\text{dist}(\Omega, \Omega') \geq \lambda_0$, then $u \in \mathcal{C}^{2,\alpha'}(\Omega_{\delta_0})$ for some $0 < \alpha' < \alpha$ depending only on n and α .
- There exists $p_0 > 2$ depending on $\delta_0, n, \alpha, \Lambda_1, \Lambda_2, \Omega$ and Ω' such that for all $2 < p \leq p_0$, if $\text{dist}(\Omega, \Omega') > C_1$ for some constant $C_1 > 0$, we also have that $u \in \mathcal{C}^{2,\alpha'}(\Omega_{\delta_0})$ for some $0 < \alpha' < \alpha$ depending only on n and α .

The $\mathcal{C}^{2,\alpha'}$ norm depends on $\delta_0, n, \Lambda_1, \Lambda_2, \alpha, \Omega, \Omega', \|f\|_{\mathcal{C}^\alpha(\Omega_{\delta_1})}, \|g\|_{\mathcal{C}^\alpha(\Omega'_{\delta_1})}, C_1$. Here δ_1 depends only on δ_0 and universal constants.

Remark. Standard elliptic theory implies that $u \in \mathcal{C}^{2,\alpha}(\Omega_{\delta_0})$.

We emphasize that, unlike the standard subdifferential for convex functions, in general the c -subdifferential associated to a power cost function is a nonlocal operator on the set of c -convex functions. A similar nonlocality is then inherited by the corresponding Monge-Ampère measures which makes the so called Alexandroff weak solutions to the equation (1.1) unstable when restricting to subdomains. This poses a serious obstacle in proving the regularity of the potential u since we loose information as we move into a smaller subregion and it is possible that the restricted function u is not a solution of the same equation in any reasonable sense. We note however that this problem would not present if we knew that u is differentiable everywhere.

As a consequence, our first problem is to determine whether the optimal transport map is continuous, i.e., that the potential u is \mathcal{C}^1 . This is a tricky question that has received a lot of attention recently. In the quadratic case, for the optimal transport with cost $c_q(z) = \frac{1}{2}|z|^2$, this problem is equivalent to proving strict convexity of the potential (see Caffarelli [7]). For the general case, Ma-Trudinger-Wang [27] have introduced condition (A.3) on the cost that is sufficient for regularity. The example by Loeper [25] shows that the weak form of (A.3) is also necessary in the general case. Further references can be found in Trudinger-Wang [30], [29], [31], Figalli-Loeper [18], Kim-McCann [26].

Note that (A.3) and its weak form are conditions on the cost that involve fourth order derivatives but, in particular, they do not include the power cost $\frac{1}{p}|z|^p$, $p > 1$. More precisely, (A.3) is written as

$$\sum_{i,j,k,l,p,q,r,s} (c^{p,q}c_{ij,p}c_{q,rs} - c_{ij,rs}) c^{r,k} c^{s,l} \xi_i \xi_j \eta_k \eta_l \geq c_0 |\xi|^2 |\eta|^2$$

for all $\xi, \eta \in \mathbb{R}^n$ satisfying $\xi \perp \eta$, where $c_{i,j}(x, y) = D_{ij}c(x, y)$, $(c^{i,j})$ is the inverse matrix of $(c_{i,j})$ and c_0 is a positive constant. Thus the so called “double mountain above sliding mountain” property established in [26] is not true in our case and it is interesting to know whether the local c -subdifferential is still the same as the global c -subdifferential which is especially important in localizing our problem to understand the smoothness of u . As an attempt at answering this nontrivial question, we explore the problem of strict convexity in depth in order to obtain more refined information of the possible degeneracies and to describe the geometry of optimal transportation. By comparing with the optimal transport problem for the quadratic cost, we are able to show that the sections of u have good shapes and hence their eccentricity is controlled. We then employ this to demonstrate that: for a given subdomain, any local c -subdifferential of u with respect to the subdomain at x is also a global c -subdifferential of u at x as long as the point x is sufficiently away from the boundary of the subdomain. That is, the a priori lost information can only happen in a very small region near the boundary of the subdomain. This indicates that although the restricted potential u might no longer be an Alexandroff weak solution of the equation (1.1) in the subdomain, it should be close to such a solution of the equation. We quantify this kind of almost solutions by introducing notions of Alexandroff weak solutions and viscosity solutions in one region relative to another region. Then we give a careful description of these very weak solutions, so that the machinery for second order elliptic equations can be applied.

The next step is to show second derivative Hölder estimates. In the case of quadratic cost, they were obtained by Caffarelli in [5], where, more generally, he showed $W^{2,p}$ estimates; also the Park City notes [11] contain a sketch of the $\mathcal{C}^{1,1}$ estimates. The recent note by Gutiérrez, Huang and Nguyen [23] thoroughly explores the ideas in Caffarelli’s work to give a $\mathcal{C}^{2,\alpha}$ estimate directly. These proofs consist in the understanding of the regularity of the Monge-Ampère equation

$$g(x + \nabla w) \det(I + D^2 w) = f(x). \quad (1.4)$$

The purpose of this paper is to treat (1.1) as a perturbation of (1.4), that is small when the distance λ between Ω and Ω' is very big, or when $p \rightarrow 2$. The ideas first appeared in Caffarelli’s work [4] in the setting of fully nonlinear uniformly elliptic equations, (see also the book [12]). However, some difficulties arise in our work: first of all we need to work with almost solutions of (1.1) and estimate explicitly the closeness between them and solutions of the Dirichlet problem for the standard Monge-Ampère equation. We overcome this obstacle by proving a comparison principle which allows us to compare very weak solutions coming

from two different equations. Second, estimates are usually given in terms of good Dirichlet boundary data. To handle this, we localize the problem and work with sections, which is doable thanks to our previous discussion on restriction and localization. Third, the estimates deteriorate at the boundary of the sections; this deterioration will compensate with the fact that the sections become more and more round, but it needs to be quantified explicitly.

Some related works to the present that have recently appeared are the following: asymptotic regularity of optimal transport on manifolds (Delanoë [15], Delanoë-Ge [16]), regularity for Gaussian-to-Gaussian maps with near Euclidean cost (Warren [34]).

The paper is structured as follows: Section 2 is the basic set up for our problem, together with a quick review on optimal transportation for convex costs, while Section 3 contains some geometrical properties of our special cost that will be needed at a later time. Section 4 is the starting point in the proof of the main theorem: a comparison to the optimal transport problem for quadratic cost to get good control of the sections of u . Then, in Section 5 we give an (almost) global c -subdifferential result, and further explore the concepts of Alexandroff and viscosity solutions in a relative sense. Next, Section 6 deals with the comparison principle for viscosity solutions, that allows to show that the potential u is close to a quadratic polynomial locally. Finally, in the last Section 7 we put together all these results in order to complete the proof of Theorem 1.1.

Acknowledgements. M.d.M. González and T. Nguyen gratefully acknowledge the support provided by the Mathematical Sciences Research Institute at Berkeley and the Institute for Advanced Study at Princeton where parts of this work were carried out. T. Nguyen also would like to thank Neil Trudinger for fruitful discussions on regularity of optimal maps.

2 Set up

2.1 A review of convex costs

Here we review some standard background for equation (1.1) when the cost function $c : \mathbb{R}^n \rightarrow [0, \infty)$ is strictly convex. The basic references for optimal transportation are the books by Ambrosio-Gigli-Savaré [1] and Villani ([32], or its most recent book [33]). In this case, it is well known that the optimal transport map exists and is unique. This was studied independently by Caffarelli [9] and Gangbo-McCann [19] following the ideas introduced by Brenier in [2] for the quadratic cost $c(z) = \frac{1}{2}|z|^2$. To prove the existence of such map, they considered the dual functional of Kantorovich

$$I(\phi, \psi) = - \int_{\Omega} f(x)\phi(x)dx - \int_{\Omega'} g(y)\psi(y)dy \quad (2.1)$$

for pairs (ϕ, ψ) of Lipschitz functions in the set

$$K := \{(\phi, \psi) : -\phi(x) - \psi(y) \leq c(x - y) \text{ for all } x \in \Omega, y \in \Omega'\}.$$

Now, it is possible to show that a maximizing pair (u, v) for the Kantorovich functional exists and is unique up to a constant. Furthermore if we let

$$\Theta_u(x) := x - \nabla c^*(-\nabla u),$$

then this is the solution of the transport problem, where we have defined c^* , the Legendre transform of c , as

$$c^*(y) = \sup_{x \in \mathbb{R}^n} \{\langle x, y \rangle - c(x)\}.$$

The function $u : \Omega \rightarrow \mathbb{R}$ is known as the potential, and from the proof we obtain a representation of the form

$$u(x) = \sup_{y \in \Omega'} \{-c(x - y) - v(y)\}.$$

This potential u is locally Lipschitz and satisfies a second order fully non-linear PDE of Monge-Ampère type:

$$M[u] = f \quad \text{in } \Omega, \tag{2.2}$$

where

$$M[u] := g(x - \nabla c^*(-\nabla u)) \det(I + D^2 c^*(-\nabla u) D^2 u).$$

Equation (2.2) can be understood in the a.e. sense (where the gradient is an L^∞ function), in the sense of Alexandroff, or in the viscosity setting. The book of Gutiérrez [22] gives a good introduction to these concepts in the quadratic case. Let us give here the necessary definitions for general strictly convex cost function (cf. Gutiérrez-Nguyen [24]).

Let u be a function on Ω . We define the (*global*) c -subdifferential of u at x_0 by

$$\partial_c u(x_0) := \{y \in \mathbb{R}^n : u(x) \geq u(x_0) - c(x - y) + c(x_0 - y) \quad \text{for all } x \in \Omega\} \tag{2.3}$$

and the (*local*) c -subdifferential of u at x_0 as

$$\partial_c^{\text{loc}} u(x_0) := \{p \in \mathbb{R}^n : u(x) \geq u(x_0) - c(x - p) + c(x_0 - p) \text{ in some neighborhood of } x_0\}. \tag{2.4}$$

For each $u \in \mathcal{C}(\bar{\Omega})$ satisfying $\partial_c u(\Omega) \subset \bar{\Omega}'$, the *generalized Monge-Ampère measure* of u associated to the cost function c and the weight g is the Borel measure defined by

$$w_c(g, u)(E) = \int_{\partial_c u(E)} g(y) dy \quad \text{for all Borel sets } E \subset \Omega.$$

If, in addition $u \in \mathcal{C}^2(\Omega)$ and $c^* \in \mathcal{C}^2(\mathbb{R}^n)$, then we can give an alternate representation

$$w_c(g, u)(E) = \int_{E \cap \Gamma_u} g(x - \nabla c^*(-\nabla u)) \det(I + D^2 c^*(-\nabla u) D^2 u) dx \tag{2.5}$$

for $\Gamma_u = \{x \in \Omega : \partial_c u(x) \neq \emptyset\}$.

On the other hand, a function $u : \Omega \rightarrow \mathbb{R} \cup \{+\infty\}$ is called c -convex in Ω if there exists $A \subset \mathbb{R}^n \times \mathbb{R}$ such that

$$u(x) = \sup_{(y, a) \in A} \{-c(x - y) + a\}$$

for all $x \in \Omega$.

With these two ingredients in mind, we say that a c -convex function $u \in \mathcal{C}(\bar{\Omega})$ is a solution of equation (2.2) in the sense of Alexandroff if $\partial_c u(\Omega) \subset \bar{\Omega}'$ and $w_c(g, u)(E) = \int_E f$ for every Borel set $E \subset \Omega$. It can be shown that a c -convex function u with $\partial_c u(\Omega) \subset \bar{\Omega}'$ is a solution

in the sense of Alexandroff if and only if it satisfies (2.2) a.e. in Ω and $w_c(g, u)$ is absolutely continuous with respect to the Lebesgue measure.

However, one of the main difficulties in this problem is that the restriction to a subdomain of an Alexandroff solution may not be an Alexandroff solution any longer which makes it extremely difficult to study regularity of such solutions. This is so because the global c -subdifferential (2.3) may not agree with the local one (2.4).

We need one more concept (see Ma-Trudinger-Wang [27]): a set $E' \subset \mathbb{R}^n$ is c^* -convex relative to a set $E \subset \mathbb{R}^n$ if for each $x \in E$, the image $\{\nabla_x c(x - y) : y \in E'\}$ is convex. They proved that if u is a generalized solution of (2.2) in the L^∞ sense, $f > 0$ in Ω , and Ω' is c^* -convex relative to Ω , then u is also a solution in the sense of Alexandroff.

2.2 “Almost” quadratic costs

In the following, we set up the proof for the first statement of Theorem 1.1. Let Ω be a bounded domain. We will assume that $\Omega' = \lambda \mathbf{a} + \Omega^*$ with Ω^* is a bounded domain and \mathbf{a} is a unit vector. Thus, Ω' is nothing but Ω^* translated by a distance $\lambda > 0$ in the direction of the vector \mathbf{a} . We will rewrite the transport equation when $\lambda \geq \lambda_0$ for λ_0 big enough in the following proposition, which plays a key role in our approach. Before stating and proving the result, we need to introduce some notations. As the matrix $I + (p - 2)\mathbf{a} \otimes \mathbf{a}$ is symmetric and positive definite, there is an invertible matrix A such that $I + (p - 2)\mathbf{a} \otimes \mathbf{a} = A^T A$. Also if E is a subset in \mathbb{R}^n , we shall use $A(E)$ to denote the set $\{Ax : x \in E\}$.

Proposition 2.1. *There exists $\lambda_0 > 0$ depending only on $\text{diam}(\Omega \cup \Omega^*)$ such that for all $\lambda \geq \lambda_0$, we have: $T(x)$ is the optimal map for the transport problem between (Ω, f) and $(\lambda \mathbf{a} + \Omega^*, g)$ with cost $c(z) = \frac{1}{p}|z|^p$ if and only if $\tilde{T}(x) := A[T(A^{-1}x) - \lambda \mathbf{a}]$ is the optimal map for the transport problem between $(A(\Omega), \tilde{f})$ and $(A(\Omega^*), \tilde{g})$ for the convex cost*

$$c_\lambda(z) = \frac{1}{2}|z|^2 + \frac{1}{\lambda}E(z), \quad (2.6)$$

where the term E is a smooth function satisfying the following

$$\begin{aligned} E(z) &= O(|z|^3), \quad \nabla E(z) = O(|z|^2), \quad D^2 E(z) = O(|z|), \\ E(Ax - Ay) &= \lambda^3 \left\{ \frac{1}{p} |\mathbf{a} - \frac{x-y}{\lambda}|^p - \frac{1}{p} + \langle \mathbf{a}, \frac{x-y}{\lambda} \rangle - \frac{1}{2} \left| \frac{Ax-Ay}{\lambda} \right|^2 \right\} \quad \forall x \in \Omega, \forall y \in \Omega^*. \end{aligned} \quad (2.7)$$

Here we have set $\tilde{f}(x) = f(A^{-1}x)$ and $\tilde{g}(y) = g(\lambda \mathbf{a} + A^{-1}y)$.

Proof. We first observe that a Borel map $y : \Omega \rightarrow \lambda \mathbf{a} + \Omega^*$ satisfies $y_\# f dx = g dx$ if and only if $\tilde{y} : A(\Omega) \rightarrow A(\Omega^*)$ satisfies $\tilde{y}_\# \tilde{f} dx = \tilde{g} dx$, where $\tilde{y}(x) := A[y(A^{-1}x) - \lambda \mathbf{a}]$. Moreover, by the definition of \tilde{f} , we have

$$\begin{aligned} \int_\Omega \frac{1}{p} |x - y(x)|^p f(x) dx &= \int_\Omega \frac{1}{p} |x - \lambda \mathbf{a} - A^{-1} \tilde{y}(Ax)|^p \tilde{f}(Ax) dx \\ &= \frac{\lambda^{p-2}}{\det A} \int_{A(\Omega)} \frac{\lambda^2}{p} \left| \mathbf{a} - \frac{A^{-1}(x - \tilde{y}(x))}{\lambda} \right|^p \tilde{f}(x) dx. \end{aligned} \quad (2.8)$$

We now want to expand the integrand in the above integral. Note that $|A^{-1}(x - \tilde{y}(x))| \leq c := \text{diam}(\Omega \cup \Omega^*)$ and hence $\left| \frac{A^{-1}(x - \tilde{y}(x))}{\lambda} \right| \leq \frac{c}{\lambda} \ll 1$ for λ big enough. Let $w = \frac{A^{-1}(x - \tilde{y}(x))}{\lambda}$

and consider the Taylor expansion of

$$h(w) = \frac{1}{p} |\mathbf{a} - w|^p$$

around the origin:

$$h(w) = \frac{1}{p} - \langle \mathbf{a}, w \rangle + \frac{1}{2} \langle [I + (p-2)\mathbf{a} \otimes \mathbf{a}] w, w \rangle + F(w) = \frac{1}{p} - \langle \mathbf{a}, w \rangle + \frac{1}{2} \langle A^T A w, w \rangle + F(w),$$

where $F(w) = O(|w|^3)$. Define

$$c_\lambda(z) := \frac{1}{2} |z|^2 + \frac{1}{\lambda} E(z)$$

with $E(z) = \lambda^3 F(\frac{A^{-1}z}{\lambda})$. Then it follows from the above expansion that

$$\begin{aligned} \frac{\lambda^2}{p} \left| \mathbf{a} - \frac{A^{-1}(x - \tilde{y}(x))}{\lambda} \right|^p &= \lambda^2 h \left(\frac{A^{-1}(x - \tilde{y}(x))}{\lambda} \right) = \frac{\lambda^2}{p} - \lambda \langle \mathbf{a}, A^{-1}x \rangle + \lambda \langle \mathbf{a}, A^{-1}\tilde{y}(x) \rangle \\ &\quad + \frac{1}{2} |x - \tilde{y}(x)|^2 + \lambda^2 F \left(\frac{A^{-1}(x - \tilde{y}(x))}{\lambda} \right) \\ &= \frac{\lambda^2}{p} - \lambda \langle \mathbf{a}, A^{-1}x \rangle + \lambda \langle \mathbf{a}, A^{-1}\tilde{y}(x) \rangle + c_\lambda(x - \tilde{y}(x)). \end{aligned}$$

Plug this into (2.8) and by using the fact $\tilde{y}_\# \tilde{f} dx = \tilde{g} dx$, we obtain

$$\begin{aligned} \int_{\Omega} \frac{1}{p} |x - y(x)|^p f(x) dx &= \frac{\lambda^{p-2}}{\det A} \left\{ \int_{A(\Omega)} \left(\frac{\lambda^2}{p} - \lambda \langle \mathbf{a}, A^{-1}x \rangle \right) \tilde{f}(x) dx \right. \\ &\quad \left. + \lambda \int_{A(\Omega)} \langle \mathbf{a}, A^{-1}\tilde{y}(x) \rangle \tilde{f}(x) dx + \int_{A(\Omega)} c_\lambda(x - \tilde{y}(x)) \tilde{f}(x) dx \right\} \\ &= \lambda^{p-1} \left\{ \int_{\Omega} \left(\frac{\lambda}{p} - \langle \mathbf{a}, x \rangle \right) f(x) dx - \int_{\lambda \mathbf{a} + \Omega^*} (\lambda - \langle \mathbf{a}, y \rangle) g(y) dy \right\} \\ &\quad + \frac{\lambda^{p-2}}{\det A} \left\{ \int_{A(\Omega)} c_\lambda(x - \tilde{y}(x)) \tilde{f}(x) dx \right\}. \end{aligned}$$

Since the first term in the last expression is independent of $y(x)$ and $\tilde{y}(x)$, the proposition is proved. Notice that the last identity in (2.7) follows from the above definition of E and the fact that $F(w) = h(w) - \frac{1}{p} + \langle \mathbf{a}, w \rangle - \frac{1}{2} |Aw|^2$. \square

Remark. By Proposition 2.1, it is enough to prove regularity for the potential u in the optimal transportation problem between $A(\Omega)$ and $A(\Omega^*)$ with respect to the convex cost function

$$c(z) = \frac{1}{2} |z|^2 + \frac{1}{\lambda} E(z),$$

where E is a smooth function satisfying (2.7).

Lemma 2.2. *Let λ_0 be as in Proposition 2.1. Then for all $\lambda \geq \lambda_0$ we have: Ω' is c^* -convex with respect to Ω if and only if $A(\Omega^*)$ is c_λ^* -convex with respect to $A(\Omega)$.*

Proof. Let $x \in \Omega$, and consider the set

$$E_x := \{\nabla_x c(x - y) : y \in \Omega'\} = \{\nabla_x c(x - \lambda \mathbf{a} - y) : y \in \Omega^*\}.$$

From the proof of Proposition 2.1, we know that

$$c(x - \lambda \mathbf{a} - y) = \frac{\lambda^p}{p} - \lambda^{p-1} \langle \mathbf{a}, x - y \rangle + \frac{\lambda^{p-2}}{2} \langle A^T A(x - y), x - y \rangle + \lambda^p F\left(\frac{x-y}{\lambda}\right).$$

Therefore we obtain

$$E_x = -\lambda^{p-1} \mathbf{a} + \lambda^{p-2} A^T \left(\left\{ A(x - y) + \lambda(A^T)^{-1} \nabla F\left(\frac{x-y}{\lambda}\right) : y \in \Omega^* \right\} \right).$$

On the other hand, if we let $z = Ax$ and consider the set

$$K_{Ax} = K_z := \{\nabla_z c_\lambda(z - w) : w \in A(\Omega^*)\},$$

then since $c_\lambda(z) = \frac{1}{2} |z|^2 + \lambda^2 F\left(\frac{A^{-1}z}{\lambda}\right)$, we have

$$\begin{aligned} K_{Ax} &= \left\{ z - w + \lambda(A^{-1})^T \nabla F\left(\frac{A^{-1}(z - w)}{\lambda}\right) : w \in A(\Omega^*) \right\} \\ &= \left\{ A(x - y) + \lambda(A^T)^{-1} \nabla F\left(\frac{x-y}{\lambda}\right) : y \in \Omega^* \right\}. \end{aligned}$$

It follows that

$$E_x = -\lambda^{p-1} \mathbf{a} + \lambda^{p-2} A^T (K_{Ax}),$$

and hence the set E_x is convex if and only if the set K_{Ax} is convex. \square

We will also need:

Lemma 2.3. *The potential u in the optimal transportation problem of Theorem 1.1 is a weak solution in the Alexandroff sense, for λ big enough depending on the convexity of Ω' and p . Moreover, $\partial_c^{\text{loc}} u(\Omega) \subset \overline{\Omega'}$.*

Proof. Since Ω' is strongly convex by the assumption, so is Ω^* . Let $\tilde{\rho}$ be a C^2 defining function for Ω^* such that (1.2) holds. Then $\rho(x) := \tilde{\rho}(A^{-1}x)$ is a defining function for $A(\Omega^*)$ satisfying

$$\begin{aligned} \langle D^2 \rho(x) \xi, \xi \rangle &= \langle D^2 \tilde{\rho}(A^{-1}x) A^{-1} \xi, A^{-1} \xi \rangle \\ &\geq C \langle A^{-1} \xi, A^{-1} \xi \rangle = C \langle (A^{-1})^t A^{-1} \xi, \xi \rangle \geq c_p |\xi|^2 \end{aligned}$$

for all $x \in \partial A(\Omega^*)$ and all $\xi \in \mathbb{R}^n$. Thus $A(\Omega^*)$ is also a strongly convex domain. Next we claim that $A(\Omega^*)$ is c_λ^* -convex with respect to $A(\Omega)$ for λ big enough (depending on the convexity of Ω' and p). Indeed, $\nabla c_\lambda(x - y) = x - y + \frac{1}{\lambda} O(|x - y|^2)$ so the image $\{\nabla_x c_\lambda(x - y) : y \in A(\Omega^*)\}$ is the same as $x - U_x$ with $U_x := \{y - \frac{1}{\lambda} O(|x - y|^2) : y \in A(\Omega^*)\}$. Let $F(z)$ be the inverse of the map $y \mapsto y - \frac{1}{\lambda} O(|x - y|^2)$. Then $\bar{\rho}(z) := \rho(F(z))$ is a defining function for U_x . Since $F(z) = z + \frac{1}{\lambda} O(|x - z|^2)$, $DF(z) = I + \frac{1}{\lambda} O(|x - z|)$, $D^2 F_k(z) = \frac{1}{\lambda} O(1)$ and

$$\langle D^2 \bar{\rho}(z) \xi, \xi \rangle = \langle D^2 \rho(F(z)) DF(z) \xi, DF(z) \xi \rangle + \sum_{k=1}^n \frac{\partial \rho}{\partial y_k}(F(z)) \langle D^2 F_k(z) \xi, \xi \rangle,$$

we infer that for all $\lambda > 0$ sufficiently large

$$\langle D^2 \bar{\rho}(z) \xi, \xi \rangle \geq c_p \langle DF(z)^t DF(z) \xi, \xi \rangle - \frac{C}{\lambda} |\xi|^2 \geq \left(\frac{c_p}{2} - \frac{C}{\lambda} \right) |\xi|^2 \geq 0 \quad \forall z \in \partial U_x, \forall \xi \in \mathbb{R}^n.$$

Therefore the connected set U_x is convex and the claim is proved. Consequently, Ω' is c^* -convex with respect to Ω by Lemma 2.2. It then follows from [27] that $\partial_c^{\text{loc}} u(\Omega) \subset \overline{\Omega'}$ and u is an Alexandroff solution. \square

Let us summarize our problem: $u \in \mathcal{C}(\bar{\Omega})$ is a weak solution in the sense of Alexandroff of the equation

$$\begin{cases} g(x - \nabla c^*(-\nabla u)) \det(I + D^2 c^*(-\nabla u) D^2 u) = f & \text{in } \Omega, \\ \partial_c u(\Omega) = \Omega', \end{cases} \quad (2.9)$$

for the cost $c = c_\lambda$, when λ is big enough. This PDE has a very precise structure. Indeed, consider the operator

$$M_\lambda[u] = g(x - \nabla c^*(-\nabla u)) \det(I + D^2 c^*(-\nabla u) D^2 u).$$

Because of the expression for the cost (2.6), it can be rewritten as

$$M_\lambda[u] = g(x + \nabla u + G(\nabla u)) \det(I + D^2 u + H(\nabla u)(D^2 u)), \quad (2.10)$$

where G and H in particular satisfy

$$\|H(\nabla u)\|_{L^\infty} \leq \frac{C}{\lambda} \|\nabla u\|_{L^\infty(\Omega)}, \quad \|G(\nabla u)\|_{L^\infty} \leq \frac{C}{\lambda} \|\nabla u\|_{L^\infty(\Omega)}^2. \quad (2.11)$$

We will need one more important concept for a c -convex function $u \in \mathcal{C}(\bar{\Omega})$: the notion of *sections*. Let $y_0 \in \partial_c u(\Omega)$. Define $\bar{u}(x) := u(x) + c(x - y_0)$ and let $x_0 \in \Omega$ be a minimum point of \bar{u} in Ω . We note that this choice of x_0 is equivalent to assuming that $y_0 \in \partial_c u(x_0)$ since

$$u(x) + c(x - y_0) \geq u(x_0) + c(x_0 - y_0) \text{ for all } x \in \Omega.$$

In particular, if u is differentiable at x_0 then $y_0 := x_0 - \nabla c^*(-\nabla u(x_0))$. For any $\mu > 0$ define the μ -section of u at x_0 as

$$S_\mu(u, x_0, y_0) := \{x \in \Omega : u(x) \leq u(x_0) - c(x - y_0) + c(x_0 - y_0) + \mu\}. \quad (2.12)$$

In order to simplify the notation, we simply write $S_\mu(u, x_0)$ for the section, while the dependence on y_0 is made implicitly. With some abuse of notation, this section can be rewritten as

$$S_\mu(\bar{u}, x_0) = \{x \in \Omega : \bar{u}(x) \leq \bar{u}(x_0) + \mu\}. \quad (2.13)$$

We would like to study its geometrical properties when $\mu \rightarrow 0$.

In our last remark here, we provide the main idea for the proof of the second statement of Theorem 1.1. When $p > 2$, $p \rightarrow 2$, we have the following estimates for the cost $c(z) = \frac{1}{p} |z|^p$. For simplicity, assume that $1 < \text{dist}(\Omega, \Omega') < C$. Then

$$\begin{aligned} \frac{1}{p} |z|^2 &\leq c(z) \leq \frac{1}{p} |z|^2 \left[1 + (p-2) (\log |z|)^{p-1}\right], \\ \frac{2}{p} z_i &\leq \nabla_i c(z) \leq \frac{2}{p} z_i \left[1 + (p-2) (\log |z|)^{p-1}\right] \end{aligned}$$

and

$$\frac{2}{p} \delta_{ij} + (p-2) \left[|z|^{p-4} z_i z_j\right] \leq D_{ij}^2 c(z) \leq \frac{2}{p} \delta_{ij} + (p-2) \left[\delta_{ij} (\log |z|)^{p-1} + |z|^{p-4} z_i z_j\right].$$

In particular, there exists $p_0 > 2$, close to 2, and depending only on the distance between Ω and Ω' , such that we are in the same situation as in the first statement of Theorem 1.1.

3 On the geometry of the equation

In this section we summarize some miscellaneous geometric properties that will be needed for the proof of the main theorem, in particular, a precise understanding of the almost convexity of the cost.

3.1 Geometric interpretation of the cost c_λ - almost convexity

Let $u \in \mathcal{C}(\bar{\Omega})$ be a c -convex function in Ω . Fix $x_0 \in \Omega$ and $y_0 \in \partial_c u(x_0)$. Then

$$u(x) \geq -c(x - y_0) + c(x_0 - y_0) + u(x_0) \quad \text{for all } x \in \Omega. \quad (3.1)$$

If $c = c_q$ is precisely the quadratic cost, then (3.1) reduces to

$$u(x) \geq -\frac{1}{2}|x|^2 + \langle x - x_0, y_0 \rangle + \frac{1}{2}|x_0|^2 + u(x_0) \quad \text{for all } x \in \Omega, \quad (3.2)$$

i.e., $\bar{u}(x) := u(x) + \frac{1}{2}|x|^2$ is a convex function. If $c = c_\lambda$ as defined in (2.6), then (3.1) reads

$$u(x) \geq -\frac{1}{2}|x|^2 + \langle x - x_0, p_0 \rangle - \frac{1}{2\lambda} \langle D^2 E(\zeta - y_0)(x - x_0), x - x_0 \rangle + \frac{1}{2}|x_0|^2 + u(x_0), \quad (3.3)$$

where ζ belongs to the segment $[x, x_0]$ and $p_0 := y_0 - \frac{1}{\lambda} \nabla E(x_0 - y_0)$. Essentially, it is almost quadratic as in (3.2) plus a small perturbation that goes to zero as $\lambda \rightarrow \infty$.

Next, we fix $x_0 \in \Omega$, $y_0 \in \partial_c u(x_0)$. Take any other pair $\tilde{x} \in \Omega$, $\tilde{y} \in \partial_c u(\tilde{x})$. Then there is a supporting cost that touches u from below at \tilde{x} , and it is written as

$$z = u(\tilde{x}) - c(x - \tilde{y}) + c(\tilde{x} - \tilde{y}).$$

Now we let

$$\bar{u}(x) := u(x) + c(x - y_0). \quad (3.4)$$

The supporting cost at \tilde{x} for \bar{u} is given by

$$\begin{aligned} z &= \bar{u}(\tilde{x}) - c(x - \tilde{y}) + c(\tilde{x} - \tilde{y}) - c(\tilde{x} - y_0) + c(x - y_0) \\ &= \bar{u}(\tilde{x}) + \langle \tilde{p}, x - \tilde{x} \rangle + \langle \tilde{Q}(x - \tilde{x}), x - \tilde{x} \rangle \end{aligned} \quad (3.5)$$

where $\tilde{p} := \nabla c(\tilde{x} - y_0) - \nabla c(\tilde{x} - \tilde{y})$ and $\tilde{Q} = \frac{1}{2} [D^2 c(\xi - y_0) - D^2 c(\xi - \tilde{y})]$ for some $\xi \in [x, \tilde{x}]$.

Since our cost is of the form $c(z) = \frac{1}{2}|z|^2 + \frac{1}{\lambda} E(z)$, an estimate for \tilde{Q} is

$$\|\tilde{Q}\| = \frac{1}{2\lambda} \|D^2 E(\xi - y_0) - D^2 E(\xi - \tilde{y})\| \leq \frac{C}{\lambda} |\tilde{y} - y_0|. \quad (3.6)$$

On the other hand, by its construction

$$\tilde{p} = \nabla c(\tilde{x} - y_0) - \nabla c(\tilde{x} - \tilde{y}) = \tilde{y} - y_0 + \frac{1}{\lambda} [\nabla E(\tilde{x} - y_0) - \nabla E(\tilde{x} - \tilde{y})],$$

i.e.,

$$\tilde{p} = (\tilde{y} - y_0) \left(1 + O\left(\frac{C}{\lambda}\right)\right), \quad (3.7)$$

where the constants in both (3.6) and (3.7) depend on $C = C(y_0, \tilde{y})$.

If the cost is quadratic, then $\tilde{Q} \equiv 0$, and the supporting cost for \bar{u} is just a hyperplane; this is the standard construction for convex functions. We have proved that:

Lemma 3.1. *Let $u \in \mathcal{C}(\bar{\Omega})$ be c -convex and $y_0 \in \partial_c u(x_0)$ for some $x_0 \in \Omega$. Then the function \bar{u} defined by (3.4) is almost a convex function, precisely, at any point $\tilde{x} \in \Omega$, there exists (at least) one supporting almost-hyperplane of the form*

$$z = \bar{u}(\tilde{x}) + \langle \tilde{p}, x - \tilde{x} \rangle + \langle \tilde{Q}(x - \tilde{x}), x - \tilde{x} \rangle,$$

where \tilde{p} and \tilde{Q} are given by (3.7) and (3.6), respectively.

In the light of the previous arguments, it is expected that a section S of u , as defined in (2.12), is not going to be too far from its convex hull. We will use the notation $[S]$ for the convex hull of a set S .

Proposition 3.2. *Fix the cost $c = c_\lambda$. Suppose that $u \in \mathcal{C}(\bar{\Omega})$ is c -convex satisfying $\partial_c u(\Omega) \subset \Omega'$. Let $x_1 \in \Omega$, $y_1 \in \partial_c u(x_1)$, $\mu > 0$, and consider the section $S := S_\mu(u, x_1, y_1)$ as defined in (2.12). Assume that $S \Subset \Omega$. Then there exists a constant $C > 0$ such that*

$$\text{dist}(x, \partial[S]) \leq \frac{C}{\lambda} \text{ for all } x \in \partial S. \quad (3.8)$$

The constant C depends on $\partial_c u(\Omega)$, but not on λ .

Proof. Let $x_0 \in \partial S \subset \Omega$, and $y_0 \in \partial_c u(x_0)$. Then we have for all x in Ω ,

$$\begin{aligned} u(x) &\geq u(x_0) - c(x - y_0) + c(x_0 - y_0) \\ &= u(x_1) - c(x_0 - y_1) + c(x_1 - y_1) + \mu - c(x - y_0) + c(x_0 - y_0). \end{aligned}$$

Thus it follows from the definition of S that for all $x \in S$,

$$\begin{aligned} 0 &\geq c(x - y_1) - c(x_0 - y_1) + c(x_0 - y_0) - c(x - y_0) \\ &= \langle \nabla c(x_0 - y_1) - \nabla c(x_0 - y_0), x - x_0 \rangle + \frac{1}{2} \langle [D^2 c(\zeta - y_1) - D^2 c(\zeta - y_0)](x - x_0), x - x_0 \rangle \end{aligned}$$

for some ζ in the segment joining x to x_0 . Or

$$0 \geq \langle p, x - x_0 \rangle + \frac{1}{2} \langle Q(x - x_0), x - x_0 \rangle \text{ for all } x \in S,$$

where $p = \nabla c(x_0 - y_1) - \nabla c(x_0 - y_0)$ and $Q = D^2 c(\zeta - y_1) - D^2 c(\zeta - y_0)$. Since our cost is of the form $c(z) = \frac{1}{2}|z|^2 + \frac{1}{\lambda}E(z)$, an estimate for Q is

$$\|Q\| = \frac{1}{\lambda} \|D^2 E(\zeta - y_1) - D^2 E(\zeta - y_0)\| \leq \frac{C}{\lambda} |y_0 - y_1|.$$

Therefore, S is the intersection of almost-halfplanes,

$$0 \geq \langle p, x - x_0 \rangle + \frac{1}{2} \langle Q(x - x_0), x - x_0 \rangle.$$

The section S would be convex if $Q \equiv 0$ for all $x_0 \in \partial S$. In our case, these supporting functions are planes plus a small quadratic perturbation; we would like to measure the distance of how far is S from $[S]$. If S is not convex, for $x_0 \in \partial S$, we can find $x \in S$ such that x is outside the half plane

$$H_{x_0} : 0 \geq \langle \frac{p}{|p|}, x - x_0 \rangle.$$

Then if we let $d = \text{dist}(x, H_{x_0})$ then $d = \langle \frac{p}{|p|}, x - x_0 \rangle > 0$. We also have

$$\langle p, x - x_0 \rangle \leq -\frac{1}{2} \langle Q(x - x_0), x - x_0 \rangle.$$

Hence by using the estimate of Q ,

$$d \leq \frac{C}{\lambda} \frac{|y_0 - y_1|}{|p|} |x - x_0|^2 \leq \frac{C}{\lambda} \frac{|y_0 - y_1|}{|p|}.$$

On the other hand, by its construction

$$p = \nabla c(x_0 - y_1) - \nabla c(x_0 - y_0) = y_0 - y_1 + \frac{1}{\lambda} [\nabla E(x_0 - y_1) - \nabla E(x_0 - y_0)].$$

Thus, $(1 - \frac{C}{\lambda})|y_0 - y_1| \leq |p| \leq (1 + \frac{C}{\lambda})|y_0 - y_1|$ and by combining with the above estimate we obtain $d \leq \frac{C}{\lambda}$. We have in fact shown for each $x_0 \in \partial S$,

$$\text{dist}(x, H_{x_0}) \leq \frac{C}{\lambda} \text{ for all } x \in S.$$

It follows immediately from this that

$$\text{dist}(x_0, \partial[S]) \leq \frac{C}{\lambda}.$$

□

3.2 On the local c -subdifferential

In Lemma 3.1 we have given a characterization of the global c -subdifferentials of u . It is well known that this concept may be different from the local c -subdifferential, and this is one of the main difficulties in the study of the regularity for optimal transportation. More precisely, the restriction of an Alexandroff solution of the transport problem to a smaller domain may not be an Alexandroff solution any longer because extra mass may appear (*global* \neq *local*). We give here the first attempt to understand the behavior of the subdifferential under restriction to a smaller domain. In particular, we show that the characterization of Lemma 3.1 on “almost”-convexity is still true for (local) c -subgradients.

We have defined in Section 2 the (*global*) c -subdifferential of u at \tilde{x} by

$$\partial_c u(\tilde{x}) := \{\tilde{y} \in \mathbb{R}^n : u(x) \geq u(\tilde{x}) - c(x - \tilde{y}) + c(\tilde{x} - \tilde{y}) \text{ for all } x \in \Omega\}$$

and the (*local*) c -subdifferential of u at \tilde{x} as

$$\partial_c^{\text{loc}} u(\tilde{x}) := \{p \in \mathbb{R}^n : u(x) \geq u(\tilde{x}) - c(x - p) + c(\tilde{x} - p) \text{ in some neighborhood of } \tilde{x}\}.$$

More generally for an open set $V \subset \Omega$ and $\tilde{x} \in V$, we write

$$\partial_c(u, V)(\tilde{x}) := \{q \in \mathbb{R}^n : u(x) \geq u(\tilde{x}) - c(x - q) + c(\tilde{x} - q) \text{ for all } x \in V\}.$$

It is clear that $\partial_c^{\text{loc}} u(\tilde{x}) \supset \partial_c u(\tilde{x})$, but in general they are not equal.

We will also need to work with the following set of subgradients and c -subgradients. For $\tilde{x} \in \Omega$, define

$$\begin{aligned} \partial^- u(\tilde{x}) &:= \{p \in \mathbb{R}^n : u(x) \geq u(\tilde{x}) + \langle p, x - \tilde{x} \rangle + o(|x - \tilde{x}|) \text{ for } x \text{ near } \tilde{x}\}, \\ \partial_c^- u(\tilde{x}) &:= \{\tilde{y} \in \mathbb{R}^n : u(x) \geq u(\tilde{x}) - c(x - \tilde{y}) + c(\tilde{x} - \tilde{y}) + o(|x - \tilde{x}|) \text{ for } x \text{ near } \tilde{x}\}. \end{aligned}$$

It follows that

$$\partial^- u(\tilde{x}) = -\nabla c(\tilde{x} - \partial_c^- u(\tilde{x})). \quad (3.9)$$

The above mappings can be extended to boundary points as follows. Let $\tilde{x} \in \partial\Omega$, we denote $\partial^-u(\tilde{x}) := \{p \in \mathbb{R}^n : p = \lim_{k \rightarrow \infty} p_k\}$, where $p_k \in \partial^-u(x_k)$ and $\{x_k\}$ is a sequence of interior points of Ω converging to \tilde{x} , and let $\partial_c^-u(\tilde{x})$ be given by (3.9).

First we present two well known lemmas that relate the c -subdifferentials of two functions:

Lemma 3.3. *Let $u, v \in \mathcal{C}(\overline{\Omega})$ and $E \subset \Omega$ be an open set such that*

$$\begin{cases} u = v & \text{on } \partial E, \\ u \geq v & \text{in } E. \end{cases}$$

Then $\partial_c(u, E)(E) \subset \partial_c(v, E)(E)$. Consequently, $\partial_c u(E) \subset \partial_c^{loc} v(E)$.

Proof. This result can be found in [24, Lemma 5.1] but we include the proof here for convenience. Let $p \in \partial_c(u, E)(E)$. Then there exists $x_0 \in E$ such that $u(x) \geq u(x_0) - c(x - p) + c(x_0 - p)$ for all $x \in E$. Let

$$a := \sup_{x \in \bar{E}} \{u(x_0) - c(x - p) + c(x_0 - p) - v(x)\}.$$

Since $u(x_0) \geq v(x_0)$, we get $a \geq 0$. Also there is $x_1 \in \bar{E}$ such that $a = u(x_0) - c(x_1 - p) + c(x_0 - p) - v(x_1)$. Thus $v(x) \geq v(x_1) - c(x - p) + c(x_1 - p)$ for all $x \in E$. Then two things can happen:

If $a > 0$, then since $u(x_1) \geq a + v(x_1)$ we must have $x_1 \notin \partial E$ and hence, $p \in \partial_c(v, E)(x_1) \subset \partial_c(v, E)(E)$. On the other hand if $a = 0$, then we already have that $v(x) \geq v(x_0) - c(x - p) + c(x_0 - p)$ in E which in turn gives $p \in \partial_c(v, E)(x_0) \subset \partial_c(v, E)(E)$. \square

For the rest of the paper, $\mathcal{N}_\nu(E)$ will denote the ν -neighborhood of the set E , i.e.,

$$\mathcal{N}_\nu(E) := \{x \in \mathbb{R}^n : \text{dist}(x, E) \leq \nu\}.$$

Lemma 3.4. *Let $u, v \in \mathcal{C}(\overline{B_R(\bar{x})})$ and suppose that u is c_λ -convex and v is c_q -convex. There exists a universal constant $C > 0$ such that if $\|u - v\|_{L^\infty(B_R(\bar{x}))} \leq \epsilon$, then*

$$\partial_c u(B_{R-\frac{\epsilon}{\delta}}(\bar{x})) \subset \mathcal{N}_{(4\delta+\frac{C}{\lambda})} \{\partial_{c_q} v(B_R(\bar{x}))\} \quad \text{for all } \delta > \frac{\epsilon}{R}. \quad (3.10)$$

Consequently,

$$\partial_c^{loc} u(B_{R-\frac{\epsilon}{\delta}}(\bar{x})) \subset \mathcal{N}_{(4\delta+\frac{C}{\lambda})} \{[\partial_{c_q} v(B_R(\bar{x}))]\}, \quad (3.11)$$

where $[\partial_{c_q} v(B_R(\bar{x}))]$ is the convex hull of $\partial_{c_q} v(B_R(\bar{x}))$.

Proof. We note that a version of this result for quadratic cost appears in [11] (Lemma 2 of part 4). Let $p \in \partial_c u(B_{R-\frac{\epsilon}{\delta}}(\bar{x}))$, then there exists $x_0 \in B_{R-\frac{\epsilon}{\delta}}(\bar{x})$ such that

$$u(x) \geq u(x_0) - c(x - p) + c(x_0 - p) \quad \text{for all } x \in B_R := B_R(\bar{x}).$$

Consider $v^*(x) := v(x) + \epsilon + \frac{2\delta}{R}(|x - \bar{x}|^2 - R^2)$, then $v^*|_{\partial B_R} \geq u$ and $v^*|_{B_{R-\frac{\epsilon}{\delta}}(\bar{x})} \leq u$. Define

$$a := \sup_{x \in \bar{B}_R} \{u(x_0) - c(x - p) + c(x_0 - p) - v^*(x)\}.$$

Since $u(x_0) \geq v^*(x_0)$, we get $a \geq 0$. On the other hand, there exists $x_1 \in \overline{B}_R$ such that $a = u(x_0) - c(x_1 - p) + c(x_0 - p) - v^*(x_1)$, and thus $v^*(x) \geq v^*(x_1) - c(x - p) + c(x_1 - p)$ for all $x \in B_R$. Moreover, $u(x_1) \geq u(x_0) - c(x_1 - p) + c(x_0 - p) = v^*(x_1) + a$. Therefore, if $a > 0$, then x_1 does not belong to the boundary of B_R , and we conclude $p \in \partial_c v^*(x_1) \subset \partial_c v^*(B_R)$. If, on the contrary, $a = 0$, then $v^*(x) \geq v^*(x_0) - c(x - p) + c(x_0 - p)$ for all $x \in B_R$, so automatically $p \in \partial_c v^*(x_0) \subset \partial_c v^*(B_R)$. Thus we have shown that $\partial_c u(B_{R-\frac{\epsilon}{8}}(\bar{x})) \subset \partial_c v^*(B_R)$.

Next suppose $y^* \in \partial_c v^*(x_0)$, $x_0 \in B_R$. Then by (3.9), $y^* = x_0 - \nabla c^*(x_0 - p^*)$ for some $p^* \in \partial^- v^*(x_0)$. Define $p := p^* - \frac{4\delta}{R}(x_0 - \bar{x})$ and we claim that $x_0 + p \in \partial \bar{v}(x_0)$, where $\bar{v} := v + \frac{1}{2}|x|^2$ is a convex function and $\partial \bar{v}(x_0) := \{q : \bar{v}(x) \geq \bar{v}(x_0) + \langle q, x - x_0 \rangle \forall x \in B_R\}$. To see this, let us write $p^* = \sum_{i=1}^N t_i p_i^*$ where p_i^* are extremal points of $\partial^- v^*(x_0)$ and $0 \leq t_i \leq 1$ satisfying $\sum_{i=1}^N t_i = 1$. For each i , by a result in [7] we can find a sequence $\{x_k\}$ of differentiable points of v such that $x_k \rightarrow x_0$ and $p_i^* = \lim_{k \rightarrow \infty} \nabla v^*(x_k) = \lim_{k \rightarrow \infty} \nabla v(x_k) + \frac{4\delta}{R}(x_0 - \bar{x})$. This yields $x_0 + p_i^* - \frac{4\delta}{R}(x_0 - \bar{x}) = \lim_{k \rightarrow \infty} \nabla \bar{v}(x_k) \in \partial \bar{v}(x_0)$ for each i . We then obtain the claim since $x_0 + p = \sum_{i=1}^N t_i [x_0 + p_i^* - \frac{4\delta}{R}(x_0 - \bar{x})]$ and $\partial \bar{v}(x_0)$ is a convex set. Observe that $y^* = x_0 + p^* + \frac{1}{\lambda} \nabla E(x_0 - y^*)$ by the definitions of p^* and c_λ . The relation (3.10) now follows from the claim, the fact $\partial \bar{v}(x_0) = \partial_{c_q} v(x_0)$ and

$$|y^* - [x_0 + p]| = \left| x_0 + p^* + \frac{1}{\lambda} \nabla E(x_0 - y^*) - [x_0 + p^* - \frac{4\delta}{R}(x_0 - \bar{x})] \right| \leq 4\delta + \frac{C}{\lambda}.$$

To see (3.11), let $y \in \partial_c^{\text{loc}} u(x)$ for $x \in B_{R-\frac{\epsilon}{8}}(\bar{x})$. Then there exists $p \in \partial^- u(x)$ such that $y = x - \nabla c^*(-p)$. We write $p = \sum_{i=1}^N t_i p_i$ as a finite convex combination of extremal points p_i of $\partial^- u(x_0)$. Since u is c -convex, $y_i := x - \nabla c^*(-p_i) \in \partial_c u(x)$ for all i which implies that

$$\sum_{i=1}^N t_i y_i \in \mathcal{N}_{\left(4\delta + \frac{C}{\lambda}\right)} \left\{ [\partial_{c_q} v(B_R(\bar{x}))] \right\}$$

by (3.10). This yields (3.11) because

$$\left| y - \sum_{i=1}^N t_i y_i \right| = \left| x + p + \frac{1}{\lambda} \nabla E(x - y) - \sum_{i=1}^N t_i \left[x + p_i + \frac{1}{\lambda} \nabla E(x - y_i) \right] \right| \leq \frac{C}{\lambda}.$$

□

Next, we present a characterization of subgradients. Although the so called “double mountain above sliding mountain” lemma is not true anymore without assuming the (A.3w) condition, we still have a useful characterization using the asymptotic behavior of the cost $c := c_\lambda$, that generalizes Lemma 3.1.

Proposition 3.5. *Suppose that $u \in \mathcal{C}(\overline{\Omega})$ is a c -convex function in Ω . Then for $\tilde{x} \in \overline{\Omega}$, $\tilde{p} \in \partial^- u(\tilde{x})$, we have*

$$u(x) + \frac{1}{2}|x|^2 \geq u(\tilde{x}) + \frac{1}{2}|\tilde{x}|^2 + \langle \tilde{p} + \tilde{x}, -\tilde{x} \rangle + \frac{1}{\lambda} O(|x - \tilde{x}|^2) \quad \text{for all } x \in \Omega.$$

Consequently if $\tilde{y} \in \partial_c^- u(\tilde{x})$ with $\tilde{x} \in \overline{\Omega}$, then

$$u(x) \geq u(\tilde{x}) - c(x - \tilde{y}) + c(\tilde{x} - \tilde{y}) + \frac{1}{\lambda} O(|x - \tilde{x}|^2) \quad \text{for all } x \in \Omega.$$

Proof. First assume that $\tilde{x} \in \Omega$. By construction, $\partial^-u(\tilde{x})$ is a closed convex set and hence we can write \tilde{p} as a convex combination of finite extremal points, that is, $\tilde{p} = \sum_{i=1}^N t_i q_i$, where $t_i \geq 0$, $\sum_{i=1}^N t_i = 1$ and q_i are extremal points in $\partial^-u(\tilde{x})$.

For each i , let $\bar{y}_i \in \mathbb{R}^n$ be such that $q_i = -\nabla c(\tilde{x} - \bar{y}_i)$. Since $q_i = \lim_{m \rightarrow \infty} \nabla u(z_m)$ for a sequence $z_m \rightarrow \tilde{x}$ which are points of differentiability of u , it follows that $\bar{y}_i \in \partial_c u(\tilde{x})$. Then, for each \bar{y}_i , using the fact $c(z) = \frac{1}{2}|z|^2 + \frac{1}{\lambda}E(z)$ given by (2.6), we have that

$$u(x) + \frac{1}{2}|x|^2 \geq u(\tilde{x}) + \frac{1}{2}|\tilde{x}|^2 + \langle \bar{y}_i, x - \tilde{x} \rangle + \frac{1}{\lambda} [E(\tilde{x} - \bar{y}_i) - E(x - \bar{y}_i)] \quad \text{for all } x \in \Omega,$$

and consequently, as $\bar{y}_i = q_i + \tilde{x} + \frac{1}{\lambda} \nabla E(\tilde{x} - \bar{y}_i)$ by our choice of \bar{y}_i ,

$$\begin{aligned} & u(x) + \frac{1}{2}|x|^2 \\ & \geq u(\tilde{x}) + \frac{1}{2}|\tilde{x}|^2 + \langle q_i + \tilde{x}, x - \tilde{x} \rangle - \frac{1}{\lambda} \left[E(x - \bar{y}_i) - E(\tilde{x} - \bar{y}_i) - \langle \nabla E(\tilde{x} - \bar{y}_i), x - \tilde{x} \rangle \right] \\ & = u(\tilde{x}) + \frac{1}{2}|\tilde{x}|^2 + \langle q_i + \tilde{x}, x - \tilde{x} \rangle + \frac{1}{\lambda} O(|x - \tilde{x}|^2) \quad \text{for all } x \in \Omega. \end{aligned}$$

Therefore if we take the convex combination with respect to i , we obtain

$$u(x) + \frac{1}{2}|x|^2 \geq u(\tilde{x}) + \frac{1}{2}|\tilde{x}|^2 + \langle \tilde{p} + \tilde{x}, x - \tilde{x} \rangle + \frac{1}{\lambda} O(|x - \tilde{x}|^2) \quad \text{for all } x \in \Omega. \quad (3.12)$$

Now assume that $\tilde{x} \in \partial\Omega$. Then by the definition of subdifferential at boundary points, we have $\tilde{p} = \lim_{k \rightarrow \infty} p_k$ where $p_k \in \partial^-u(x_k)$ and $\{x_k\}$ is a sequence of interior points of Ω converging to \tilde{x} . The above proof shows that (3.12) holds if \tilde{x} and \tilde{p} are replaced by x_k and p_k , and therefore by letting k tend to infinity we see that (3.12) is still true when $\tilde{x} \in \partial\Omega$. Thus the first statement of the lemma follows.

To prove the second statement, let $\tilde{p} := -\nabla c(\tilde{x} - \tilde{y})$. Then we obtain (3.12) since $\tilde{p} \in \partial^-u(\tilde{x})$ by the assumption $\tilde{y} \in \partial_c^-u(\tilde{x})$. It follows from (3.12) and by a similar argument to the one used above that

$$u(x) \geq u(\tilde{x}) - c(x - \tilde{y}) + c(\tilde{x} - \tilde{y}) + \frac{1}{\lambda} O(|x - \tilde{x}|^2) \quad \text{for all } x \in \Omega.$$

□

3.3 An Alexandroff type estimate

Although it is not needed in the proof of our main result, we believe that it would be interesting to include here an Alexandroff type estimate for the cost original $c(x) = \frac{1}{p}|x|^p$ analogous to the one in Monge-Ampère setting [22, Theorem 1.4.2].

Proposition 3.6. *Suppose $c(x) = \frac{1}{p}|x|^p$, $1 < p < 2 + \frac{1}{n-1}$. Let $\Omega \subset \mathbb{R}^n$ be a bounded open convex set and $u \in C(\bar{\Omega})$ be a c -convex function on Ω satisfying $u \geq 0$ on $\partial\Omega$. We have:*

(i) *If $1 < p \leq 2$ then*

$$u(x) \geq -C(n, p) \left[\text{diam}(\Omega)^{\frac{n-(p-1)}{p-1}} \text{dist}(x, \partial\Omega) |\partial_c u(\Omega)| \right]^{\frac{p-1}{n}} \quad \forall x \in \Omega.$$

(ii) If $2 \leq p < 2 + \frac{1}{n-1}$ then

$$u(x) \geq -C(n, p) \left[\text{diam}(\Omega)^{n-1} \text{dist}(x, \partial\Omega)^{\frac{(p-1)-n(p-2)}{p-1}} |\partial_c u(\Omega)| \right]^{\frac{p-1}{n}} \quad \forall x \in \Omega,$$

where $C(n, p)$ is a constant depending only on the dimension n and p .

Proof. Let $x_0 \in \Omega$ be such that $u(x_0) < 0$. In order to prove the theorem we first define a c -convex function, whose graph is the upside-down c -cone surface with vertex at $(x_0, u(x_0))$, as follows

$$w(x) = \sup_{v \in \mathcal{F}} v(x) \quad \forall x \in \bar{\Omega},$$

where

$$\mathcal{F} = \left\{ v \in \mathcal{C}(\bar{\Omega}) : v \text{ is } c\text{-convex, } v(x_0) \leq u(x_0) \text{ and } v|_{\partial\Omega} \leq u|_{\partial\Omega} \right\}.$$

Then since $u \in \mathcal{F}$ we have $w \geq u$ on $\bar{\Omega}$, $w(x_0) = u(x_0)$, w is c -convex on Ω and $w|_{\partial\Omega} = u|_{\partial\Omega}$. We now claim that $w \in \mathcal{C}(\bar{\Omega})$. Indeed, since Ω is convex it is q -regular with q is the conjugate of p . Hence, there exists $h \in W^{1,q}(\Omega) \cap \mathcal{C}(\bar{\Omega})$ weak solution to the q -Laplacian

$$-\text{div}(|\nabla h(x)|^{q-2} \nabla h(x)) + n = 0 \quad \text{in } \Omega, \text{ and } h = -u \text{ on } \partial\Omega.$$

Notice that $\text{div}(|\nabla h(x)|^{q-2} \nabla h(x)) = \text{div}(\nabla c^*(\nabla h(x)))$. For each $f(x) = -c(x-y) - \lambda \in \mathcal{F}$, we have $-f(x) \geq -u(x)$ on $\partial\Omega$, and $-\text{div}(\nabla c^*(-\nabla f(x))) + n = 0$. Hence by the comparison principle for the q -Laplacian we get that $-f \geq h$ in $\bar{\Omega}$, and therefore

$$w(x) = \sup_{v \in \mathcal{F}} v(x) \leq -h(x) \quad \text{for all } x \in \bar{\Omega}.$$

Thus we obtain $u \leq w \leq -h$ in $\bar{\Omega}$ and hence the claim follows since $u, -h \in \mathcal{C}(\bar{\Omega})$, $u|_{\partial\Omega} = -h|_{\partial\Omega}$ and w is c -convex on Ω .

From the above claim and by Lemma 3.3, we have

$$\begin{aligned} \partial_c u(\Omega) &\supset \partial_c w(\Omega) \supset \partial_c w(x_0) \\ &= \{y : u(x_0) - c(x-y) + c(x_0-y) \leq u(x) \quad \forall x \in \partial\Omega\} \\ &\supset \{y : u(x_0) - c(x-y) + c(x_0-y) \leq 0 \quad \forall x \in \partial\Omega\} \\ &\supset \{y : u(x_0) + \langle \nabla c(x_0-y), x_0-x \rangle \leq 0 \quad \forall x \in \partial\Omega\} =: E. \end{aligned}$$

Let $x^* \in \partial\Omega$ be the point such that $|x^* - x_0| = \text{dist}(x_0, \partial\Omega)$. Define $G(z) := x_0 - \nabla c^*(-z)$ and

$$z_0 := \frac{-u(x_0)}{\text{dist}(x_0, \partial\Omega)} \frac{x^* - x_0}{|x^* - x_0|}.$$

Then, we claim that

1. $G\left(B\left(0, \frac{-u(x_0)}{\text{diam}(\Omega)}\right)\right) \subset E$.
2. $G(z_0) \in E$.
3. If $G(z_1), G(z_2) \in E$, then $G(tz_1 + (1-t)z_2) \in E$ for all $t \in [0, 1]$.

The first and the third claims can be checked readily. To obtain the second claim, we need to use the fact that

$$\langle x^* - x_0, x - x_0 \rangle \leq |x^* - x_0|^2 \quad \forall x \in \partial\Omega,$$

which follows from the convexity of Ω .

Let H be the convex hull in \mathbb{R}^n of the point z_0 and the ball $B\left(0, \frac{-u(x_0)}{\text{diam}(\Omega)}\right)$. From these three claims we can conclude that $G(H) \subset E$, and thus,

$$|\partial_c u(\Omega)| \geq |E| \geq |G(H)| = \int_H \det D^2 c^*(-y) dy = (q-1) \int_H |y|^{n(q-2)} dy.$$

The last integral above was computed in Gutiérrez-Nguyen [24] (equations (67) and (68)), and the proposition follows from these estimates. \square

4 Initial step

This is the first ingredient in the proof of Theorem 1.1. Let $u \in \mathcal{C}(\overline{\Omega})$ be a solution of the optimal transport problem between Ω and Ω'

$$\begin{cases} M_\lambda[u] = f & \text{in } \Omega, \\ \partial_{c_\lambda} u(\Omega) = \Omega', \end{cases} \quad (4.1)$$

where M_λ is defined in (2.10) and satisfies (2.11), with the hypothesis stated in the main theorem.

The main idea is to compare the solution u of the optimal transport equation (2.9) to w , where w is the solution of the optimal transport problem between Ω and Ω' with quadratic cost $c_q(z) = \frac{1}{2}|z|^2$, and Ω' convex. Note that w is unique up to an additive constant, strictly convex and smooth (see the article by Caffarelli [7], further regularity is contained in [8], [10]). The equation satisfied by w is the following

$$\begin{cases} g(x + \nabla w) \det(I + D^2 w) = f & \text{in } \Omega, \\ \partial_{c_q} w(\Omega) = \Omega'. \end{cases} \quad (4.2)$$

4.1 An approximation result

In this subsection we prove a comparison result between the solution u of (4.1) to the solution w of (4.2):

Proposition 4.1. *For any $\epsilon_0 > 0$, there exists $\lambda_0 > 0$ depending only on ϵ_0 , n , Λ_1 , Λ_2 , Ω and Ω' such that if $\lambda \geq \lambda_0$, f and g satisfy (1.3), $w \in \mathcal{C}(\overline{\Omega})$ is a solution of (4.2) and $u \in \mathcal{C}(\overline{\Omega})$ is a solution of the transport equation (4.1) with $u(x_0) = w(x_0)$ for some $x_0 \in \Omega$, then we have*

$$\|u - w\|_{L^\infty(\Omega)} \leq \epsilon_0.$$

The proof of this proposition is a simple consequence of the following strengthened result where the constant λ_0 is independent of the domains Ω and Ω' , as long as they lie between two universal balls.

Proposition 4.2. *Let Ω and Ω' be two closed subsets of \mathbb{R}^n satisfying $B_1(0) \subset \Omega \subset \overline{B}_n(0)$ and $B_1(0) \subset \Omega' \subset \overline{B}_n(0)$. Let f and g satisfy (1.3).*

For any $\epsilon_0 > 0$, there exists $\lambda_0 > 0$ depending only on ϵ_0 , n , Λ_1 and Λ_2 such that if $\lambda \geq \lambda_0$, $w \in \mathcal{C}(\overline{\Omega})$ is a solution of (4.2) and $u \in \mathcal{C}(\overline{\Omega})$ is a solution of the transport equation (4.1) with $u(x_0) = w(x_0)$ for some $x_0 \in \Omega$, then we have

$$\|u - w\|_{L^\infty(B_1(0))} \leq \epsilon_0. \quad (4.3)$$

Proof. By contradiction, assume that there exist some ϵ_0 , $\{\lambda_k\} \subset \mathbb{R}$, $\{\Omega_k\}$, $\{\Omega'_k\}$, $\{f_k\}$, $\{g_k\}$, $\{w_k\}$, $\{u_k\}$ and $\{x_k\}$ with $x_k \in \Omega_k$ such that $\lambda_k \rightarrow \infty$, $B_1(0) \subset \Omega_k \subset \overline{B}_n(0)$ closed, $B_1(0) \subset \Omega'_k \subset \overline{B}_n(0)$ closed, f_k and g_k satisfy (1.3) on its domains Ω_k and Ω'_k , $w_k \in \mathcal{C}(\overline{\Omega}_k)$ is a solution of

$$\begin{cases} g_k(x + \nabla w_k) \det(I + D^2 w_k) = f_k & \text{in } \Omega_k, \\ \partial_{c_q} w_k(\Omega_k) = \Omega'_k, \end{cases}$$

and $u_k \in \mathcal{C}(\overline{\Omega}_k)$ is a solution of

$$\begin{cases} g_k(x - \nabla c^*(-\nabla u_k)) \det(I + D^2 c_k^*(-\nabla u_k) D^2 u_k) = f_k & \text{in } \Omega_k, \\ \partial_{c_k} u_k(\Omega_k) = \Omega'_k, \end{cases}$$

where $c_k = c_{\lambda_k}$, $u_k(x_k) = w_k(x_k)$ and

$$\|u_k - w_k\|_{L^\infty(B_1(0))} \geq \epsilon_0 \quad \text{for all } k \geq 1. \quad (4.4)$$

By selecting a subsequence, we know from [28, Theorem 1.8.4] that $\{\Omega_k\}$ converges to Ω and $\{\Omega'_k\}$ converges to Ω' in the Hausdorff distance topology. Moreover, Ω and Ω' are closed sets satisfying $B_1(0) \subset \Omega \subset \overline{B}_n(0)$ and $B_1(0) \subset \Omega' \subset \overline{B}_n(0)$. It also follows that $x_k \rightarrow \bar{x}$ for some $\bar{x} \in \Omega$.

Note that $\|u_k - w_k\|_{L^\infty(B_1(0))} = \|[u_k - u_k(x_k)] - [w_k - w_k(x_k)]\|_{L^\infty(B_1(0))}$. By replacing u_k and w_k by $u_k - u_k(x_k)$ and $w_k - w_k(x_k)$ respectively, we can assume in addition that $u_k(x_k) = w_k(x_k) = 0$. Next for convenience, we extend u_k as a c_k -convex function on $\overline{B}_n(0)$. Let us describe precisely this natural extension. Define the Kantorovich functional as in (2.1),

$$I_k(\phi, \psi) := - \int_{\Omega_k} f_k(x) \phi(x) dx - \int_{\Omega'_k} g_k(y) \psi(y) dy,$$

on the set

$$K_{c_k} := \{(\phi, \psi) \in Lip(\Omega_k) \times Lip(\Omega'_k) : -\phi(x) - \psi(y) \leq c_k(x - y) \quad \text{for all } x \in \Omega_k, y \in \Omega'_k\}.$$

Let v_k be the c_k -transform of u_k , that is,

$$v_k(y) := \sup_{x \in \Omega_k} \{-c_k(x - y) - u_k(x)\} \quad \text{for } y \in \mathbb{R}^n. \quad (4.5)$$

We know that (u_k, v_k) is a maximizer of the functional $I_k(\phi, \psi)$ over K_{c_k} , and furthermore

$$u_k(x) = \sup_{y \in \Omega'_k} \{-c_k(x - y) - v_k(y)\} \quad \text{for all } x \in \Omega_k.$$

Now let us extend u_k by defining

$$u_k^*(x) := \sup_{y \in \overline{B}_n(0)} \{-c_k(x - y) - v_k(y)\} \quad \text{for } x \in \overline{B}_n(0) \supset \Omega_k. \quad (4.6)$$

Since $\Omega'_k \subset \overline{B}_n(0)$, it is clear that $u_k^*(x) \geq u_k(x)$ for all $x \in \Omega_k$. On the other hand, $\sup_{y \in \overline{B}_n(0)} \{-c_k(x-y) - v_k(y)\} \leq u_k(x)$ for $x \in \Omega_k$ by (4.5). Thus, $u_k^*|_{\Omega_k} \equiv u_k$. In particular, we have

$$\begin{aligned} \sup_{x \in \overline{B}_n(0)} \{-c_k(x-y) - u_k^*(x)\} &\geq \sup_{x \in \Omega_k} \{-c_k(x-y) - u_k^*(x)\} \\ &= \sup_{x \in \Omega_k} \{-c_k(x-y) - u_k(x)\} = v_k(y). \end{aligned}$$

Also, for $y \in \overline{B}_n(0)$, it follows from the definition of u_k^* in (4.6) that $u_k^*(x) \geq -c_k(x-y) - v_k(y)$ for all $x \in \overline{B}_n(0)$ giving $v_k(y) \geq \sup_{x \in \overline{B}_n(0)} \{-c_k(x-y) - u_k^*(x)\}$. Hence we obtain

$$v_k(y) = \sup_{x \in \overline{B}_n(0)} \{-c_k(x-y) - u_k^*(x)\} \quad \forall y \in \overline{B}_n(0).$$

From now on, we identify u_k with its extension u_k^* given by (4.6). Likewise, we also identify the function w_k with its extension on $\overline{B}_n(0)$. Let us check that $\{u_k\}$ is equicontinuous on $\overline{B}_n(0)$: fix $x_1, x_2 \in \overline{B}_n(0)$. By the extension of u_k , there exists $y_k \in \partial_{c_k} u_k(x_2)$ such that $y_k \in \overline{B}_n(0)$. Then by definition,

$$u_k(x_1) - u_k(y_k) \geq -c_k(x_1 - y_k) + c_k(y_k - x_2).$$

Because $\|y_k\| \leq n$, we get

$$|c_k(x_1 - y_k) - c_k(x_2 - y_k)| \leq C|x_1 - x_2|$$

for some constant C independent of k . Thus $u_k(x_1) - u_k(x_2) \geq -C|x_1 - x_2|$. It then follows by interchanging x_1 and x_2 that

$$|u_k(x_1) - u_k(x_2)| \leq C|x_1 - x_2|,$$

and we have proved that $\{u_k\}$ is equicontinuous on $\overline{B}_n(0)$. Moreover $\{u_k\}$ is uniformly bounded, because we have just seen that $|u_k(x)| = |u_k(x) - u_k(x_k)| \leq C|x - x_k| \leq C$. Therefore by Arzela-Ascoli theorem, we have that a subsequence, still denoted by u_k , converges uniformly on $\overline{B}_n(0)$ to some $u_\infty \in \mathcal{C}(\overline{B}_n(0))$. Similarly, there exists a subsequence $\{w_k\}$ converging uniformly on $\overline{B}_n(0)$ to some $w_\infty \in \mathcal{C}(\overline{B}_n(0))$. These yield $u_\infty(\bar{x}) = w_\infty(\bar{x})$ for some \bar{x} as a consequence.

Notice that f_k and g_k can be viewed as functions defined on $B_n(0)$ by extending them to be zero outside Ω_k and Ω'_k respectively. Then by extracting subsequences, we can assume that $f_k \xrightarrow{*} f$ weakly* in $L^\infty(B_n(0))$ and $g_k \xrightarrow{*} g$ weakly* in $L^\infty(B_n(0))$. Let $\mu_k := f_k(x)dx$, $\mu := f(x)dx$, $\nu_k := g_k(y)dy$ and $\nu := g(y)dy$.

Claim 1: $\overline{B}_1(0) \subset \text{supp } \mu \subset \Omega$ and $\overline{B}_1(0) \subset \text{supp } \nu \subset \Omega'$. Indeed since $f_k \xrightarrow{*} f$ weakly* in $L^\infty(B_n(0))$, we get $\{\mu_k\}$ converges weakly* to μ as measures, i.e.,

$$\lim_{k \rightarrow \infty} \int_{B_n(0)} \varphi(x) d\mu_k = \int_{B_n(0)} \varphi(x) d\mu \quad \text{for all } \varphi \in \mathcal{C}_c(B_n(0)).$$

Now let $x \in \text{supp } \mu$. Then by following the proof of [1, Proposition 5.1.8] where probability measures are considered, we see that there exists $x_k \in \text{supp } \mu_k = \Omega_k$ such that $x_k \rightarrow x$. But then we must have $x \in \Omega$ since $\{\Omega_k\}$ converges to Ω in the Hausdorff topology. Thus

$\text{supp } \mu \subset \Omega$. Next if $x_0 \in B_1(0)$, then $\Lambda_1 \leq \frac{1}{|B(x_0, r)|} \int_{B(x_0, r)} f \, dx \leq \Lambda_2$ for all $r > 0$ sufficiently small because $\Lambda_1 \leq \frac{1}{|B(x_0, r)|} \int_{B(x_0, r)} f_k \, dx \leq \Lambda_2$ and

$$\frac{1}{|B(x_0, r)|} \int_{B(x_0, r)} f_k \, dx \longrightarrow \frac{1}{|B(x_0, r)|} \int_{B(x_0, r)} f \, dx \quad \text{as } k \rightarrow \infty.$$

Hence by letting $r \rightarrow 0^+$, we obtain $\Lambda_1 \leq f(x_0) \leq \Lambda_2$ for a.e. $x_0 \in B_1(0)$. This implies $\bar{B}_1(0) \subset \text{supp } \mu$ and we have shown that $\bar{B}_1(0) \subset \text{supp } \mu \subset \Omega$. Thus Claim 1 is proved since the case for ν is completely similar.

We would like to show next that u_∞ and w_∞ are both solutions of (4.2). Let us prove this only for u_∞ as the case of w_∞ is much simpler and follows in a similar way. Recall that $v_k(y) = \sup_{x \in \bar{B}_n(0)} \{-c_k(x - y) - u_k(x)\}$ for all $y \in \bar{B}_n(0)$, and let

$$v_\infty(y) := \sup_{x \in \bar{B}_n(0)} \{-c_q(x - y) - u_\infty(x)\} \quad \text{for } y \in \bar{B}_n(0),$$

where c_q denotes the quadratic cost. Then since $c_k(z) = c_q(z) + \frac{E_k(z)}{\lambda_k}$, it is easy to see that

$$\|v_k - v_\infty\|_{L^\infty(\bar{B}_n(0))} \leq \|u_k - u_\infty\|_{L^\infty(\bar{B}_n(0))} + \frac{C}{\lambda_k}$$

yielding $v_k \rightarrow v_\infty$ uniformly in $\bar{B}_n(0)$.

Claim 2: We have

$$\lim_{k \rightarrow \infty} \int_{\Omega_k} f_k u_k = \int_{\Omega} f u_\infty, \quad \lim_{k \rightarrow \infty} \int_{\Omega'_k} g_k v_k = \int_{\Omega'} g v_\infty \quad \text{and} \quad \int_{\Omega} f = \int_{\Omega'} g.$$

To see this, observe that $\int_{\Omega} u_\infty f = \int_{\text{supp } \mu} u_\infty f = \int_{B_n(0)} u_\infty f$ since $\text{supp } \mu \subset \Omega$ by Claim 1. Therefore,

$$\begin{aligned} \int_{\Omega_k} f_k u_k - \int_{\Omega} f u_\infty &= \int_{\Omega_k} f_k (u_k - u_\infty) + \left[\int_{\Omega_k} f_k u_\infty - \int_{\Omega} f u_\infty \right] \\ &= \int_{\Omega_k} f_k (u_k - u_\infty) + \left[\int_{B_n(0)} f_k u_\infty - \int_{B_n(0)} f u_\infty \right] \longrightarrow 0 \end{aligned}$$

giving $\lim_{k \rightarrow \infty} \int_{\Omega_k} f_k u_k = \int_{\Omega} f u_\infty$ as desired. Similarly, $\lim_{k \rightarrow \infty} \int_{\Omega'_k} g_k v_k = \int_{\Omega'} g v_\infty$. Also

$$\int_{\Omega} f = \int_{B_n(0)} f = \lim_{k \rightarrow \infty} \int_{B_n(0)} f_k = \lim_{k \rightarrow \infty} \int_{B_n(0)} g_k = \int_{B_n(0)} g = \int_{\Omega'} g,$$

and the claim follows.

We are ready to show that u_∞ is a solution of (4.2). Let $I(\phi, \psi)$ be the Kantorovich functional given by (2.1), and define

$$K_{c_q} := \{(\phi, \psi) \in \text{Lip}(\Omega) \times \text{Lip}(\Omega') : -\phi(x) - \psi(y) \leq c_q(x - y) \quad \text{for all } x \in \Omega, y \in \Omega'\}.$$

Let $(\phi_0, \psi_0) \in K_{c_q}$ be a maximizer of the functional $I(\phi, \psi)$ over K_{c_q} and recall that (u_k, v_k) is a maximizer of the functional $I_k(\phi, \psi)$ over K_{c_k} . As done for (u_k, v_k) , we can extend ϕ_0 and ψ_0 as Lipschitz functions on $\bar{B}_n(0)$ satisfying $\psi_0(y) = \sup_{x \in \bar{B}_n(0)} \{-c_q(x - y) - \phi_0(x)\}$

for all $y \in \overline{B}_n(0)$. Using the facts $(u_k + \frac{C}{\lambda_k}, v_k) \in K_{c_q}$ and $(\phi_0 + \frac{C}{\lambda_k}, \psi_0) \in K_{c_k}$, we obtain $I_k(\phi_0 + \frac{C}{\lambda_k}, \psi_0) - I(\phi_0, \psi_0) \leq I_k(u_k, v_k) - I(\phi_0, \psi_0) \leq I_k(u_k, v_k) - I(u_k + \frac{C}{\lambda_k}, v_k)$ giving

$$\begin{aligned} & -\frac{C}{\lambda_k} \int_{\Omega_k} f_k + \int_{\Omega} f \phi_0 - \int_{\Omega_k} f_k \phi_0 + \int_{\Omega} g \psi_0 - \int_{\Omega'_k} g_k \psi_0 \\ & \leq I_k(u_k, v_k) - I(\phi_0, \psi_0) \leq \frac{C}{\lambda_k} \int_{\Omega} f + \int_{\Omega} f u_k - \int_{\Omega_k} f_k u_k + \int_{\Omega'} g v_k - \int_{\Omega'_k} g_k v_k \quad \text{for all } k. \end{aligned}$$

Hence it follows from Claim 2 that $\lim_{k \rightarrow \infty} I_k(u_k, v_k) = I(\phi_0, \psi_0)$. Consequently,

$$\begin{aligned} I(u_{\infty}, v_{\infty}) &= - \int_{\Omega} f(x) u_{\infty}(x) dx - \int_{\Omega'} g(y) v_{\infty}(y) dy \\ &= \lim_{k \rightarrow \infty} \left\{ - \int_{\Omega_k} f_k(x) u_k(x) dx - \int_{\Omega'_k} g_k(y) v_k(y) dy \right\} \\ &= \lim_{k \rightarrow \infty} I_k(u_k, v_k) = I(\phi_0, \psi_0) = \sup_{(\phi, \psi) \in K_{c_q}} I(\phi, \psi), \end{aligned}$$

where we have again used Claim 2 to obtain the second equality. Since $\int_{\Omega} f = \int_{\Omega'} g$, we infer that u_{∞} is a solution of (4.2) as desired. Similarly, w_{∞} is also a solution of (4.2). Thus $u_{\infty} \equiv w_{\infty}$ on $\text{supp } \mu$ as $u_{\infty}(\bar{x}) = w_{\infty}(\bar{x})$ and (4.2) has a unique solution modulo additive constants on the support of $\mu = f(x) dx$. This together with the fact $B_1(0) \subset \text{supp } \mu$ yields $\lim_{k \rightarrow \infty} \|u_k - w_k\|_{L^{\infty}(B_1(0))} = \|u_{\infty} - w_{\infty}\|_{L^{\infty}(B_1(0))} = 0$, a contradiction to (4.4). \square

Remark. If we assume in addition that Ω is convex in Proposition 4.2, then the conclusion (4.3) can be strengthened by $\|u - w\|_{L^{\infty}(\Omega)} \leq \epsilon_0$. The reason is that in this case we actually have $\text{supp } \mu = \Omega$ instead of $\text{supp } \mu \subset \Omega$. Indeed, let $x \in \text{Int}(\Omega)$ and $B(x, r) \subset \Omega$. Then $B(x, r) \subset B_{\frac{r}{4}}(\Omega_k) := \cup_{y \in \Omega_k} B(y, \frac{r}{4})$ for all k sufficiently large because $\Omega \subset B_{\frac{r}{4}}(\Omega_k)$ from the definition of the Hausdorff distance. Since Ω_k are convex, it is clear that $\text{dist}(x, \partial\Omega_k) = \text{dist}(x, \partial B_{\frac{r}{4}}(\Omega_k)) - \frac{r}{4} \geq \text{dist}(x, \partial\Omega) - \frac{r}{4} \geq r - \frac{r}{4} = \frac{3r}{4}$ yielding in particular $B(x, \frac{r}{2}) \subset \Omega_k$ for all k sufficiently large. Consequently,

$$\mu(B(x, r)) = \lim_{k \rightarrow \infty} \mu_k(B(x, r)) \geq \Lambda_1 \lim_{k \rightarrow \infty} |B(x, r) \cap \Omega_k| = \Lambda_1 |B(x, r)|$$

implying $x \in \text{supp } \mu$, i.e., $\text{Int}(\Omega) \subset \text{supp } \mu$. This and the Claim 1 above give $\text{supp } \mu = \Omega$.

4.2 Some initial estimates

With the information given by Proposition 4.1, we can start localizing the problem. Indeed, we will show that points far from the boundary of Ω are mapped (uniformly) to points far from the boundary of Ω' . We remind the reader that we are using the notation

$$\Omega_{\delta_0} = \{x \in \Omega : \text{dist}(x, \partial\Omega) > \delta_0\}.$$

Proposition 4.3. *Given any $\delta_0 > 0$, there exists $\lambda_0 > 0$ depending only on $\delta_0, n, \Lambda_1, \Lambda_2, \Omega$ and Ω' such that for all $\lambda \geq \lambda_0$ and $u \in \mathcal{C}(\overline{\Omega})$ is a solution of (4.1), we have*

$$\text{dist}(\partial_c u(\Omega_{\delta_0}), \partial\Omega') \geq C \delta_0^{\gamma} - \frac{C}{\lambda}, \quad (4.7)$$

for some universal constants $C, \gamma > 0$. Consequently,

$$\text{dist}(\partial_c u(\Omega_{\delta_0}), \partial_c u(\partial\Omega)) \geq C \delta_0^{\gamma} - \frac{C}{\lambda}. \quad (4.8)$$

Proof. For any $\epsilon > 0$, Proposition 4.1 gives that $\|u - [w + u(x_0) - w(x_0)]\|_{L^\infty(\Omega)} < \epsilon$ for all $\lambda \geq \lambda_0 = \lambda_0(\epsilon, n, \Lambda_1, \Lambda_2, \Omega, \Omega')$, where $x_0 \in \Omega$ and $w \in \mathcal{C}(\bar{\Omega})$ is a solution of (4.2). By working with $\tilde{w} := w + u(x_0) - w(x_0)$ if necessary, we can assume that $\|u - w\|_{L^\infty(\Omega)} < \epsilon$. Let ϕ be the smooth convex function satisfying

$$\begin{cases} \det D^2\phi = 1 & \text{in } \Omega_{\frac{\delta_0}{2}}, \\ \phi = 0 & \text{on } \partial\Omega_{\frac{\delta_0}{2}}. \end{cases}$$

It is known that $-C_1(n, \Omega_{\frac{\delta_0}{2}}) \leq \min_{\Omega_{\frac{\delta_0}{2}}} \phi \leq -C_2(n, \Omega_{\frac{\delta_0}{2}})$ (see [22, Proposition 3.2.3]). Hence by using [22, Lemma 5.1.6] we obtain

$$\Omega_{\delta_0} \subset \left\{ x \in \Omega_{\frac{\delta_0}{2}} : \phi(x) \leq -c\delta_0 \right\}, \quad (4.9)$$

where $c > 0$ depending on n and $\Omega_{\frac{\delta_0}{2}}$. In fact, we can choose c to be a constant which depends only on n and Ω for all $\delta_0 > 0$ sufficiently small. Next, we construct the function

$$w^*(x) = w(x) + \epsilon + 2\epsilon(c\delta_0)^{-1}\phi(x).$$

Then we have that $w^*(x) > u(x)$ for $x \in \partial\Omega_{\frac{\delta_0}{2}}$. We take V to be the set

$$V := \left\{ x \in \Omega_{\frac{\delta_0}{2}} : w^*(x) < u(x) \right\}.$$

It is clear from (4.9) that $\Omega_{\delta_0} \subset V \Subset \Omega_{\frac{\delta_0}{2}}$. Therefore, Lemma 3.3 assures that

$$\partial_c u(\Omega_{\delta_0}) \subset \partial_c u(V) \subset \partial_c(w^*, V)(V) = \partial_c w^*(V). \quad (4.10)$$

On the other hand, by straightforward computation and using that our cost c_λ is just a C/λ -perturbation of the quadratic cost c_q , then

$$\partial_c w^*(V) \subset \mathcal{N}_{\frac{2\epsilon}{c\delta_0} \|\nabla\phi\|_{L^\infty(\Omega_{\delta_0/2})} + \frac{C}{\lambda}} \left\{ \partial_{c_q} w(V) \right\} = \mathcal{N}_{\frac{2\epsilon}{c\delta_0} \|\nabla\phi\|_{L^\infty(\Omega_{\delta_0/2})} + \frac{C}{\lambda}} \left\{ \nabla\hat{w}(V) \right\}, \quad (4.11)$$

where $\hat{w} = w + \frac{|x|^2}{2}$ and we note that $\|\nabla\phi\|_{L^\infty(\Omega_{\frac{\delta_0}{2}})} \leq C$ (see [21, Theorem 17.21]).

Next, for the solution of the optimal transportation problem \hat{w} with quadratic cost, Caffarelli [8] has proved that

1. $\nabla\hat{w}(\Omega) = \Omega'$.
2. $\nabla\hat{w}(\partial\Omega) \subset \partial\Omega'$.
3. $|\nabla\hat{w}(x) - \nabla\hat{w}(y)| \geq C|x - y|^\gamma$ for all $x, y \in \Omega$, where $\gamma > 0$ is a universal constant.

As a consequence, $\nabla\hat{w}(\partial\Omega) = \partial\Omega'$ and

$$\text{dist}(\nabla\hat{w}(V), \partial\Omega') \geq \text{dist}(\nabla\hat{w}(\Omega_{\frac{\delta_0}{2}}), \partial\Omega') \geq C\delta_0^\gamma. \quad (4.12)$$

Combining (4.10), (4.11) and (4.12) we get

$$\text{dist}(\partial_c u(\Omega_{\delta_0}), \partial\Omega') \geq C\delta_0^\gamma - \left[\frac{2\epsilon}{c\delta_0} \|\nabla\phi\|_{L^\infty(\Omega_{\delta_0/2})} + \frac{C}{\lambda} \right].$$

Choosing suitable ϵ , we obtain the estimate (4.7) of the lemma. On the other hand, (4.8) follows from (4.7) and the fact that

$$\partial_c u(\partial\Omega) \subset \partial\Omega',$$

which was proven by Figalli-Kim-McCann [17]. \square

Using the above two propositions, one has a very good control of the sections of u : they are very near the sections of w , and as a consequence, they stay away from the boundary $\partial\Omega$. Let $x_0 \in \Omega_{\delta_0}$ and $y_0 \in \partial_c u(x_0)$. The μ -section of u at x_0 is defined in (2.12), or for $\bar{u}(x) := u(x) + c(x - y_0)$ as in (2.13). One can also define the (quadratic) t -sections for w . Let \bar{x} be the minimum point of the function $\bar{w}(x) := w(x) + \frac{1}{2}|x - y_0|^2$. We define

$$T_t(w, \bar{x}) = \{x \in [\Omega] : w(x) \leq w(\bar{x}) - c_q(x - \bar{y}) + c_q(\bar{x} - \bar{y}) + t\}, \quad (4.13)$$

with

$$\partial_{c_q} w(\bar{x}) = \{\bar{y} := \bar{x} + \nabla w(\bar{x})\}.$$

By using the function \bar{w} we can write

$$T_t(\bar{w}, \bar{x}) = \{x \in [\Omega] : w(x) \leq w(\bar{x}) + \langle \nabla \bar{w}(\bar{x}), x - \bar{x} \rangle + t\}, \quad (4.14)$$

that is a convex set.

Proposition 4.4. *Given $\delta_0 > 0$, there exists $\mu_1 = \mu_1(\delta_0, n, \Lambda_1, \Lambda_2, \Omega, \Omega') > 0$ satisfying that, for all $h, \epsilon > 0$ with $h + \epsilon \leq \mu_1$, we can find $\lambda_0 = \lambda_0(\epsilon, \delta_0, n, \Lambda_1, \Lambda_2, \Omega, \Omega') > 0$ such that if $\lambda \geq \lambda_0$ and u is a solution of (4.1), then*

$$T_{h-\epsilon}(w, \bar{x}) \subset S_h(u, x_0, y_0) \subset T_{h+\epsilon}(w, \bar{x}) \subset \Omega_{c\delta_0^{\theta}} \quad (4.15)$$

for all $x_0 \in \Omega_{\delta_0}$ and $y_0 \in \partial_c u(x_0)$, where w is a solution of the quadratic transport problem (4.2) and \bar{x} is the minimum point of $\bar{w} := w + \frac{1}{2}|x - y_0|^2$. Here c and θ are positive constants depending only on $n, \Lambda_1, \Lambda_2, \Omega$ and Ω' .

Proof. Let $x_0 \in \Omega_{\delta_0}$ and $y_0 \in \partial_c u(x_0)$. It follows from Proposition 4.3 that $y_0 \in \Omega'_{\frac{c}{2}\delta_0^{\gamma}}$ for all $\lambda \geq \lambda_0 = \lambda_0(\delta_0, n, \Lambda_1, \Lambda_2, \Omega, \Omega')$. Define $\bar{u} := u + c(x - y_0)$ and $\hat{w} := w + \frac{1}{2}|x|^2$. Notice that x_0 is a minimum point for \bar{u} in Ω . By definition,

$$\begin{aligned} T_t(w, \bar{x}) &= \{x \in \Omega : \bar{w}(x) \leq \bar{w}(\bar{x}) + \langle \nabla \bar{w}(\bar{x}), x - \bar{x} \rangle + t\} \\ &= \{x \in \Omega : \hat{w}(x) \leq \hat{w}(\bar{x}) + \langle \nabla \hat{w}(\bar{x}), x - \bar{x} \rangle + t\}, \end{aligned}$$

a convex set. As $\nabla \bar{w}(\bar{x}) = 0$, we obtain $\nabla \hat{w}(\bar{x}) = \nabla w(\bar{x}) + \bar{x} = y_0 \in \Omega'_{\frac{c}{2}\delta_0^{\gamma}}$. On the other hand, we know from [8] (see the discussion before (4.12)) that $\nabla \hat{w}(\partial\Omega) = \partial\Omega'$ and $|\nabla \hat{w}(x) - \nabla \hat{w}(y)| \leq C|x - y|^{\gamma}$ for all $x, y \in \bar{\Omega}$. Therefore we conclude $\text{dist}(\bar{x}, \partial\Omega) \geq c\delta_0^{\gamma^2} > 0$. But since \hat{w} is a strictly convex function (c.f. [7]), there exists $\mu_1 > 0$ depending on $\delta_0, n, \Lambda_1, \Lambda_2, \Omega$ and Ω' such that

$$T_{\mu_1}(w, \bar{x}) \subset \Omega_{\frac{c}{2}\delta_0^{\gamma^2}}, \quad (4.16)$$

which gives the last inclusion in (4.15) whenever $h + \epsilon \leq \mu_1$.

Next for any $\epsilon > 0$, by Proposition 4.1 there exists λ_0 depending only on $\epsilon, n, \Lambda_1, \Lambda_2, \Omega$ and Ω' such that for all $\lambda \geq \lambda_0$ we have $\|u - [w + u(x_0) - w(x_0)]\|_{L^\infty(\Omega)} \leq \epsilon/4$ (of course, we are abusing the notation, just take λ_0 to be the smallest of all the choices). Since the sections of w are the same as the corresponding sections of $w + u(x_0) - w(x_0)$, we can assume that $\|u - w\|_{L^\infty(\Omega)} \leq \epsilon/4$. This implies that

$$\|\bar{u} - \bar{w}\|_{L^\infty(\Omega)} \leq \frac{\epsilon}{4} + \frac{C}{\lambda} \leq \frac{\epsilon}{2} \quad \text{for all } \lambda \geq \lambda_0. \quad (4.17)$$

Now it is easy to see that the first two inclusions in (4.15) hold true. Indeed if $x \in T_{h-\epsilon}(w, \bar{x}) = \{x \in \Omega : \bar{w}(x) \leq \bar{w}(\bar{x}) + h - \epsilon\}$, then

$$\bar{u}(x) \leq \bar{w}(x) + \frac{\epsilon}{2} \leq \bar{w}(\bar{x}) + h - \frac{\epsilon}{2} \leq \bar{w}(x_0) + h - \frac{\epsilon}{2} \leq \bar{u}(x_0) + h.$$

On the other hand, if $x \in S_h(u, x_0, y_0) = \{x \in \Omega : \bar{u}(x) \leq \bar{u}(x_0) + h\}$ then

$$\bar{w}(x) \leq \bar{u}(x) + \frac{\epsilon}{2} \leq \bar{u}(x_0) + h + \frac{\epsilon}{2} \leq \bar{u}(\bar{x}) + h + \frac{\epsilon}{2} \leq \bar{w}(\bar{x}) + h + \epsilon.$$

These yield $T_{h-\epsilon}(w, \bar{x}) \subset S_h(u, x_0, y_0) \subset T_{h+\epsilon}(w, \bar{x})$ as desired. \square

As a corollary, we obtain the following result which is the starting point for the induction process in Section 7:

Corollary 4.5. *Given $\delta_0 > 0$, let $\mu_1 > 0$ be the corresponding constant given by Proposition 4.4. Then for all $h \leq \frac{\mu_1}{2}$, there exists $\lambda_0 = \lambda_0(h, \delta_0, n, \Lambda_1, \Lambda_2, \Omega, \Omega') > 0$ such that if $\lambda \geq \lambda_0$ and u is a solution of (4.1), we have the following: for all $x_0 \in \Omega_{\delta_0}$ and $y_0 \in \partial_c u(x_0)$,*

(i) $S_h(u, x_0, y_0) \subset \Omega_{c\delta_0^\theta}$;

(ii) *There is an affine transformation $Tx = Ax + b$ with $C_1 \leq |\det A|^{\frac{2}{n}} h \leq C_2$ and $\|A\| \leq Ch^{-1}$ such that*

$$B_K(0) \subset \tilde{\Omega} := TS_h(u, x_0, y_0) \subset B_n(0).$$

Here all the constants are universal (depending on n, Λ_1 and Λ_2). As a consequence,

$$\text{dist}(x, \partial[\tilde{\Omega}]) \leq \frac{Ch^{-1}}{\lambda} \quad \forall x \in \partial\tilde{\Omega}.$$

Proof. Applying Proposition 4.4 for $\epsilon = h/2$, we obtain

$$T_{\frac{h}{2}}(w, \bar{x}) \subset S_h(u, x_0, y_0) \subset T_{\frac{3h}{2}}(w, \bar{x}) \subset \Omega_{c\delta_0^\theta},$$

where w is the solution of the quadratic transport problem (4.2). Let $Tx = Ax + b$ be an affine transformation normalizing $T_{3h/2}(w, \bar{x})$, i.e., $B_1 \subset T(T_{3h/2}(w, \bar{x})) \subset B_n$. Then as $|T_{3h/2}(w, \bar{x})| \approx h^{\frac{n}{2}}$ (see [22, Corollary 3.2.4]), it is clear that $C_1 \leq |\det A|^{\frac{2}{n}} h \leq C_2$. Also it follows from Theorem 3.3.8 in [22] that $B_K \subset T(T_{h/2}(w, \bar{x}))$ with $K > 0$ a universal constant. Thus $B_K \subset TS_h(u, x_0, y_0) \subset B_n$. Note that as $B_{Ch}(\bar{x}) \subset T_{3h/2}(w, \bar{x})$, we get $T(B_{Ch}(\bar{x})) \subset B_n$ implying that $\|A\| \leq Ch^{-1}$. This together with Proposition 3.2 yields the last estimate. \square

In the above corollary, f and g are only required to satisfy the condition (1.3). However if in addition $f \in C_{loc}^\alpha(\Omega)$ and $g \in C_{loc}^\alpha(\Omega')$ as assumed in Theorem 1.1, then we in fact have the following better estimates

$$S_h(u, x_0, y_0) \subset B_{C\sqrt{h}}(x_0), \quad \|A\| \leq Ch^{-\frac{1}{2}}, \quad \|A^{-1}\| \leq Ch^{\frac{1}{2}} \quad (4.18)$$

and

$$\text{dist}(x, \partial[\tilde{\Omega}]) \leq \frac{Ch^{-\frac{1}{2}}}{\lambda} \quad \forall x \in \partial\tilde{\Omega}. \quad (4.19)$$

Indeed the regularity of w in [7] gives $CI \leq D^2\bar{w} \leq C'I$ in $\Omega_{c\delta_0^\theta}$ which in turn implies that

$$B_{C_1\sqrt{h}}(\bar{x}) \subset T_{3h/2}(w, \bar{x}) \subset B_{C_2\sqrt{h}}(\bar{x}). \quad (4.20)$$

Then (4.18) and (4.19) follow from this, the fact $T^{-1}(B_K(0)) \subset S_h(u, x_0, y_0) \subset T_{3h/2}(w, \bar{x})$ and the arguments used in the proof of Corollary 4.5.

Another consequence of our construction is a gradient bound for points in the section $S_h := S_h(u, x_0, y_0)$. We denote

$$\tilde{u}(x) := \frac{1}{h} [u(T^{-1}x) - u(x_0) - c(x_0 - y_0) - h], \quad \tilde{\Omega} := TS_h(u, x_0, y_0) \quad \text{and} \quad \tilde{c}(x) := \frac{1}{h} c(T^{-1}x), \quad (4.21)$$

where T is the transformation from the previous corollary. Then:

Corollary 4.6. *Under the same assumptions as in Corollary 4.5, assume further that $f \in \mathcal{C}_{loc}^\alpha(\Omega)$ and $g \in \mathcal{C}_{loc}^\alpha(\Omega')$, then we have that*

$$\partial_c^{loc} u(S_h) \subset B_{C(\sqrt{h} + \frac{1}{\lambda})}(y_0)$$

and

$$\partial_{\tilde{c}}^{loc} \tilde{u}(\tilde{\Omega}) \subset B_{C(1 + \frac{1}{\lambda\sqrt{h}})}(Ty_0).$$

Proof. As in the proof of Corollary 4.5, we choose $\epsilon = h/2$ and have $\|u - w\|_{L^\infty(\Omega)} \leq \epsilon/4$. Also it follows from (4.20) that $S_h \subset T_{3h/2}(w, \bar{x}) \subset B_{C_2\sqrt{h}}(\bar{x})$. Now applying Lemma 3.4 with $\delta := C\sqrt{h}$, we then obtain $\partial_c^{loc} u(S_h) \subset \mathcal{N}_{(C\sqrt{h} + \frac{\epsilon}{\lambda})} \left\{ [\nabla \hat{w}(B_{C\sqrt{h}}(\bar{x}))] \right\}$, where $\hat{w} := w + \frac{1}{2}|x|^2$. Since w is the solution of optimal transport with quadratic cost from Ω to Ω' and $B_{C\sqrt{h}}(\bar{x})$ is away from the boundary of Ω , we get $\|D^2 w\|_{L^\infty(B_{C\sqrt{h}}(\bar{x}))} \leq C$ and hence

$$|\nabla \hat{w}(x) - y_0| = |\nabla \hat{w}(x) - \nabla \hat{w}(\bar{x})| \leq C|x - \bar{x}| \leq C\sqrt{h} \quad \forall x \in B_{C\sqrt{h}}(\bar{x}).$$

Therefore, we conclude that $\partial_c^{loc} u(S_h) \subset B_{C(\sqrt{h} + \frac{1}{\lambda})}(y_0)$. The last conclusion follows from the fact $\partial_{\tilde{c}}^{loc} \tilde{u}(TS_h) = T\partial_c^{loc} u(S_h)$ and the estimate (4.18) for $\|A\|$. \square

4.3 A first (but rough) control of the sections

In this subsection we present a result comparing sections of the potential u with Euclidean balls.

Theorem 4.7. *Assume that $f \in \mathcal{C}_{loc}^\alpha(\Omega)$, $g \in \mathcal{C}_{loc}^\alpha(\Omega')$ and (1.3) holds. Given $\delta_0 > 0$, there exist $\mu_0 = \mu_0(\delta_0, n, \Lambda_1, \Lambda_2, \Omega, \Omega') > 0$ and $\lambda_0 = \lambda_0(\delta_0, n, \Lambda_1, \Lambda_2, \Omega, \Omega') > 0$ such that if $\lambda \geq \lambda_0$ and u is a solution of (4.1), then*

$$B_{C^{-1}\sqrt{\rho}}(x_0) \subset S_\rho(u, x_0, y_0) \subset B_{C\sqrt{\rho}}(x_0) \quad (4.22)$$

for all $x_0 \in \Omega_{\delta_0}$, $y_0 \in \partial_c u(x_0)$ and $0 < \rho \leq \mu_0$, where $C > 0$ depends only on $n, \Lambda_1, \Lambda_2, \|f\|_{\mathcal{C}^\alpha(\Omega_{\delta_1})}$ and $\|g\|_{\mathcal{C}^\alpha(\Omega'_{\delta_1})}$ with $\delta_1 = \delta_1(\delta_0, n, \Lambda_1, \Lambda_2)$.

Proof. Let $\mu_0 := \mu^2$, with $\mu := \mu_1/2$ and $\mu_1 > 0$ is the constant given by Proposition 4.4. By applying Proposition 4.1 with $\epsilon_0 := \mu^3$ we obtain

$$\|u - w\|_{L^\infty(\Omega)} \leq \epsilon_0,$$

where w is an optimal transport solution for the quadratic cost between Ω and Ω' , and with densities f and g , respectively.

Step 1: Since $x_0 \in \Omega_{\delta_0}$ and $\mu + \epsilon_0 \leq \mu_1$, Proposition 4.4 gives

$$T_{\mu-\epsilon_0}(w, \bar{x}) \subset S_\mu(u, x_0) := S_\mu(u, x_0, y_0) \subset T_{\mu+\epsilon_0}(w, \bar{x}) \subset \Omega_{\delta_1},$$

where \bar{x} is the minimum point of $\bar{w} := w + \frac{1}{2}|x - y_0|^2$. Moreover as Ω' is convex and f, g are Hölder continuous we know from Caffarelli $C^{2,\alpha}$ estimate in [7] that w is regular in the interior. It then follows by a similar argument yielding (4.20) that $B_{C_1\sqrt{\mu-\epsilon_0}}(\bar{x}) \subset T_{\mu-\epsilon_0}(w, \bar{x})$ and $T_{\mu+\epsilon_0}(w, \bar{x}) \subset B_{C_2\sqrt{\mu+\epsilon_0}}(\bar{x})$. Thus $B_{C_1\sqrt{\mu-\epsilon_0}}(\bar{x}) \subset S_\mu(u, x_0) \subset B_{C_2\sqrt{\mu+\epsilon_0}}(\bar{x})$. This together with the facts $|\bar{x} - x_0| \leq C\sqrt{\epsilon_0}$ (see (6.31) for a proof of this fact) and ϵ_0 is very small compared to μ gives

$$B_{C\sqrt{\mu-\epsilon_0}}(x_0) \subset S_\mu(u, x_0) \subset B_{C'\sqrt{\mu+\epsilon_0}}(x_0).$$

For simplicity and without loss of generality we assume that $x_0 = 0$. Therefore, $B_1(0) \subset \mu^{-\frac{1}{2}}S_\mu(u, x_0) \subset B_n(0)$. Let $\Theta(x) := x - \nabla c^*(-\nabla u(x))$ be the optimal map corresponding to u and let $\Theta^*(y) := y - \nabla c^*(-\nabla u^*(y))$ be the dual optimal map which is known to be the inverse of Θ . Then since w is regular, we know from Lemma 3.4 and the proof of Corollary 4.6 that

$$\Theta B_R(x_0) \subset B_{C(R+\frac{1}{\lambda}+\frac{\epsilon_0}{R})}(y_0) \quad \forall R \leq C\sqrt{\mu}. \quad (4.23)$$

Likewise as $\|u^* - w^*\|_{L^\infty(\Omega')} \leq \epsilon_0 + \frac{C}{\lambda}$ and the dual potential w^* for the quadratic cost is regular thanks to the convexity of Ω , we also have

$$\Theta^* B_R(y_0) \subset B_{C(R+\frac{1}{\lambda}+\frac{\epsilon_0}{R})}(x_0) \quad \text{or} \quad B_R(y_0) \subset \Theta B_{C(R+\frac{1}{\lambda}+\frac{\epsilon_0}{R})}(x_0) \quad \forall R \leq C\sqrt{\mu}. \quad (4.24)$$

Let $\Omega_1 := \mu^{-\frac{1}{2}}\Theta^*[\Theta(S_\mu(u, x_0))]$, $\Omega'_1 := \mu^{-\frac{1}{2}}[\Theta(S_\mu(u, x_0))]$ and

$$u_1(x) := \frac{1}{\mu}u(\mu^{\frac{1}{2}}x), \quad x \in \mu^{-\frac{1}{2}}\Omega \supset \Omega_1. \quad (4.25)$$

Here we are using the notation $[A]$ for the convex hull of the set A . The advantage of this choice of domains lies in the fact that Ω'_1 is convex. On the other hand, by Lemma 4.8, $u_1|_{\Omega_1}$ is the optimal transport solution between Ω_1 and Ω'_1 with densities $f_1(x) := f(\mu^{\frac{1}{2}}x)$, $g_1(y) := g(\mu^{\frac{1}{2}}y)$ and with cost

$$c_1(x) := \frac{1}{\mu}c(\mu^{\frac{1}{2}}x) = \frac{1}{2}|x|^2 + \frac{1}{\lambda_1}E_1(x) \quad \text{with} \quad \frac{1}{\lambda_1} := \frac{\mu^{1/2}}{\lambda}.$$

Note that all the dilations of regions or variables in Ω are with respect to $x_0 = 0$ and in Ω' are always with respect to y_0 . Moreover by using (4.23) and (4.24) we obtain

$$B_1(0) \subset \mu^{-\frac{1}{2}}S_\mu(u, x_0) \subset \Omega_1 \subset \mu^{-\frac{1}{2}}\Theta^*B_{C'\sqrt{\mu+\epsilon_0}}(y_0) \subset \mu^{-\frac{1}{2}}B_{C\sqrt{\mu+\epsilon_0}}(x_0) \subset B_n(0)$$

and

$$B_1(y_0) \subset \mu^{-\frac{1}{2}}B_{C\sqrt{\mu-\epsilon_0}}(y_0) \subset \mu^{-\frac{1}{2}}\Theta S_\mu(u, x_0) \subset \Omega'_1 \subset \mu^{-\frac{1}{2}}B_{C'\sqrt{\mu+\epsilon_0}}(y_0) \subset B_n(y_0).$$

Hence we can apply Proposition 4.2 to get

$$\|u_1 - w_1\|_{L^\infty(B_1)} \leq \epsilon_0, \quad (4.26)$$

where w_1 is the optimal transport solution for the quadratic cost between Ω_1 and Ω'_1 , and with densities f_1 and g_1 .

Step 2: Let $\Theta_1(x) := x - \nabla c_1^*(-\nabla u_1(x)) = \mu^{-\frac{1}{2}} \Theta(\mu^{\frac{1}{2}}x)$ be the optimal map corresponding to u_1 and let $\Theta_1^*(y) := y - \nabla c_1^*(-\nabla u_1^*(y)) = \mu^{-\frac{1}{2}} \Theta^*(\mu^{\frac{1}{2}}y)$ be its dual optimal map. As the target domain Ω'_1 is convex and f_1, g_1 are Hölder continuous with C^α norms are controlled by those of f and g , we know from Caffarelli $C^{2,\alpha}$ estimate in [7] for quadratic cost that w_1 is regular in the interior. Therefore, by using (4.26) and arguing as in Step 1 we get

$$\Theta_1 B_R(x_0) \subset B_{C(R+\frac{1}{\lambda_1}+\frac{\epsilon_0}{R})}(y_0) \quad \forall R \leq C\sqrt{\mu}. \quad (4.27)$$

We would like to have (4.24) for Θ_1^* but this is not immediate since we do not know whether w_1^* is regular due to the nonconvexity of Ω_1 . Observe that we do not really need Θ_1^* to be well behaved in every interior subregion of Ω'_1 , but only in the subregion $B_{C\sqrt{\mu}}(y_0) \subset \Omega'_1$. To obtain this regularity we proceed as follows.

Let $\tilde{\Omega}_1 := \mu^{-\frac{1}{2}} [S_\mu(u, x_0)]$ and $\tilde{\Omega}'_1 := \mu^{-\frac{1}{2}} \Theta[S_\mu(u, x_0)]$. Then $u_1|_{\tilde{\Omega}_1}$ is the optimal transport solution between $\tilde{\Omega}_1$ and $\tilde{\Omega}'_1$ with cost c_1 and with densities f_1 and g_1 . As above, by using (4.23) and (4.24) we also have $B_1(0) \subset \tilde{\Omega}_1 \subset B_n(0)$ and $B_1(y_0) \subset \tilde{\Omega}'_1 \subset B_n(y_0)$. Note that even though we can not compare $\tilde{\Omega}'_1$ to Ω'_1 , we know that $\tilde{\Omega}'_1$ contains the big ball $B_1(y_0)$ centered at y_0 and as mentioned above this is enough for our purpose. Hence we can apply Proposition 4.2 to get $\|u_1 - \tilde{w}_1\|_{L^\infty(B_1)} \leq \epsilon_0$ implying

$$\|u_1^* - \tilde{w}_1^*\|_{L^\infty(B_1)} \leq \epsilon_0 + \frac{C}{\lambda_1}.$$

Here \tilde{w}_1^* is the optimal transport solution for the quadratic cost between $\tilde{\Omega}'_1$ and $\tilde{\Omega}_1$, and with densities g_1 and f_1 . Moreover as $\tilde{\Omega}_1$ is convex, we know that \tilde{w}_1^* is regular in the interior of $\tilde{\Omega}'_1$. Therefore, we also have

$$\Theta_1^* B_R(y_0) \subset B_{C(R+\frac{1}{\lambda_1}+\frac{\epsilon_0}{R})}(x_0) \quad \text{or} \quad B_R(y_0) \subset \Theta_1 B_{C(R+\frac{1}{\lambda_1}+\frac{\epsilon_0}{R})}(x_0). \quad (4.28)$$

From (4.27) and (4.28), we see that Θ_1 and Θ_1^* enjoy the same property as Θ and Θ^* given by (4.23) and (4.24). This allows us to carry the next level of the rescaling process. Indeed, let $\Omega_2 := \mu^{-1} \Theta^*[\Theta(S_{\mu^2}(u, x_0))]$, $\Omega'_2 := \mu^{-1} \Theta(S_{\mu^2}(u, x_0))$ and

$$u_2(x) := \frac{1}{\mu^2} u(\mu x) = \frac{1}{\mu} u_1(\mu^{\frac{1}{2}}x) \quad \text{for} \quad x \in \mu^{-1}\Omega \supset \mu^{-\frac{1}{2}}\Omega_1 \supset \Omega_2.$$

Then by Lemma 4.8, $u_2|_{\Omega_2}$ is the optimal transport solution between Ω_2 and Ω'_2 with cost $c_2(x) := \frac{1}{\mu^2} c(\mu x)$, and with densities $f_2(x) := f(\mu x)$ and $g_2(y) := g(\mu y)$. Notice that $\Omega_2 = \mu^{-\frac{1}{2}} \Theta_1^*[\Theta_1(S_\mu(u_1, x_0))]$ and $\Omega'_2 = \mu^{-\frac{1}{2}} [\Theta_1(S_\mu(u_1, x_0))]$. Hence by arguing as in Step 1 and using (4.27) and (4.28) we obtain

$$\|u_2 - w_2\|_{L^\infty(B_1)} \leq \epsilon_0,$$

with w_2 is an optimal transport solution for the quadratic cost between Ω_2 and Ω'_2 , and with densities f_2 and g_2 .

Step k : Keep repeating the above process we get

$$\|u_k - w_k\|_{L^\infty(B_1)} \leq \epsilon_0,$$

where $u_k(x) := \frac{1}{\mu^k} u(\mu^{\frac{k}{2}} x)$ and w_k is the optimal transport solution for the quadratic cost between $\Omega_k := \mu^{-\frac{k}{2}} \Theta^*[\Theta(S_{\mu^k}(u, x_0))]$ and $\Omega'_k := \mu^{-\frac{k}{2}} [\Theta(S_{\mu^k}(u, x_0))]$, and with densities $f_k(x) := f(\mu^{\frac{k}{2}} x)$ and $g_k(y) := g(\mu^{\frac{k}{2}} y)$. Since $0 < \rho \leq \mu_0 = \mu^2$, we can pick the positive integer number k such that $\mu^{k+2} < \rho \leq \mu^{k+1}$. Thus

$$\mu^{\frac{k}{2}} T_{\frac{\rho}{\mu^k} - \epsilon_0}(w_k, \bar{x}_k) \subset S_\rho(u, x_0) = \mu^{\frac{k}{2}} S_{\frac{\rho}{\mu^k}}(u_k, x_0) \subset \mu^{\frac{k}{2}} T_{\frac{\rho}{\mu^k} + \epsilon_0}(w_k, \bar{x}_k).$$

Here \bar{x}_k denotes the minimum point of $\bar{w}_k := w_k + \frac{1}{2}|x - y_0|^2$. As the target domain Ω'_k is convex and f_k, g_k are Hölder continuous with C^α norms are controlled by those of f and g , we know that $B_{C_1 \sqrt{\frac{\rho}{\mu^k} - \epsilon_0}}(\bar{x}_k) \subset T_{\frac{\rho}{\mu^k} - \epsilon_0}(w_k, \bar{x}_k)$ and $T_{\frac{\rho}{\mu^k} + \epsilon_0}(w_k, \bar{x}_k) \subset B_{C_2 \sqrt{\frac{\rho}{\mu^k} + \epsilon_0}}(\bar{x}_k)$. It follows from these and the fact $|\bar{x}_k - x_0| \leq C\sqrt{\epsilon_0}$ that

$$S_\rho(u, x_0) \subset \mu^{\frac{k}{2}} B_{C \sqrt{\frac{\rho}{\mu^k} + \epsilon_0}}(x_0) \subset B_{C \sqrt{\rho + \mu^k \epsilon_0}}(x_0) \subset B_{C \sqrt{\rho + \rho \frac{\epsilon_0}{\mu^2}}}(x_0) \subset B_{C \sqrt{\rho}}(x_0).$$

Similarly, we also have $B_{C' \sqrt{\rho}}(x_0) \subset S_\rho(u, x_0)$ yielding, (4.22) as desired. \square

In the above proof, we have used the following simple lemma:

Lemma 4.8. *Let u_1 be defined by (4.25). Then u_1 is the optimal transport solution between $\Omega_1 := \mu^{-\frac{1}{2}} \Theta^*[\Theta(S_\mu(u, x_0))]$ and $\Omega'_1 := \mu^{-\frac{1}{2}} [\Theta(S_\mu(u, x_0))]$ with cost $c_1(x) := \frac{1}{\mu} c(\mu^{\frac{1}{2}} x)$, and with densities $f_1(x) := f(\mu^{\frac{1}{2}} x)$ and $g_1(y) := g(\mu^{\frac{1}{2}} y)$. The similar statement also holds for the function u_k .*

Proof. Let $\Theta_1(x) := x - \nabla c_1^*(-\nabla u_1(x))$. Then as u_1 is c_1 -convex, it suffices to show that $\Theta_{1\#} f_1 \chi_{\Omega_1} dx = g_1 \chi_{\Omega'_1} dy$. Now since $c_1^*(x) = \frac{1}{\mu} c^*(\mu^{\frac{1}{2}} x)$, we obtain

$$\Theta_1(x) = x - \mu^{-\frac{1}{2}} \nabla c^*(-\nabla u(\mu^{\frac{1}{2}} x)) = \mu^{-\frac{1}{2}} \left[\mu^{\frac{1}{2}} x - \nabla c^*(-\nabla u(\mu^{\frac{1}{2}} x)) \right] = \mu^{-\frac{1}{2}} \Theta(\mu^{\frac{1}{2}} x).$$

It follows that for any function $h \in \mathcal{C}(\Omega'_1)$, by letting $\tilde{h}(x) := h(\mu^{-\frac{1}{2}} x)$ we have

$$\begin{aligned} \int_{\Omega_1} h(\Theta_1(x)) f_1(x) dx &= \int_{\mu^{-\frac{1}{2}} \Theta^*[\Theta(S_\mu(u, x_0))]} h\left(\mu^{-\frac{1}{2}} \Theta(\mu^{\frac{1}{2}} x)\right) f(\mu^{\frac{1}{2}} x) dx \\ &= \mu^{-\frac{n}{2}} \int_{\Theta^*[\Theta(S_\mu(u, x_0))]} \tilde{h}(\Theta(z)) f(z) dz \\ &= \mu^{-\frac{n}{2}} \int_{[\Theta(S_\mu(u, x_0))]} \tilde{h}(y') g(y') dy' \\ &= \int_{\mu^{-\frac{1}{2}} [\Theta(S_\mu(u, x_0))]} \tilde{h}(\mu^{\frac{1}{2}} y) g(\mu^{\frac{1}{2}} y) dy = \int_{\Omega'_1} h(y) g_1(y) dy. \end{aligned}$$

Thus $\Theta_{1\#} f_1 \chi_{\Omega_1} dx = g_1 \chi_{\Omega'_1} dy$ and the proof is completed. \square

5 Local vs. Global

We consider now with the second ingredient for the proof of Theorem 1.1, where we look at the rescaled solution \tilde{u} from (4.21) in the rescaled set $\tilde{\Omega}$. It still satisfies the same PDE as u , with the advantage that the value of \tilde{u} on the boundary of $\tilde{\Omega}$ is controlled. Then our strategy is to compare \tilde{u} with the solution of a standard Monge-Ampère equation with given Dirichlet data at the boundary, using the standard elliptic theory for viscosity solutions. However, the main difficulty is that although u is an Alexandroff solution of the original problem, \tilde{u} comes from the restriction of u to a section, so \tilde{u} may not be an Alexandroff or viscosity solution in $\tilde{\Omega}$. In the following, we deal with this crucial issue.

5.1 A notion of viscosity solutions

We next introduce a notion of viscosity solution for the equation $M_\lambda[u] = f$. Unlike the standard definition, a viscosity solution as defined below does not have the local property as it depends on the domain under consideration. That is if u is a viscosity solution of the equation $M_\lambda[u] = f$ in Ω , then u might not be a viscosity solution of the same equation restricted to a subdomain U of Ω .

To motivate our choice of test functions, let us seek a pointwise condition guaranteeing that a function $\phi \in \mathcal{C}^2(\Omega)$ is c_λ -convex in Ω . Note that ϕ is c_λ -convex at x_0 if and only if the function $h_0(x) := \phi(x) + c_\lambda(x - y_0)$ has a minimum point in Ω at x_0 or equivalently h_0 has a supporting plane in Ω at x_0 , where $y_0 := x_0 - \nabla c_\lambda^*(-\nabla\phi(x_0))$. But for $\Omega \subset B_n$, we have

$$\begin{aligned} D^2h_0(x) &= I + D^2\phi(x) + \frac{1}{\lambda}D^2E(x - y_0) \geq I + D^2\phi(x) - \frac{C}{\lambda}I|x - y_0| \\ &\geq I + D^2\phi(x) - \frac{C}{\lambda}I \left(|x_0 + \nabla\phi(x_0)| + 2n \right) \quad \forall x \in \Omega. \end{aligned}$$

Thus if in addition Ω is convex and

$$\inf_{\Omega} (I + D^2\phi) > \frac{C}{\lambda}I \left(\|id + \nabla\phi\|_{L^\infty(\Omega)} + 2n \right),$$

then for any $x_0 \in \Omega$, the function h_0 has a supporting plane in Ω at x_0 . This implies that ϕ is c_λ -convex in Ω and in particular $I + D^2c_\lambda^*(-\nabla\phi(x))D^2\phi(x) \geq 0$ in Ω . We will choose the class of test functions to be those satisfying the above inequality.

Definition 5.1. *Let $\Omega \subset B_n$ be an open set, $f \in \mathcal{C}(\Omega)$ and $g \in \mathcal{C}(\overline{\Omega'})$. Let $u \in \mathcal{C}(\overline{\Omega})$ be a c_λ -convex function satisfying $\partial_{c_\lambda}u(\Omega) \subset \overline{\Omega'}$.*

- (i) *A function ϕ is said to be in the admissible class for the operator M_λ in Ω if $\phi \in \mathcal{C}^2([\Omega])$ and*

$$\inf_{[\Omega]} (I + D^2\phi) > \frac{C}{\lambda}I \left(\|id + \nabla\phi\|_{L^\infty([\Omega])} + 2n \right). \quad (5.1)$$

- (ii) *The function u is said to be a viscosity subsolution (supersolution) of the equation $M_\lambda[u] = f$ in Ω if for any $x_0 \in \Omega$ and for any function ϕ in the admissible class for M_λ in Ω satisfying $u - \phi$ has a maximum (minimum) at x_0 , we have*

$$g(x_0 - \nabla c_\lambda^*(-\nabla\phi(x_0))) \det (I + D^2c_\lambda^*(-\nabla\phi(x_0))D^2\phi(x_0)) \geq (\leq) f(x_0).$$

We call u a viscosity solution of the equation $M_\lambda[u] = f$ in Ω if it is both a viscosity subsolution and viscosity supersolution in Ω .

(iii) Let $U \subset \Omega$ be any open set. The function u is said to be a viscosity supersolution of the equation $M_\lambda[u] = f$ in U relative to Ω if for any $x_0 \in U$ and for any function ϕ in the admissible class for M_λ in Ω satisfying $u - \phi$ has a minimum in Ω at x_0 , we have

$$g(x_0 - \nabla c_\lambda^*(-\nabla\phi(x_0))) \det(I + D^2 c_\lambda^*(-\nabla\phi(x_0)) D^2 \phi(x_0)) \leq f(x_0).$$

Remark. Concerning the above definition we remark that if ϕ is in the admissible class for M_λ in Ω and $u - \phi$ has a maximum or minimum at $x_0 \in \Omega$, then $x_0 - \nabla c_\lambda^*(-\nabla\phi(x_0)) \in \partial_{c_\lambda} u(x_0)$. Thus $x_0 - \nabla c_\lambda^*(-\nabla\phi(x_0))$ belongs to the domain of g as we require $\partial_{c_\lambda} u(\Omega) \subset \overline{\Omega'}$.

We wonder what the relation is between weak/viscosity solutions in our case. The following result is a straightforward modification of Proposition 1.3.4 in [22].

Proposition 5.2. Let $\Omega \subset B_n$ be an open set and $f \in \mathcal{C}(\Omega)$, $g \in \mathcal{C}(\Omega')$. Suppose $u \in \mathcal{C}(\overline{\Omega})$ is a c -convex function satisfying $\partial_c u(\Omega) \subset \Omega'$. Then if u is an Alexandroff weak solution of $M_\lambda[u] = f$ in Ω , we have u is also a viscosity subsolution of the equation in any open subset U of Ω .

Proof. For simplicity, we present the proof only for the case $U = \Omega$. However the arguments clearly work for any open set $U \subset \Omega$. Let $\tilde{x} \in \Omega$ and $\phi \in \mathcal{C}^2([\Omega])$ be a function in the admissible class such that $u - \phi$ has a maximum at \tilde{x} . We can assume that $u(\tilde{x}) = \phi(\tilde{x})$ and $u(x) < \phi(x)$ for all $x \in \Omega \setminus \{\tilde{x}\}$. This can be achieved by adding $r|x - \tilde{x}|^2$ to ϕ and later taking $r \rightarrow 0^+$. Indeed, the only thing we need to check is the function $\tilde{\phi}(x) := \phi(x) + r|x - \tilde{x}|^2$ belongs to the admissible class for M_λ in Ω . But this is true because

$$\begin{aligned} \inf_{[\Omega]} (I + D^2 \tilde{\phi}) &= \inf_{[\Omega]} (I + D^2 \phi) + 2rI > \frac{C}{\lambda} I \left(\|id + \nabla \phi\|_{L^\infty([\Omega])} + 2n \right) + 2rI \\ &\geq \frac{C}{\lambda} I \left(\|id + \nabla \tilde{\phi}\|_{L^\infty([\Omega])} + 2n \right) + 2r \left(1 - \frac{2nC}{\lambda} \right) I \\ &\geq \frac{C}{\lambda} I \left(\|id + \nabla \tilde{\phi}\|_{L^\infty([\Omega])} + 2n \right). \end{aligned} \quad (5.2)$$

Let $m = \min_{\delta/2 \leq |x - \tilde{x}| \leq \delta} \{\phi(x) - u(x)\}$, where $\delta > 0$ is a small number. Note that $m > 0$. Fix $0 < \epsilon < m$ and set

$$E_\epsilon := \{x \in B_\delta(\tilde{x}) : u(x) + \epsilon > \phi(x)\}.$$

It is easy to see that $E_\epsilon \subset B_{\delta/2}(\tilde{x})$, because if $\delta/2 \leq |x - \tilde{x}| < \delta$, then $\phi(x) - u(x) \geq m$, so that $x \notin E_\epsilon$. Note also that $E_\epsilon \downarrow \{\tilde{x}\}$ as $\epsilon \rightarrow 0^+$, and $u + \epsilon = \phi$ on ∂E_ϵ .

Because u is a weak Alexandroff solution in Ω , we have

$$\int_{E_\epsilon} f(x) dx = w_c(g, u)(E_\epsilon) = w_c(g, u + \epsilon)(E_\epsilon) = \int_{\partial_c(u + \epsilon)(E_\epsilon)} g(y) dy. \quad (5.3)$$

But $\partial_c(u + \epsilon)(E_\epsilon) \subset \partial_c(\phi, E_\epsilon)(E_\epsilon) = \partial_c \phi(E_\epsilon)$ by Lemma 3.3 and the fact ϕ is a smooth c -convex function in Ω . Thus the chain of equalities (5.3) continues as follows:

$$(5.3) \leq \int_{\partial_c \phi(E_\epsilon)} g(y) dy = w_c(g, \phi)(E_\epsilon) = \int_{E_\epsilon} g(x - Dc^*(-\nabla\phi)) \det(I + D^2 c^*(-\nabla\phi) D^2 \phi) dx,$$

where the last equality is due to the representation (2.5). Next, because f and g are continuous, we can let $\epsilon \rightarrow 0$ to obtain

$$M_\lambda[\phi] \geq f \quad \text{at } \tilde{x},$$

as desired. \square

5.2 The restriction of a weak solution is almost a weak solution

Notice that under the assumption in Proposition 5.2, the function u might not be a viscosity supersolution. This is because the mass may increase when we restrict to a subdomain (i.e. local but not global supporting costs may appear). However, we are going to show that this extra mass can only come from the subdifferential of u at points near the boundary of the subdomain. In particular, we are able to quantify how bad the restriction of u to a section $S := S_\mu(u, x_0, y_0)$ is. In order not to confuse with the height of a section, we modify our previous notation and denote

$$S^\gamma := \{x \in S : \text{dist}(x, \partial S) > \gamma\}.$$

Proposition 5.3. *Suppose $\Omega \subset B_n$ is open and $u \in \mathcal{C}(\Omega)$ is a c_λ -convex function. Let $S := S_\mu(u, x_0, y_0) \Subset \Omega$ and assume there exists a constant $K > 0$ such that*

$$S_\rho(u, x_1, y_1) \subset B_{K\sqrt{\rho}}(x_1) \quad \text{for all } x_1 \in S, y_1 \in \partial_c u(x_1) \text{ and } \rho > 0 \text{ small.} \quad (5.4)$$

Then we have

$$\partial_c(u, S)(x) = \partial_c u(x) \quad \text{for every } x \in S^\gamma,$$

where $\gamma := \frac{CK}{\sqrt{\lambda}}$. That is, local with respect to S is global as long as we stay sufficiently away from the boundary. Here the constant C depends only on n and the norm $\|\partial_c^{loc} u(S)\|_{L^\infty}$.

Proof. Let $\tilde{x} \in S$, $\tilde{y} \in \partial_c(u, S)(\tilde{x})$ and define $\bar{u}(x) := u(x) + \frac{1}{2}|x|^2$. From the definition, we know that (see the proof of Proposition 3.5 for more details) $\tilde{y} \in \partial_c u(\tilde{x})$ if and only if

$$\bar{u}(x) \geq \bar{u}(\tilde{x}) + \langle \tilde{p} + \tilde{x}, x - \tilde{x} \rangle - \frac{1}{\lambda} \left[E(x - \tilde{y}) - E(\tilde{x} - \tilde{y}) - \langle \nabla E(\tilde{x} - \tilde{y}), x - \tilde{x} \rangle \right] \quad \forall x \in \Omega,$$

where $\tilde{p} := -\nabla c_\lambda(\tilde{x} - \tilde{y}) \in \partial^- u(\tilde{x})$. Suppose that $\tilde{y} \notin \partial_c u(\tilde{x})$, i.e., \tilde{y} is local with respect to S but not global. Since the set

$$\{y \in \mathbb{R}^n : y \in \partial_c(u, S)(z_1) \cap \partial_c(u, S)(z_2) \quad \text{for some } z_1, z_2 \in \bar{S}, z_1 \neq z_2\}$$

has measure zero (see [24, Lemma 2.11]), we can assume without loss of generality that

$$\bar{u}(x) > \bar{u}(\tilde{x}) + \langle \tilde{p} + \tilde{x}, x - \tilde{x} \rangle - \frac{1}{\lambda} \left[E(x - \tilde{y}) - E(\tilde{x} - \tilde{y}) - \langle \nabla E(\tilde{x} - \tilde{y}), x - \tilde{x} \rangle \right]$$

for all $x \in \bar{S}$ satisfying $x \neq \tilde{x}$. As $\tilde{y} \notin \partial_c u(\tilde{x})$, the almost plane

$$z = \tilde{f}(x) := \bar{u}(\tilde{x}) + \langle \tilde{p} + \tilde{x}, x - \tilde{x} \rangle - \frac{1}{\lambda} [E(x - \tilde{y}) - E(\tilde{x} - \tilde{y}) - \langle \nabla E(\tilde{x} - \tilde{y}), x - \tilde{x} \rangle]$$

must cross the graph of \bar{u} at some point $x^* \in \Omega \setminus \bar{S}$.

Among all the points on ∂S that lie on the segment $\overline{\tilde{x}x^*}$ connecting \tilde{x} and x^* , take \hat{x} to be the point closest to \tilde{x} . Let x_1 be the midpoint of the segment $\overline{\tilde{x}\hat{x}}$, $y_1 \in \partial_c u(x_1)$ and $p_1 := -\nabla c_\lambda(x_1 - y_1)$. In particular, we have $|\tilde{x} - x_1| \geq \frac{1}{2} \text{dist}(\tilde{x}, \partial S)$ and $|x^* - x_1| \geq \frac{1}{2} \text{dist}(\tilde{x}, \partial S)$. Because of our choices, it is clear that the almost plane $z = \tilde{f}(x)$ crosses the global supporting cost

$$z = f_1(x) := \bar{u}(x_1) + \langle p_1 + x_1, x - x_1 \rangle - \frac{1}{\lambda} [E(x - y_1) - E(x_1 - y_1) - \langle \nabla E(x_1 - y_1), x - x_1 \rangle]$$

at two different points z_1 and z_2 which lie on the line passing through \tilde{x} and x^* (one point, say z_1 , lies on the segment $\overline{\tilde{x}x_1}$ and the other z_2 lies on the segment $\overline{x_1x^*}$).

Let us first estimate the heights $a := \bar{u}(\tilde{x}) - f_1(\tilde{x})$ and $b := \bar{u}(x^*) - f_1(x^*)$. Assume that $\tilde{x} \in S$ satisfies $\text{dist}(\tilde{x}, \partial S) > 2K\sqrt{\rho}$. Then $|\tilde{x} - x_1| > K\sqrt{\rho}$ and hence $\tilde{x} \notin S_\rho(u, x_1, y_1)$ by the assumption (5.4). But $\tilde{x} \notin S_\rho(u, x_1, y_1)$ means that

$$\bar{u}(\tilde{x}) > f_1(\tilde{x}) + \rho$$

yielding $a = \bar{u}(\tilde{x}) - f_1(\tilde{x}) > \rho$. For the same reason, we also have $b = \bar{u}(x^*) - f_1(x^*) > \rho$.

Next observe that since $E(x - \tilde{y}) = E(\tilde{x} - \tilde{y}) + \langle \nabla E(\tilde{x} - \tilde{y}), x - \tilde{x} \rangle + O(|x - \tilde{x}|^2)$, we get

$$\tilde{f}(x) = \bar{u}(\tilde{x}) + \langle \tilde{p} + \tilde{x}, x - \tilde{x} \rangle + \frac{1}{\lambda} O(|x - \tilde{x}|^2) \geq \tilde{g}(x) := \bar{u}(\tilde{x}) + \langle \tilde{p} + \tilde{x}, x - \tilde{x} \rangle - \frac{C}{\lambda} |x - \tilde{x}|^2.$$

Similarly,

$$\begin{aligned} f_1(x) &= \bar{u}(x_1) + \langle p_1 + x_1, x - x_1 \rangle + \frac{1}{\lambda} O(|x - \tilde{x}|^2) \\ &\quad - \frac{1}{\lambda} \left[E(\tilde{x} - y_1) - E(x_1 - y_1) + \langle \nabla E(\tilde{x} - y_1), x - \tilde{x} \rangle - \langle \nabla E(x_1 - y_1), x - x_1 \rangle \right] \\ &\leq g_1(x) := \bar{u}(x_1) + \langle p_1 + x_1, x - x_1 \rangle + \frac{C}{\lambda} |x - \tilde{x}|^2 \\ &\quad - \frac{1}{\lambda} \left[E(\tilde{x} - y_1) + \langle \nabla E(\tilde{x} - y_1), x - \tilde{x} \rangle - E(x_1 - y_1) - \langle \nabla E(x_1 - y_1), x - x_1 \rangle \right]. \end{aligned}$$

We knew that $a, b > \rho$ and the equation $\tilde{f}(x) = f_1(x)$ has two solutions in the segment $\overline{\tilde{x}x^*}$. Therefore if we select $\rho := \frac{6C}{\lambda}(2n)^2$ ensuring that $\|\tilde{f} - \tilde{g}\|_{L^\infty(\Omega)} < \rho/3$ and $\|f_1 - g_1\|_{L^\infty(\Omega)} < \rho/3$, then the equation $\tilde{g}(x) = g_1(x)$ still has two solutions in the segment $\overline{\tilde{x}x^*}$. The reason for this is that we still have $\tilde{g}(\tilde{x}) > g_1(\tilde{x})$, $\tilde{g}(x_1) < g_1(x_1)$ and $\tilde{g}(x^*) > g_1(x^*)$ because the graph $\tilde{g}(x)$ lies below the graph of \tilde{f} by an amount of at most $\rho/3$ and the graph $g_1(x)$ lies above the graph of f_1 by an amount of at most $\rho/3$.

We now consider the equation $\tilde{g}(x) = g_1(x)$. That is,

$$\begin{aligned} \langle \tilde{p} + \tilde{x}, x - \tilde{x} \rangle - \frac{2C}{\lambda} |x - \tilde{x}|^2 &= [\bar{u}(x_1) - \bar{u}(\tilde{x})] + \langle p_1 + x_1, x - x_1 \rangle \\ &\quad - \frac{1}{\lambda} \left[E(\tilde{x} - y_1) + \langle \nabla E(\tilde{x} - y_1), x - \tilde{x} \rangle - E(x_1 - y_1) - \langle \nabla E(x_1 - y_1), x - x_1 \rangle \right]. \end{aligned}$$

Since $\bar{u}(x_1) + \langle y_1, \tilde{x} - x_1 \rangle - \frac{1}{\lambda} [E(\tilde{x} - y_1) - E(x_1 - y_1)] - \bar{u}(\tilde{x}) = f_1(\tilde{x}) - \bar{u}(\tilde{x}) = -a$, this is equivalent to

$$\langle [\tilde{p} + \tilde{x} + \frac{1}{\lambda} \nabla E(\tilde{x} - y_1)] - y_1, x - \tilde{x} \rangle - \frac{2C}{\lambda} |x - \tilde{x}|^2 = -a.$$

Let us look at what happens along the above mentioned line passing through \tilde{x} and x^* . On this line, we can write $x - \tilde{x} = te$ for $t \in \mathbb{R}$, where e is some unit vector. Then if we denote $\alpha := \langle [\tilde{p} + \tilde{x} + \frac{1}{\lambda} \nabla E(\tilde{x} - y_1)] - y_1, e \rangle$, we conclude that the quadratic equation

$$\frac{2C}{\lambda} t^2 - \alpha t = a$$

has two distinct solutions t_1, t_2 satisfying $z_i := \tilde{x} + t_i e \in \overline{\tilde{x}x^*}$. Notice that as $a > 0$, this is impossible since the two solutions t_1 and t_2 have opposite signs. Thus we have shown that \tilde{y} must be global whenever

$$\text{dist}(\tilde{x}, \partial S) > CK\sqrt{\frac{1}{\lambda}}.$$

□

As a consequence of the previous proposition, we have:

Corollary 5.4. *Let $\Omega \subset B_n$ be open and $f \in \mathcal{C}(\Omega)$, $g \in \mathcal{C}(\overline{\Omega})$. Let $S := S_\mu(u, x_0, y_0) \Subset \Omega$ be a section satisfying $\partial_c u(S) \subset \Omega'$ and (5.4) for some constant $K > 0$. Suppose that $u \in \mathcal{C}(\overline{\Omega})$ is a c_λ -convex function that is an Alexandroff solution in S relative to Ω of the equation $M_\lambda[u] = f$, i.e.,*

$$\int_{\partial_c(u, \Omega)(E)} g(y) dy = \int_E f(x) dx \quad \text{for all Borel sets } E \subset S. \quad (5.5)$$

Then there exists $\gamma = CK/\sqrt{\lambda}$ such that u restricted to S is an Alexandroff weak solution in the set S^γ relative to S , i.e.,

$$\int_{\partial_c(u, S)(E)} g(y) dy = \int_E f(x) dx \quad \text{for all Borel sets } E \subset S^\gamma. \quad (5.6)$$

Moreover, u is a viscosity supersolution in S^γ relative to S .

Proof. By Proposition 5.3, we have $\partial_c(u, S)(E) = \partial_c(u, \Omega)(E)$ for all $E \subset S^\gamma$. This together with (5.5) yields (5.6). It remains to show that u is a viscosity supersolution in S^γ relative to S .

Let $\tilde{x} \in S^\gamma$ and $\phi \in \mathcal{C}^2([S])$ be a function in the admissible class for M_λ in S such that $u - \phi$ has a minimum in S at \tilde{x} . We can assume in addition that $u(\tilde{x}) = \phi(\tilde{x})$ and $u(x) > \phi(x)$ for all x in $\overline{S} \setminus \{\tilde{x}\}$. This can be achieved by adding $-r|x - \tilde{x}|^2$ to ϕ and later taking $r \rightarrow 0^+$. Indeed, we only need to check that the function $\tilde{\phi}(x) := \phi(x) - r|x - \tilde{x}|^2$ is in the admissible class for M_λ in S for all $r > 0$ sufficiently small. But as done in (5.2), we have

$$\begin{aligned} & \inf_{[S]} (I + D^2 \tilde{\phi}) - \frac{C}{\lambda} I \left(\|id + \nabla \tilde{\phi}\|_{L^\infty([S])} + 2n \right) \\ & \geq \inf_{[S]} (I + D^2 \phi) - \frac{C}{\lambda} I \left(\|id + \nabla \phi\|_{L^\infty([S])} + 2n \right) - 2r \left(1 + \frac{2nC}{\lambda} \right) I \\ & > 0 \quad \text{for all } r > 0 \text{ small enough.} \end{aligned}$$

Let $m = \min_{\overline{S} \setminus B_\delta(\tilde{x})} \{u(x) - \phi(x)\}$, where $\delta > 0$ is small. Note that $m > 0$. Fix $0 < \epsilon < m$ and set

$$E_\epsilon := \{x \in S : u(x) - \epsilon < \phi(x)\}.$$

It is easy to see that $E_\epsilon \subset B_\delta(\tilde{x}) \subset S^\gamma$, because if $|x - \tilde{x}| \geq \delta$, then $u(x) - \phi(x) \geq m$, so that $x \notin E_\epsilon$. Note also that $E_\epsilon \downarrow \{\tilde{x}\}$ as $\epsilon \rightarrow 0^+$ and $u - \epsilon = \phi$ on ∂E_ϵ .

Now it follows from (5.6) that $\int_{\partial_c(u, S)(E_\epsilon)} g(y) dy = \int_{E_\epsilon} f(x) dx$. But as a consequence of the proof of Lemma 3.3 and the fact $\phi(x) \leq u(x) - \epsilon$ for $x \in S \setminus E_\epsilon$, we get

$$\partial_c(\phi, S)(E_\epsilon) \subset \partial_c(u - \epsilon, S)(E_\epsilon) = \partial_c(u, S)(E_\epsilon).$$

Therefore we obtain

$$\int_{E_\epsilon} g(x - Dc^*(-\nabla\phi)) \det(I + D^2c^*(-\nabla\phi)D^2\phi) dx = \int_{\partial_c(\phi, S)(E_\epsilon)} g(y) dy \leq \int_{E_\epsilon} f(x) dx,$$

where the first equality is due to (2.5) and the fact ϕ is c -convex in S . Next, because f and g are continuous, we can let $\epsilon \rightarrow 0$ to obtain

$$M_\lambda[\phi] \leq f \quad \text{at } \tilde{x},$$

as desired. \square

6 Comparison arguments

After a careful understanding of viscosity solutions for the equation $M_\lambda[u] = f$, our next aim is to show a comparison principle, that is one of the main ingredients in the proof of Theorem 1.1. We remind the reader of our setting:

$$M_\lambda[u] := g(x + \nabla u + G(\nabla u)) \det(I + D^2u + H(\nabla u)(D^2u)), \quad (6.1)$$

where G and H in particular satisfy

$$\|H(\nabla u)\|_{L^\infty} \leq \frac{C}{\lambda} \|\nabla u\|_{L^\infty(\Omega)}, \quad \|G(\nabla u)\|_{L^\infty} \leq \frac{C}{\lambda} \|\nabla u\|_{L^\infty(\Omega)}^2. \quad (6.2)$$

Since viscosity solutions are expected to be stable with respect to a small perturbation, u should be close to a Monge-Ampère solution if $[\Omega]$ is close to Ω (that is, Ω is almost convex) and λ is large enough.

6.1 Comparison principle for viscosity solutions

It is well known that for viscosity solutions:

Proposition 6.1. *Let $f \in \mathcal{C}(\overline{\Omega})$ and $g \in \mathcal{C}(\overline{\Omega'})$. Suppose that $w \in \mathcal{C}^2([\Omega]) \cap \mathcal{C}(\overline{\Omega})$ is a function in the admissible class for the operator M_λ in Ω . Then:*

- (i) *Let $U \subset \Omega$ be open. If $u \in \mathcal{C}(\overline{\Omega})$ is a viscosity supersolution of $M_\lambda[u] = f$ in U relative to Ω and $M_\lambda[w] > f$ at any point $x \in U$ satisfying $x - \nabla c_\lambda^*(-\nabla w(x)) \in \overline{\Omega'}$, then*

$$\min_{\overline{\Omega}}\{u - w\} = \min_{\overline{\Omega} \setminus U}\{u - w\}.$$

- (ii) *If $u \in \mathcal{C}(\overline{\Omega})$ is a viscosity subsolution of $M_\lambda[u] = f$ and $M_\lambda[w] < f$ at any point $x \in \Omega$ satisfying $x - \nabla c_\lambda^*(-\nabla w(x)) \in \overline{\Omega'}$, then*

$$\max_{\overline{\Omega}}\{u - w\} = \max_{\partial\Omega}\{u - w\}.$$

Proof. For the first statement, suppose by contradiction $\min_{\overline{\Omega}}\{u - w\} < \min_{\overline{\Omega} \setminus U}\{u - w\}$. Then there exists $x_0 \in U$ such that $u(x_0) - w(x_0) = \min_{\overline{\Omega}}\{u - w\}$, and so $u - w$ has a minimum in Ω at x_0 . Notice that $x_0 - \nabla c_\lambda^*(-\nabla w(x_0)) \in \overline{\Omega'}$ by the remark after Definition 5.1. Therefore it follows from the definition of viscosity supersolution that

$$M_\lambda[w] \leq f \text{ at } x_0.$$

This is a contradiction with the initial hypothesis. The second statement follows in a similar manner. \square

Then our main comparison result is:

Theorem 6.2. *Let $B_1 \subset \Omega \subset B_n$ be a domain in \mathbb{R}^n , not necessarily convex, but such that*

$$\text{dist}(x, \partial[\Omega]) \leq \frac{d}{\lambda} \quad \forall x \in \partial\Omega. \quad (6.3)$$

Let $u \in \mathcal{C}(\bar{\Omega})$ be a solution (in the sense precised below) of the problem

$$\begin{cases} M_\lambda[u] = f & \text{in } \Omega, \\ u(x) = -c_\lambda(x) & \text{on } \partial\Omega, \end{cases} \quad (6.4)$$

where M_λ is the operator defined in (2.10) satisfying (2.11). Consider also the c_q -convex solution $w \in \mathcal{C}^3(\Omega) \cap \mathcal{C}(\bar{\Omega})$ of

$$\begin{cases} \det(I + D^2w) = 1 & \text{in } V := \mathcal{N}_{(\frac{d}{\lambda})^\beta}[\Omega], \\ w = -\frac{1}{2}|x|^2 & \text{on } \partial V, \end{cases} \quad (6.5)$$

where $0 < \beta \leq 1$. Assume also that

1. For some $\theta > 0$ and $0 < \alpha < 1$,

$$|f(x) - 1| \leq \theta \quad \text{in } \Omega \quad \text{and} \quad |g(y) - 1| \leq \theta |y|^\alpha \quad \text{in } \partial_c^{loc} u(\Omega). \quad (6.6)$$

2. For a universal constant C_n given precisely in the proof,

$$C_n \left(\frac{d}{\lambda}\right)^{-\beta((n-1)^2 + \frac{n-2}{n}\alpha)} \left[\theta + \frac{1}{\lambda} \left(\frac{d}{\lambda}\right)^{\frac{n-2}{n}(\alpha-1)\beta} \right] \leq 1. \quad (6.7)$$

Then we have that:

(i) If u is a viscosity subsolution in Ω , then

$$u \leq w + C \left(\frac{d}{\lambda}\right)^{\frac{2\beta}{n}} \quad \text{in } \Omega. \quad (6.8)$$

(ii) If u is a viscosity supersolution in Ω^γ relative to Ω for some $\gamma \geq 0$, then

$$u \geq w - \left[m\gamma + C \left(\frac{d}{\lambda}\right)^{\beta(n-1)^2} \right] \quad \text{in } \Omega, \quad (6.9)$$

where $m := \|\partial_c u(\Omega \setminus \Omega^\gamma)\|_{L^\infty}$.

In case that $[\Omega]$ (equivalently V) not strictly convex, we can modify V slightly so that V is strictly convex and still $V \approx \mathcal{N}_{(\frac{d}{\lambda})^\beta}[\Omega]$. Now we note that w exists because $\bar{w} = \frac{1}{2}|x|^2 + w$ is the solution of a standard Monge-Ampère equation with smooth right hand side and homogeneous Dirichlet boundary data, and thus, $\bar{w} \in \mathcal{C}^3(V)$ and it is strictly convex. However, the estimates on \bar{w} (in particular, Pogorelov estimates for the second derivatives) are not uniform up to the boundary. To handle this issue, we will use Proposition 6.3 and Proposition 6.4 below concerning the behavior of \bar{w} near the boundary when we do not have extra information on the domain.

The following is a barrier argument by Caffarelli [6]:

Proposition 6.3. *Let $\Omega_1 \subset B_n$ be a convex domain in \mathbb{R}^n and $\bar{w} \in \mathcal{C}(\bar{\Omega}_1)$ be a generalized solution of the Monge-Ampère equation*

$$\begin{cases} \det(D^2\bar{w}) \leq 1 & \text{in } \Omega_1, \\ \bar{w} = 0 & \text{on } \partial\Omega_1. \end{cases}$$

Then

$$\bar{w}(x) \geq -C(n) \operatorname{dist}(x, \partial\Omega_1)^{\frac{2}{n}} \quad \text{if } n \geq 3,$$

and

$$\bar{w}(x) \geq -C(\rho) \operatorname{dist}(x, \partial\Omega_1)^\rho \quad \text{for any } \rho \in (0, 1) \text{ if } n = 2.$$

We also have a Pogorelov type estimate for the second derivatives of solutions for a Monge-Ampère equation. Although it is well known that in the interior of the domain, the Hessian matrix of the solution is positive definite, the constants deteriorate at the boundary. The next lemma by Caffarelli-Li [3] quantifies this deterioration.

Proposition 6.4. *Let $\Omega_1 \subset \mathbb{R}^n$ be an open convex subset satisfying $B_1 \subset \Omega_1 \subset B_n$, and let $\bar{w} \in \mathcal{C}^2(\Omega_1) \cap \mathcal{C}(\bar{\Omega}_1)$ be a convex solution of*

$$\begin{cases} \det(D^2\bar{w}) = 1 & \text{in } \Omega_1, \\ \bar{w} = 0 & \text{on } \partial\Omega_1. \end{cases}$$

Then

$$|D^2\bar{w}(x)| \leq C'(n) \operatorname{dist}(x, \partial\Omega_1)^{1-n} \quad \text{for } n \geq 3$$

and

$$|D^2\bar{w}(x)| \leq C'(\rho) \operatorname{dist}(x, \partial\Omega_1)^{1-\frac{2}{\rho}} \quad \text{for } n = 2,$$

where ρ is any real number in $(0, 1)$.

Proof of Theorem 6.2: Without loss of generality, assume that $n \geq 3$. The proof reduces to finding suitable smooth sub and supersolutions of the problem $M_\lambda[u] = f$, and using the comparison principle for viscosity solutions given in Proposition 6.1 in the domain Ω .

Let w be the solution of problem (6.5). We will show that the functions

$$w^- = w + \frac{\delta}{2}|x|^2 \quad \text{and} \quad w^+ = w - \frac{\delta}{2}|x|^2$$

are a sub and a supersolution, respectively, of problem (6.4), for a suitable $\delta > 0$ which will be determined later. In the following, we denote $\bar{w} := \frac{1}{2}|x|^2 + w$ and $\bar{m} := \|\nabla\bar{w}\|_{L^\infty(\Omega)}$. Observe that \bar{w} is the unique convex function satisfying

$$\begin{cases} \det D^2\bar{w} = 1 & \text{in } V, \\ \bar{w} = 0 & \text{on } \partial V. \end{cases}$$

By the hypothesis for f and g , we have that

$$|f(x) - 1| \leq \theta \tag{6.10}$$

and

$$\left| g\left(x + \nabla w \pm \delta x + G(\nabla w \pm \delta x)\right) - 1 \right| \leq \theta (2\bar{m} + 2n)^\alpha. \tag{6.11}$$

We note that to obtain the estimate (6.11), we have also used the fact $C\bar{m}/\lambda \leq 1$ which is a consequence of the estimate for \bar{m} given in the Step 3 below and our assumption (6.7).

Step 1. Construction of a supersolution. We show first that w^+ is in the admissible class for the operator M_λ in Ω (see Definition 5.1) and that the matrix

$$A^+ := I + D^2c_\lambda^*(-\nabla w^+)D^2w^+ = I + D^2w - \delta I + (D^2w - \delta I)H(\nabla w - \delta x)$$

is positive definite. Indeed, we can write

$$\begin{aligned} A^+ &= D^2\bar{w}[I + H] - \delta I - (1 + \delta)H \\ &\geq \frac{1}{2}D^2\bar{w} - \delta I - 2|H|I \\ &= \left[\frac{1}{4}D^2\bar{w} - \delta I\right] + \left[\frac{1}{4}D^2\bar{w} - 2|H|I\right]. \end{aligned}$$

Moreover, it follows from the hypothesis (6.2) for H that

$$|H| \leq \frac{C(\bar{m} + 2n)}{\lambda} \leq \frac{\delta}{4}, \quad (6.12)$$

for a suitable choice of δ . Therefore,

$$A^+ \geq 2 \left[\frac{1}{4}D^2\bar{w} - \delta I\right]. \quad (6.13)$$

Since $\det(D^2\bar{w}) = 1$, an upper bound for the eigenvalues of $D^2\bar{w}$, given in Proposition 6.4, implies a lower bound. In fact,

$$D^2\bar{w}(x) \geq \left(\frac{1}{C'_n \text{dist}(x, \partial V)^{1-n}}\right)^{n-1} I \geq \frac{1}{C'_n} \left(\frac{d}{\lambda}\right)^{\beta(n-1)^2} I > 4\delta I \quad \text{for all } x \in [\Omega]. \quad (6.14)$$

For this choice of δ , looking at (6.13), we obtain that A^+ is positive definite in Ω . Furthermore, (6.14) together with the second inequality in (6.12) gives

$$\inf_{[\Omega]} D^2\bar{w} > \delta I + \frac{C}{\lambda} I \left(\bar{m} + \delta \|id\|_{L^\infty([\Omega])} + 2n\right).$$

implying that w^+ satisfies (5.1), i.e., w^+ is in the admissible class.

Now we check that w^+ is indeed a supersolution. It remains to show that

$$M_\lambda[w^+] < f \quad (6.15)$$

for

$$M_\lambda[w^+] = g\left(x + \nabla w - \delta x + G(\nabla w - \delta x)\right) \det(A^+).$$

We compute

$$(I + H)(I + D^2w) = A^+ + \frac{\delta}{2}I + \left[\frac{\delta}{2}I + (1 + \delta)H\right]. \quad (6.16)$$

Because $|H| \leq \frac{\delta}{4}$ by the choice of δ from (6.12), we have that the term $\frac{\delta}{2}I + (1 + \delta)H$ is positive definite. On the other hand, given any two positive definite matrices A_1, A_2 ,

$$\det \frac{1}{n}(A_1 + A_2) \geq \det \frac{1}{n}(A_1) + \det \frac{1}{n}(A_2).$$

Consequently, from (6.16) we obtain that

$$1 + \frac{\delta}{4} \geq \det \frac{1}{n}(I + H) \geq \det \frac{1}{n}(A^+) + \frac{\delta}{2},$$

yielding

$$\det(A^+) \leq \left(1 - \frac{\delta}{4}\right)^n \leq 1 - \frac{n}{8}\delta.$$

Finally, using the bounds for f, g from (6.10) and (6.11),

$$\begin{aligned} M_\lambda[w^+] &= g\left(x + \nabla w - \delta x + G(\nabla w - \delta x)\right) \det(A^+) \\ &\leq \left(1 + \theta(2\bar{m} + 2n)^\alpha\right) \left(1 - \frac{n}{8}\delta\right) < 1 - \theta \leq f \end{aligned}$$

given a suitable choice of

$$\delta \geq \frac{16\theta}{n}(2\bar{m} + 2n)^\alpha. \quad (6.17)$$

Then (6.15) follows as desired.

Step 2. Construction of a subsolution. We now check that w^- is a subsolution. As w^- is obviously in the admissible class for the operator M_λ in Ω , we only need to show $M_\lambda[w^-] > f$. We compute

$$M_\lambda[w^-] = g(x + \nabla w + \delta x + G(\nabla w + \delta x)) \det(A^-)$$

for

$$A^- := I + D^2w + \delta I + (D^2w + \delta I)H(\nabla w + \delta x).$$

If we choose δ satisfy (6.12), then we know that the following matrices $D^2\bar{w}$, $\frac{\delta}{4}I + H$ and $\frac{\delta}{2}I - H + \delta H$ are positive definite because of the estimate $|H(\nabla w + \delta x)| \leq \frac{C(\bar{m}+2n)}{\lambda} \leq \frac{\delta}{4}$. Consequently,

$$A^- = \left(1 - \frac{\delta}{4}\right) D^2\bar{w} + \left(\frac{\delta}{4}I + H\right) D^2\bar{w} + \frac{\delta}{2}I + \left[\frac{\delta}{2}I - H + \delta H\right] \geq \left(1 - \frac{\delta}{4}\right) D^2\bar{w} + \frac{\delta}{2}I.$$

Thus the matrix A^- is positive definite. Moreover by using the bounds for f, g from (6.10) and (6.11) we obtain the desired conclusion

$$M_\lambda[w^-] \geq \left(1 - \theta(2\bar{m} + 2n)^\alpha\right) \left(1 + \frac{\delta}{4}\right)^n \geq \left(1 - \theta(2\bar{m} + 2n)^\alpha\right) \left(1 + \frac{n\delta}{4}\right) > 1 + \theta \geq f,$$

where δ is chosen as in (6.17). Note that in the above estimate for w^- we actually do not need a uniform estimate of $D^2\bar{w}$ as in the case of w^+ .

Step 3. Choice of δ .

We now check that all the conditions on δ given in (6.12), (6.14) and (6.17) are compatible. Let us take

$$\delta := \frac{1}{4C'_n} \left(\frac{d}{\lambda}\right)^{\beta(n-1)^2} \quad (6.18)$$

which guarantees (6.14). To see that this choice of δ also satisfies (6.12) and (6.17), we first observe that Proposition 6.3 implies

$$\bar{m} \leq C_n \text{dist}([\Omega], \partial V)^{\frac{2}{n}-1} \leq C_n \left(\frac{d}{\lambda}\right)^{\frac{2-n}{n}\beta}.$$

Therefore, both (6.12) and (6.17) hold provided that

$$C_n \frac{1}{\lambda} \left(\frac{d}{\lambda}\right)^{\frac{2-n}{n}\beta} \leq \left(\frac{d}{\lambda}\right)^{\beta(n-1)^2} \quad \text{and} \quad C_n \theta \left(\frac{d}{\lambda}\right)^{\frac{2-n}{n}\alpha\beta} \leq \left(\frac{d}{\lambda}\right)^{\beta(n-1)^2}.$$

But these are true thanks to the hypothesis (6.7).

Step 4. Completion of the proof. The proof of the theorem is completed using the comparison principle for viscosity solutions of Proposition 6.1 and the initial hypothesis on u , which guarantee that

$$(u - w)(x) \leq \max_{\partial\Omega} \{u - w^+\} - \frac{\delta}{2} |x|^2, \quad \text{for all } x \in \Omega, \quad (6.19)$$

and

$$\min_{\overline{\Omega} \setminus \Omega^\gamma} \{u - w^-\} + \frac{\delta}{2} |x|^2 \leq (u - w)(x), \quad \text{for all } x \in \Omega. \quad (6.20)$$

Let us estimate the boundary values for $u - w^+$ and $u - w^-$. Recall that $\bar{w}(x) = \frac{1}{2} |x|^2 + w(x)$ and $u(x) = -c_\lambda(x) = -\frac{1}{2} |x|^2 - \frac{1}{\lambda} E(x)$ on $\partial\Omega$. Hence

$$u(x) - w^+(x) = u(x) - w(x) + \frac{\delta}{2} |x|^2 = -\bar{w}(x) - \frac{1}{\lambda} E(x) + \frac{\delta}{2} |x|^2 \quad \text{on } \partial\Omega.$$

We are going to use Proposition 6.3 to estimate the behavior of \bar{w} near ∂V . In particular, for $x \in \partial\Omega$,

$$\bar{w}(x) \geq -C \operatorname{dist}(x, \partial V)^{\frac{2}{n}} \geq -C \left(\frac{d}{\lambda} + \left(\frac{d}{\lambda} \right)^\beta \right)^{\frac{2}{n}} \geq -C \left(\frac{d}{\lambda} \right)^{\frac{2\beta}{n}},$$

where we have used the assumption (6.3) and the definition of V in the second inequality. It follows that

$$u(x) - w^+(x) \leq C \left(\frac{d}{\lambda} \right)^{\frac{2\beta}{n}} + \frac{C}{\lambda} + C\delta \leq C \left(\frac{d}{\lambda} \right)^{\frac{2\beta}{n}} \quad \text{for } x \in \partial\Omega \quad (6.21)$$

by our choice of δ in (6.18). From (6.19) and (6.21) we conclude (6.8).

To estimate the values of $u - w^-$ on $\overline{\Omega} \setminus \Omega^\gamma$, let $\bar{u} := u + c_\lambda$. Then for $x_1, x_2 \in \Omega$, by taking $y_i \in \partial_c u(x_i)$ it is easy to see that $|\bar{u}(x_1) - \bar{u}(x_2)| \leq \max\{|y_1|, |y_2|\} |x_1 - x_2| + \frac{C}{\lambda}$. Hence it follows from the definition of m that

$$|\bar{u}(x) - \bar{u}(y)| \leq m|x - y| + \frac{C}{\lambda} \quad \text{for all } x, y \in \Omega \setminus \Omega^\gamma.$$

As a consequence of this and as $u \geq \bar{u} - \frac{1}{2} |x|^2 - \frac{C}{\lambda}$, we obtain

$$u(x) - w(x) \geq \bar{u}(x) - \bar{w}(x) - \frac{C}{\lambda} \geq -\gamma m - \frac{C}{\lambda} \quad \forall x \in \Omega \setminus \Omega^\gamma \quad (6.22)$$

taking into account that $\bar{u} \equiv 0$ on $\partial\Omega$ and $\bar{w} \leq 0$. Hence, we have for $x \in \overline{\Omega} \setminus \Omega^\gamma$,

$$u(x) - w^-(x) = u(x) - w(x) - \frac{\delta}{2} |x|^2 \geq -\gamma m - \frac{C}{\lambda} - C \frac{\delta}{2} \geq -\gamma m - C\delta.$$

This together with (6.20) yields (6.9) and the theorem is proved. \square

Remark. In the proof of Theorem 6.2, it might happen that $x - \nabla c_\lambda^*(-\nabla w^\pm(x))$ does not belong to the domain of g for some x in Ω . Therefore, $M_\lambda[w^\pm]$ is not well defined at those x . However we do not have any problem since to apply Proposition 6.1 in Step 4 above we only need to check $M_\lambda[w^+] < f$ and $M_\lambda[w^-] > f$ at points x where $x - \nabla c_\lambda^*(-\nabla w^\pm(x)) \in \overline{\Omega'}$.

6.2 Sections are round

Here we compare sections $S_\mu(u, x_0)$ of u to those of a solution of a Monge-Ampère equation. Let $[\Omega]$ be the convex hull of Ω and w be the solution of

$$\begin{cases} \det(I + D^2w) = 1 & \text{in } [\Omega] \\ w = -\frac{1}{2}|x - y_0|^2 & \text{on } \partial[\Omega], \end{cases} \quad (6.23)$$

which is given by $\bar{w} = w + \frac{1}{2}|x - y_0|^2$ where \bar{w} is the unique convex solution of $\det D^2\bar{w} = 1$ in $[\Omega]$ and $\bar{w} = 0$ on $\partial[\Omega]$.

Note that in Theorem 4.7 we already have a preliminary control of the sections of u . However, we would like to obtain a more precise estimate using the comparison principle from Theorem 6.2, that allows us to say that w is a good approximation for u , and this approximation may be quantified explicitly. In the following three lemmas, we compare sections of u and w by assuming in advance that $\|u - w\|$ is small. We remind the reader that the μ -section of u was defined in (2.12) and (2.13), while the sections for w are defined in (4.13) and (4.14).

Lemma 6.5. *Let $B_1 \subset \Omega \subset [\Omega] \subset B_n$ be a domain. Let $u \in \mathcal{C}(\bar{\Omega})$ be a c_λ -convex function and $y_0 \in \partial_c u(\Omega)$. Consider w be the solution of (6.23), and \bar{u}, \bar{w} be defined as above. Assume that we have proved that*

$$\|\bar{u} - \bar{w}\|_{L^\infty(\Omega)} \leq \epsilon \quad (6.24)$$

for some $\epsilon > 0$. There exist some dimensional constants $\delta_n > 0$, $\mu_0 > 0$, $\tau_0 > 0$ and $C > 0$ such that if

$$\text{dist}(x, \partial[\Omega]) \leq \frac{\delta_n}{2} \text{ for all } x \in \partial\Omega \quad (6.25)$$

and $x_0 \in \Omega$ is a minimum point of \bar{u} in Ω , then we have

$$T_{\mu - C\epsilon^{1/2}}(w, x_0) \subset S_\mu(u, x_0) \subset T_{\mu + C\epsilon^{1/2}}(w, x_0) \subset \Omega_{\frac{\delta_n}{4}} \quad (6.26)$$

for all $\mu \leq \mu_0$, $\epsilon \leq \tau_0\mu^2$.

Proof. Recall that \bar{w} is the convex function satisfying $\det D^2\bar{w} = 1$ in $[\Omega]$ and $\bar{w} = 0$ on $\partial[\Omega]$. Let $\bar{x} \in [\Omega]$ be the point where \bar{w} attains its minimum value in $[\Omega]$; it exists because \bar{w} is a strictly convex function (see [6]). We would like to say that \bar{x} and x_0 stay away from the boundary of Ω . Indeed, [22, Proposition 3.2.3] gives an estimate for the value of \bar{w} at its minimum, $|\bar{w}(\bar{x})| \approx C_n$, C_n dimensional constant, and from Alexandroff's maximum principle [22, Theorem 1.4.2], we get that $\text{dist}(\bar{x}, \partial[\Omega]) \geq \delta_n$, δ_n universal constant. It then follows from the assumption (6.25) that $\text{dist}(\bar{x}, \partial\Omega) \geq \frac{\delta_n}{2}$. Consequently, we obtain from (6.24) that $\bar{w}(\bar{x}) - \epsilon \leq \bar{w}(x_0) - \epsilon \leq \bar{u}(x_0) \leq \bar{u}(\bar{x}) \leq \bar{w}(\bar{x}) + \epsilon$, i.e.,

$$|\bar{u}(x_0) - \bar{w}(\bar{x})| \leq \epsilon. \quad (6.27)$$

Since $\bar{w}(x_0) \leq \bar{w}(\bar{x}) + 2\epsilon \leq -C_n$ if ϵ is small enough, we conclude from Alexandroff's maximum principle that $\text{dist}(x_0, \partial[\Omega]) \geq \delta_n$. In particular, $\text{dist}(x_0, \partial\Omega) \geq \frac{\delta_n}{2}$ by (6.25).

In $[\Omega]_{\delta_n} := \{x \in [\Omega] : \text{dist}(x, \partial[\Omega]) \geq \delta_n\}$, we can use Pogorelov's estimates [22, formula (4.2.6)], to get that

$$\frac{2}{C_2^2}I \leq D^2\bar{w}(x) \leq \frac{2}{C_1^2}I \text{ for all } x \in [\Omega]_{\delta_n}. \quad (6.28)$$

With these in mind, it is known (see [23], for instance) that

$$|\nabla \bar{w}(x_0)| \leq C\epsilon^{1/2}. \quad (6.29)$$

We present the proof of (6.29) here for the sake of clarity. First of all, Taylor estimates give

$$\bar{w}(x_0) - \bar{u}(x_0) = \bar{w}(\bar{x}) - \bar{u}(x_0) + \frac{1}{2} \langle D^2 w(\xi)(x_0 - \bar{x}), x_0 - \bar{x} \rangle \quad (6.30)$$

for some ξ between x_0 and \bar{x} . Therefore, since $x_0, \bar{x} \in [\Omega]_{\delta_n}$, we can use (6.27) and Pogorelov estimates (6.28) in (6.30) to obtain

$$|\bar{x} - x_0| \leq C\epsilon^{1/2}. \quad (6.31)$$

On the other hand,

$$D_i \bar{w}(x_0) - D_i \bar{w}(\bar{x}) = - \int_0^1 D(D_i \bar{w})(x_0 + t(\bar{x} - x_0)) \cdot (\bar{x} - x_0) dt,$$

thus using (6.31) we obtain the desired (6.29).

Now, we would like to say that the sections of w do not approach the boundary of Ω . Using Alexandroff principle again we know that $T_\mu(w, \bar{x}) \subset \Omega_{\frac{\delta_n}{4}}$ for $0 < \mu < C_n$. Then, because of (6.24), (6.27) and (6.29), there exist τ_0, μ_0 universal constants such that

$$T_\mu(w, x_0) \subset T_{C\epsilon^{1/2} + \mu}(w, \bar{x}) \subset \Omega_{\frac{\delta_n}{4}}$$

for $\epsilon \leq \tau_0 \mu^2$ and $\mu \leq \mu_0$.

Next, we compare the sections of u and w . Let $x \in S_\mu(u, x_0)$. Then, using (2.13), (6.24) and (6.29), it follows,

$$\begin{aligned} \bar{w}(x) &\leq \bar{u}(x) + \epsilon \leq \bar{u}(x_0) + \mu + \epsilon \\ &\leq \bar{w}(x_0) + \langle \nabla \bar{w}(x_0), x - x_0 \rangle + \mu + 2\epsilon - \langle \nabla \bar{w}(x_0), x - x_0 \rangle \\ &\leq \bar{w}(x_0) + \langle \nabla \bar{w}(x_0), x - x_0 \rangle + \mu + 2\epsilon + 2C\epsilon^{1/2} \\ &\leq \bar{w}(x_0) + \langle \nabla \bar{w}(x_0), x - x_0 \rangle + \mu + C\epsilon^{1/2}. \end{aligned}$$

Hence, $x \in T_{\mu + C\epsilon^{1/2}}(\bar{w})$ and the second inclusion of (6.26) is proved. The other inclusion is very similar. □

We can also show that u is close to a quadratic polynomial:

Lemma 6.6. *In the hypothesis of the previous lemma, there exist a positive definite matrix $M = A^t A$ and a vector $p \in \mathbb{R}^n$ satisfying*

$$\det M = 1, \quad 0 < C_1 I \leq M \leq C_2 I, \quad \text{and} \quad |p| \leq C\epsilon^{1/2},$$

such that we have

$$B_{(1 - C(\mu^{1/2} + \mu^{-1}\epsilon^{1/2}))\sqrt{2}} \subset \mu^{-1/2} T S_\mu(u, x_0) \subset B_{(1 + C(\mu^{1/2} + \mu^{-1}\epsilon^{1/2}))\sqrt{2}} \quad (6.32)$$

and

$$|\bar{u}(x) - (\bar{u}(x_0) + \langle p, x - x_0 \rangle + \frac{1}{2} \langle M(x - x_0), x - x_0 \rangle)| \leq C(\mu^{3/2} + \epsilon) \quad (6.33)$$

for $x \in S_\mu(u, x_0)$, where $Tx = A(x - x_0)$ and $\bar{u}(x) := u(x) + c(x - y_0)$. Or equivalently for u ,

$$|u(x) - (u(x_0) + \langle p - x_0 + y_0, x - x_0 \rangle + \frac{1}{2} \langle (M - I)(x - x_0), x - x_0 \rangle)| \leq C(\mu^{3/2} + \frac{\mu^{1/2}}{\lambda} + \epsilon).$$

Proof. The behavior of the sections of the function w follows from the following three claims (that are proven in [23]).

1. There exist τ_0 and μ_0, δ_n , dimensional constants, such that, for $\epsilon \leq \tau_0 \mu^2$ and $\mu \leq \mu_0$,

$$T_\mu(w, x_0) \subset [\Omega]_{\delta_n}.$$

2. There exists μ_n such that if $\mu \leq \mu_n$ and $\gamma \leq \frac{3}{4}\mu$, then

$$\partial T_{\mu+\gamma}(w, x_0) \subset \mathcal{N}_{\frac{C\gamma}{\sqrt{\mu}}}(\partial T_\mu(w, x_0)) \quad \text{and} \quad \partial T_{\mu-\gamma}(w, x_0) \subset \mathcal{N}_{\frac{C\gamma}{\sqrt{\mu}}}(\partial T_\mu(w, x_0)). \quad (6.34)$$

3. Let $M = D^2 \bar{w}(x_0)$ and $E = \{x : \frac{1}{2} \langle D^2 \bar{w}(x_0)(x - x_0), x - x_0 \rangle \leq 1\}$. We compare $T_\mu(w, x_0)$ with ellipsoids and claim that

$$\partial T_\mu(w, x_0) \subset \mathcal{N}_{C\mu}(\partial \mu^{1/2} E), \quad (6.35)$$

for some structural constant C and all $0 < \mu \leq \mu_n$. Here the dilation is with respect to the point x_0 .

We set $\mu_0 := \min\{\mu_0, \mu_n\}$, and let $\mu \leq \mu_0$, $\epsilon \leq \tau_0 \mu^2$. From (6.26) and (6.34), we obtain

$$\partial S_\mu(u, x_0) \subset \mathcal{N}_{C\mu^{-1/2}\epsilon^{1/2}}(\partial T_\mu(w, x_0)), \quad (6.36)$$

which together with (6.35) gives

$$\partial S_\mu(u, x_0) \subset \mathcal{N}_{C(\mu+\mu^{-1/2}\epsilon^{1/2})}(\partial \mu^{1/2} E). \quad (6.37)$$

We also notice that

$$\partial \mu^{1/2} E = x_0 + A^{-1} \left(\partial B_{\sqrt{2\mu}} \right), \quad (6.38)$$

where $M = A^t A$. If we let $Tx = A(x - x_0)$, then we have that

$$T \left(\mathcal{N}_\delta \left(\partial \mu^{1/2} E \right) \right) \subset B_{\sqrt{2\mu} + \|A\|\delta}, \quad (6.39)$$

and

$$T \left(\mathcal{N}_\delta \left(\partial \mu^{1/2} E \right) \right) \subset \left(B_{\sqrt{2\mu} - \|A\|\delta} \right)^c.$$

Hence taking $\delta = C(\mu + \mu^{-1/2}\epsilon^{1/2})$ from (6.37) we get (6.32).

And for the second part of the lemma, using (6.32), the fact that $\|A\|$ is bounded, and that $\epsilon \leq \tau_0 \mu^2$, we obtain $S_\mu(u, x_0) \subset B_{C\mu^{1/2}}(x_0)$. Then, taking $p = \nabla \bar{w}(x_0)$ yields

$$\begin{aligned} & \left| \bar{u}(x) - \left(\bar{u}(x_0) + \langle p, x - x_0 \rangle + \frac{1}{2} \langle M(x - x_0), x - x_0 \rangle \right) \right| \\ & \leq |\bar{u}(x) - \bar{w}(x)| + \left| \bar{w}(x) - \bar{w}(x_0) - \langle p, x - x_0 \rangle - \frac{1}{2} \langle M(x - x_0), x - x_0 \rangle \right| + |\bar{w}(x_0) - \bar{u}(x_0)| \\ & \leq 2\epsilon + \|D^3 \bar{w}\|_{L^\infty(B_{C\mu^{1/2}})} |x - x_0|^3 \leq C(\epsilon + \mu^{3/2}), \end{aligned}$$

as desired. Notice that $|p| \leq C\epsilon^{1/2}$ by (6.29). \square

If the boundary of the convex hull of Ω is close to that of $B_{\sqrt{2}}$, then we can get better estimates for M and ∂S_μ :

Lemma 6.7. *In addition to the hypothesis of Lemma 6.5, suppose that Ω is a domain satisfying*

$$B_{(1-\sigma)\sqrt{2}} \subset [\Omega] \subset B_{(1+\sigma)\sqrt{2}}$$

for some $0 < \sigma \leq 1$. Then, there exist $\mu_0 > 0$, $\tau_0 > 0$ which are independent of σ , a positive definite matrix $M = A^t A$, and $p \in \mathbb{R}^n$ with

$$\det M = 1, \quad (1 - C\sigma)|\xi|^2 \leq \langle M\xi, \xi \rangle \leq (1 + C\sigma)|\xi|^2, \quad |p| \leq C\epsilon^{1/2} \text{ and } |p - x_0| \leq C\sigma \quad (6.40)$$

such that for $0 < \mu \leq \mu_0$ and $\epsilon \leq \tau_0\mu^2$ we have

$$B_{(1-C(\sigma\mu^{1/2}+\mu^{-1}\epsilon^{1/2}))\sqrt{2}} \subset \mu^{-1/2}TS_\mu(u, x_0) \subset B_{(1+C(\sigma\mu^{1/2}+\mu^{-1}\epsilon^{1/2}))\sqrt{2}}, \quad (6.41)$$

and in $S_\mu(u, x_0)$,

$$\left| \bar{u}(x) - (\bar{u}(x_0) + \langle p, x - x_0 \rangle + \frac{1}{2}\langle M(x - x_0), x - x_0 \rangle) \right| \leq C(\sigma\mu^{3/2} + \epsilon), \quad (6.42)$$

where $Tx = A(x - x_0)$ and $\bar{u}(x) := u(x) + c(x - y_0)$. Or equivalently,

$$\left| u(x) - (u(x_0) + \langle p - x_0 + y_0, x - x_0 \rangle + \frac{1}{2}\langle (M - I)(x - x_0), x - x_0 \rangle) \right| \leq C(\sigma\mu^{3/2} + \frac{\mu^{1/2}}{\lambda} + \epsilon)$$

in $S_\mu(u, x_0)$.

Proof. It is similar to that of Lemma 6.6. The only difference is that since now $\partial[\Omega] \subset \mathcal{N}_{\sigma\sqrt{2}}(\partial B_{\sqrt{2}})$, we will get the following improvement of (6.35):

$$\partial T_\mu(w, x_0) \subset \mathcal{N}_{C\sigma\mu}(\partial\mu^{1/2}E). \quad (6.43)$$

To do this, let $P(x) = \frac{1}{2}|x|^2 - 1$, then $\det D^2P(x) = 1$. Since $\partial[\Omega] \subset \mathcal{N}_{\sigma\sqrt{2}}(\partial B_{\sqrt{2}})$, we have $P - 3\sigma \leq 0 \leq P + 3\sigma$ on $\partial[\Omega]$ and by the comparison principle we get $\|\bar{w} - P\|_{L^\infty([\Omega])} \leq 3\sigma$. Since \bar{w} is smooth, $v := \bar{w} - P$ satisfies the following uniformly elliptic linear equation

$$\text{trace}(D(x)D^2v) = 0,$$

where $D(x) = \int_0^1 (tD^2\bar{w}(x) + (1-t)I)^{-1} \det(tD^2\bar{w}(x) + (1-t)I) dt$. Therefore, by interior Schauder estimates

$$\|\bar{w} - P\|_{C_{\text{loc}}^2} \leq C\|\bar{w} - P\|_{L_{\text{loc}}^\infty} \leq C\sigma.$$

In particular, $|p - x_0| = |D\bar{w}(x_0) - DP(x_0)| \leq C\sigma$ and $|D_{ij}\bar{w}(x_0) - \delta_{ij}| \leq C\sigma$ for $1 \leq i, j \leq n$. By differentiating $\det D^2\bar{w}(x) = 1$, we obtain that the function $v = D(\bar{w} - P)$ satisfies the linearized equation

$$\text{trace}((D^2\bar{w}(x))^{-1}D^2v) = 0.$$

Again by the interior Schauder estimates (C_{loc}^2 estimates) we get

$$\|D^3\bar{w}\|_{L_{\text{loc}}^\infty} = \|D^3(\bar{w} - P)\|_{L_{\text{loc}}^\infty} \leq C\|D(\bar{w} - P)\|_{L_{\text{loc}}^\infty} \leq C^*\sigma. \quad (6.44)$$

The rest of the proof is standard and follows as in [23, Lemma 1.2]. \square

Another ingredient that will be used in the iteration process in Section 7 is the following gradient bound for the sections:

Proposition 6.8. *Under the same assumptions and notations as in Lemma 6.5, we have*

$$\partial_c^{loc} u(S_\mu) \subset B_{\frac{C(\mu+\epsilon+\frac{1}{\lambda})}{\sqrt{\mu+\epsilon}}}(y_0)$$

for all $\mu \leq \mu_0$ and $\epsilon \leq \tau_0 \mu$.

Proof. Let $\bar{x} \in [\Omega]$ be the minimum point of \bar{w} in $[\Omega]$. Then we obtain from the regularity of w and the proofs of Lemma 6.5 and Proposition 4.4 that $S_\mu \subset T_{\mu+2\epsilon}(w, \bar{x}) \subset B_{C\sqrt{\mu+\epsilon}}(\bar{x}) \subset \Omega_{\delta_n}$. Moreover since $\|\bar{u} - \bar{w}\|_{L^\infty(\Omega)} \leq \epsilon$, we have $\|u - w\|_{L^\infty(\Omega)} \leq \epsilon + \frac{C}{\lambda}$. Thus we can apply Lemma 3.4 with $\epsilon \rightsquigarrow \epsilon + \frac{C}{\lambda}$, $R := 2C\sqrt{\mu+\epsilon}$ and $\delta := (\epsilon + \frac{C}{\lambda})/C\sqrt{\mu+\epsilon}$. The lemma then follows from the same arguments as in the proof of Corollary 4.6. \square

7 $\mathcal{C}^{2,\alpha}$ estimates

Here we present the proof of the main theorem (Theorem 1.1). Let u be a solution of (1.1), or equivalently (thanks to the results in Subsection 2.2), of

$$\begin{cases} M_\lambda[u] = f & \text{in } \Omega, \\ \partial_{c_\lambda} u(\Omega) = \Omega' \end{cases} \quad (7.1)$$

for some $\lambda > 0$. Fix $\delta_0 > 0$, and let

$$\Omega_{\delta_0} := \{x \in \Omega : \text{dist}(x, \partial\Omega) > \delta_0\}.$$

We will find λ_0 and $r > 0$ depending on δ_0 and α such that for all $\lambda \geq \lambda_0$, and for every point $x_0 \in \Omega_{\delta_0}$ it is possible to construct a second order polynomial in x , call it P_0 , such that

$$|u(x) - P_0(x)| \leq C|x - x_0|^{\alpha'} \quad \forall x \in B_r(x_0), \quad (7.2)$$

for some $\alpha' \in (0, 1)$ depending only on n , α and chosen later, and C depending only on λ_0 , n , Λ_1 , Λ_2 , α , δ_0 , $\|f\|_{\mathcal{C}^\alpha(\overline{\Omega_{\delta_1}})}$, $\|g\|_{\mathcal{C}^\alpha(\overline{\Omega'_{\delta_1}})}$, Ω and Ω' . The constant $\delta_1 > 0$ will be made precise in the proof and depends only on δ_0 and universal constants.

The proof goes by induction. In the first step, we will compare the function u with the solution of an optimal transport problem with quadratic cost. In the successive steps, once we have localized the problem, the comparison will be with respect to a solution of a Monge-Ampère equation with homogeneous Dirichlet boundary data.

7.1 Localization

We first localize the problem in a section so that we have control of the boundary data. Fix any $x_0 \in \Omega_{\delta_0}$ and $y_0 \in \partial_c u(x_0)$. The notion of sections for u at x_0 was defined in (2.12), or equivalently for $\bar{u}(x) := u(x) + c(x - y_0)$, as in (2.13). Note that x_0 is a minimum point for \bar{u} in Ω .

We make now one further simplification. It is enough to assume that $x_0 = y_0$, i.e., that $p_0 := -\nabla c(x_0 - y_0) = 0$. This can be achieved by subtracting a linear function to the potential u .

This first step in the induction argument was considered in Section 4. In fact, Proposition 4.1 shows that u is actually close in the L^∞ norm to a strictly convex function w . It also implies that the sections of u stay away from the boundary $\partial\Omega$. This is done in Proposition 4.4 and Corollary 4.5. These imply, in particular, that there exists a positive constant $\mu_1 = \mu_1(\delta_0, n, \Lambda_1, \Lambda_2, \Omega, \Omega')$ satisfying: for all $h \leq \mu_1$ we can find $\lambda_0 > 0$ depending on h such that for all $\lambda \geq \lambda_0$, if u is the solution of (7.1) for the cost c_λ , then

$$S_h(u, x_0) := S_h(u, x_0, y_0) \subset \Omega_{c\delta_0^\theta}.$$

Now we look at the image set. Note that Proposition 4.3 gives that $\partial_c u(\Omega_{c\delta_0^\theta}) \subset \Omega'_{\delta_1}$ for some small $\delta_1 > 0$ depending only on δ_0 and universal constants. Since Ω' is strongly convex by the assumption, so is Ω'_{δ_1} and hence as in Lemma 2.3 we conclude that Ω'_{δ_1} is c_λ^* -convex with respect to Ω . It then follows from [27] that $\partial_c^{\text{loc}} u(\Omega_{c\delta_0^\theta}) \subset \overline{\Omega'_{\delta_1}}$ which allows to restrict the dependence on g in the arguments to $\|g\|_{C^\alpha(\overline{\Omega'_{\delta_1}})}$, as stated in the main theorem.

Next, Corollary 4.5 and its subsequent remark (see formulas (4.18) and (4.19)) show that $S_h(u, x_0) \subset B_{C\sqrt{h}}(x_0)$ and there exists an affine transformation $Tx = Ax + b$ with $C_1 \leq |\det A|^{\frac{2}{n}} h \leq C_2$, $\|A\| \leq Ch^{\frac{-1}{2}}$ and $\|A^{-1}\| \leq Ch^{\frac{1}{2}}$ such that

$$B_K(0) \subset \tilde{\Omega} := TS_h(u, x_0) \subset B_n(0) \quad \text{and} \quad \text{dist}(x, \partial[\tilde{\Omega}]) \leq \frac{Ch^{\frac{-1}{2}}}{\lambda} \quad \text{for all } x \in \partial\tilde{\Omega}.$$

Let us consider the function

$$\tilde{u}(x) := \frac{1}{h} [u(T^{-1}x) - u(x_0) - c(x_0 - y_0) - h], \quad \text{for } x \in \tilde{\Omega}.$$

We expect that \tilde{u} satisfies the Monge-Ampère equation:

$$\begin{cases} \tilde{g}(x - \nabla \tilde{c}^*(-\nabla \tilde{u})) \det(I + D^2 \tilde{c}^*(-\nabla \tilde{u}) D^2 \tilde{u}) = \tilde{f}(x) & \text{in } \tilde{\Omega}, \\ \tilde{u}(x) = -\tilde{c}(x - Ty_0) & \text{on } \partial\tilde{\Omega}, \end{cases} \quad (7.3)$$

for $\tilde{f}(x) := f(T^{-1}x)$, $\tilde{g}(y) := g(T^{-1}y)$ and

$$\tilde{c}(x) := \frac{1}{h} c(T^{-1}x) = \frac{1}{2}|x|^2 + \frac{1}{\tilde{\lambda}} \tilde{E}(x) \quad \text{with} \quad \tilde{\lambda} := h^{\frac{-1}{2}} \lambda.$$

Equivalently, $\tilde{c}^*(z) = \frac{1}{h} c^*(hT^t z)$ by the Legendre transform.

We know from Lemma 2.3 that u is a weak solution of (7.1) in the Alexandroff sense. However, its restriction to the section $S_h(u, x_0)$ may not be an Alexandroff weak solution, and consequently, neither \tilde{u} is a weak solution of (7.3). This is, in general, the main difficulty one needs to overcome. In this regard, Proposition 5.2 easily shows that u is a viscosity subsolution of $M_\lambda[u] = f$ in $S_h(u, x_0)$. But the crucial ingredient is Corollary 5.4. As $\|\partial_c u(S_h(u, x_0))\| \leq C$ by Corollary 4.6, we obtain from Theorem 4.7 and Corollary 5.4 that u is also a viscosity supersolution of $M_\lambda[u] = f$ in $S_h(u, x_0)^{\frac{C}{\sqrt{\lambda}}}$ relative to $S_h(u, x_0)$. Therefore by rescaling and using $\|A\| \leq Ch^{\frac{-1}{2}}$, we see that \tilde{u} is a viscosity subsolution in $\tilde{\Omega}$ and a viscosity supersolution in $\tilde{\Omega}^\gamma$ relative to $\tilde{\Omega}$ of equation (7.3) with $\gamma := \frac{Ch^{\frac{-1}{2}}}{\sqrt{\lambda}} = \frac{Ch^{\frac{-3}{4}}}{\sqrt{\lambda}}$. Another conclusion that Corollary 5.4 gives us is that \tilde{u} is an Alexandroff weak solution of (7.3) in $\tilde{\Omega}^\gamma$ relative to $\tilde{\Omega}$. We also have the following observation about the relationship between sections

of \tilde{u} and sections of u : $S_\rho(\tilde{u}, x_1, y_1) = TS_{\rho h}(u, T^{-1}x_1, T^{-1}y_1)$ for all $x_1 \in \tilde{\Omega}$, $y_1 \in \partial_{\tilde{c}}\tilde{u}(x_1)$ and $\rho > 0$ as long as $S_{\rho h}(u, T^{-1}x_1, T^{-1}y_1) \subset S_h(u, x_0)$. Hence for such sections, it follows from Theorem 4.7 that

$$S_\rho(\tilde{u}, x_1, y_1) \subset TB_{C\sqrt{\rho h}}(T^{-1}x_1) \subset B_{C\|T\|\sqrt{\rho h}}(x_1) \subset B_{C\sqrt{\rho}}(x_1). \quad (7.4)$$

We will use the comparison principle for viscosity solutions from Theorem 6.2 to get a better L^∞ approximation for \tilde{u} . In order to apply this theorem, we first need to verify the assumptions. Without loss of generality, we can assume that $x_0 = y_0 = 0$ and $f(x_0) = g(y_0) = 1$. We then have

$$|\tilde{f}(x) - 1| = |f(T^{-1}x) - f(0)| \leq |A^{-1}x|^\alpha \leq \|A^{-1}\|^\alpha |x|^\alpha \leq Ch^{\frac{\alpha}{2}}|x|^\alpha \quad \text{for all } x \in \tilde{\Omega}.$$

In particular, $\|\tilde{f} - 1\|_{L^\infty(\tilde{\Omega})} \leq Ch^{\frac{\alpha}{2}}$. On the other hand for $y \in \partial_c^{\text{loc}}\tilde{u}(\tilde{\Omega}) = T\partial_c^{\text{loc}}u(S_h(u, x_0))$, we write $y = Tx$ for $x \in \partial_c^{\text{loc}}u(S_h(u, x_0))$ and thus

$$|\tilde{g}(y) - 1| = |g(x) - g(0)| \leq C|x - 0|^\alpha = C|A^{-1}y|^\alpha \leq Ch^{\frac{\alpha}{2}}|y|^\alpha.$$

Therefore, we would like to apply Theorem 6.2 with $d = Ch^{-1}$, $\lambda \rightsquigarrow \tilde{\lambda}$, $\theta \rightsquigarrow \tilde{\theta} := Ch^{\frac{\alpha}{2}}$ and $\beta > 0$ satisfying $(\frac{d}{\lambda})^\beta = \tilde{\theta}^{\frac{n\hat{\alpha}}{\alpha}}$, where $\hat{\alpha} = \hat{\alpha}(n, \alpha)$ is chosen arbitrary as long as $0 < \hat{\alpha} < \frac{\alpha}{n((n-1)^2 + \frac{n-2}{n}\alpha)}$. Notice that $\beta \leq 1$ because we can assume without loss of generality that $\lambda \geq \lambda_0 \geq \frac{C}{h^{\frac{n\hat{\alpha}+2}{2}}}$ ensuring that $\frac{d}{\lambda} \leq \tilde{\theta}^{\frac{n\hat{\alpha}}{\alpha}}$. Observe also that condition (6.7) is satisfied since

$$\begin{aligned} & C_n \left(\frac{d}{\tilde{\lambda}}\right)^{-\beta((n-1)^2 + \frac{n-2}{n}\alpha)} \left[\tilde{\theta} + \frac{1}{\tilde{\lambda}} \left(\frac{d}{\tilde{\lambda}}\right)^{\frac{n-2}{n}(\alpha-1)\beta} \right] \\ & \leq C_n \tilde{\theta}^{-\frac{n\hat{\alpha}}{\alpha}((n-1)^2 + \frac{n-2}{n}\alpha)} \left[\tilde{\theta} + \frac{1}{d} \tilde{\theta}^{\frac{n\hat{\alpha}}{\alpha}} \tilde{\theta}^{\frac{\hat{\alpha}}{\alpha}(n-2)(\alpha-1)} \right] \leq 2C_n \tilde{\theta}^{1 - \frac{n\hat{\alpha}}{\alpha}((n-1)^2 + \frac{n-2}{n}\alpha)} \leq 1. \end{aligned}$$

Thus as $\|\partial_{\tilde{c}}\tilde{u}(\tilde{\Omega})\|_{L^\infty} \leq C$ by Corollary 4.6 and $\gamma = \frac{Ch^{-\frac{3}{4}}}{\sqrt{\tilde{\lambda}}}$, we conclude from Theorem 6.2 that

$$\|\tilde{u} - w\|_{L^\infty(\tilde{\Omega})} \leq C \max \left\{ \left(\frac{d}{\tilde{\lambda}}\right)^{\frac{2\beta}{n}}, \frac{h^{-\frac{3}{4}}}{\sqrt{\tilde{\lambda}}} \right\} = C \left(\frac{d}{\tilde{\lambda}}\right)^{\frac{2\beta}{n}} = C \tilde{\theta}^{\frac{2\hat{\alpha}}{\alpha}} = Ch^{\hat{\alpha}} =: \epsilon_0, \quad (7.5)$$

where w is the c_q -convex solution of

$$\begin{cases} \det(I + D^2w) = 1 & \text{in } V := \mathcal{N}_{\tilde{\theta}^{\frac{n\hat{\alpha}}{\alpha}}}[\Omega], \\ w = -\frac{1}{2}|x - Ty_0|^2 & \text{on } \partial V. \end{cases}$$

(Note that $Ty_0 \in \partial_{\tilde{c}}\tilde{u}(Tx_0)$). The above estimate holds for any $0 < h \leq \mu_1$. We now fix $h \leq \mu_1$ (and hence λ_0 is fixed) such that $\epsilon_0 = \tau_0\mu^3$, where $0 < \mu \leq \min\{\mu_0, 1/2\}$ is a universal constant whose precise value will be determined later and $\mu_0, \tau_0 > 0$ are the constants given in Lemma 6.5. In order to prove u is $\mathcal{C}^{2, \alpha'}$ at x_0 , it is enough to prove that its normalization \tilde{u} defined in $\tilde{\Omega}$ is $\mathcal{C}^{2, \alpha'}$ at Tx_0 . However for simplicity, from now on we will abuse notations by writing $u, c, \lambda, \Omega, f, g, x_0$ and y_0 for $\tilde{u}, \tilde{c}, \tilde{\lambda}, \tilde{\Omega}, \tilde{f}, \tilde{g}, Tx_0$ and Ty_0 respectively. In

particular, we have

$$B_K(0) \subset \Omega \subset B_n(0) \quad \text{and} \quad \text{dist}(x, \partial[\Omega]) \leq \frac{Ch^{-1}}{\lambda} \quad \text{for all } x \in \partial\Omega,$$

$$u \text{ is an Alexandroff weak solution in } \Omega^{\gamma_0} \text{ relative to } \Omega \text{ with } \gamma_0 := \frac{Ch^{-\frac{3}{4}}}{\sqrt{\lambda}}, \quad (7.6)$$

$$S_\rho(u, x_1, y_1) \subset B_{C\sqrt{\rho}}(x_1) \quad \text{for all } x_1 \in S_\mu(u, x_0), \quad y_1 \in \partial_c u(x_1) \text{ and } \rho > 0 \text{ small}, \quad (7.7)$$

$$|f(x) - 1| \leq \tilde{\theta}|x|^\alpha \quad \forall x \in \Omega \quad \text{and} \quad |g(y) - 1| \leq \tilde{\theta}|y|^\alpha \quad \forall y \in \partial_c^{\text{loc}} u(\Omega),$$

$$\frac{1}{\lambda} \leq h^{\frac{n\hat{\alpha}+3}{2}}.$$

Define

$$\bar{u}(x) := u(x) + c(x - y_0), \quad \text{for } x \in \Omega.$$

7.2 Induction

To show that u (equivalently \bar{u}) is $C^{2,\alpha'}$ at the point x_0 , we will compare smaller and smaller sections $S_\mu(u, x_0)$ of u with those of a solution w of a Monge-Ampère equation with Dirichlet boundary data. The precise inductive procedure will be carried out next.

First induction step. As a consequence of (7.5) and Lemma 6.6 we can find a transformation $T_1 x = A_1(x - x_0)$, where A_1 is a unitary matrix, such that $C_1|x|^2 \leq |A_1 x|^2 \leq C_2|x|^2$ and

$$B_{(1-\sigma_1)\sqrt{2}} \subset \mu^{-1/2} T_1 S_\mu(u, x_0) \subset B_{(1+\sigma_1)\sqrt{2}}, \quad \text{with } \sigma_1 := C(\mu^{1/2} + \mu^{-1}\epsilon_0^{1/2}). \quad (7.8)$$

Moreover,

$$|\bar{u}(x) - (\bar{u}(x_0) + \langle p_1, x - x_0 \rangle + \frac{1}{2} \langle M_1(x - x_0), x - x_0 \rangle)| \leq C(\mu^{3/2} + \epsilon_0) \quad \text{in } S_\mu(u, x_0),$$

where $M_1 := A_1^t A_1$ and p_1 is some vector satisfying $|p_1| \leq C\epsilon_0^{1/2}$.

Now consider the rescaled function

$$u_1(x) := \frac{1}{\mu} \left[u(\mu^{1/2} T_1^{-1} x) - u(x_0) - c(x_0 - y_0) - \mu \right], \quad x \in \Omega_1 := \mu^{-1/2} T_1 S_\mu(u, x_0).$$

If $u|_{S_\mu(u, x_0)}$ were a viscosity solution of $g(x - \nabla c^*(-\nabla u)) \det(I + D^2 c^*(-\nabla u) D^2 u) = f(x)$ in $S_\mu(u, x_0)$, then u_1 would be a viscosity solution of the following Monge-Ampère equation:

$$\begin{cases} g_1(x - \nabla c_1^*(-\nabla u_1)) \det(I + D^2 c_1^*(-\nabla u_1) D^2 u_1) = f_1(x) & \text{in } \Omega_1, \\ u_1(x) = -c_1(x - \mu^{-\frac{1}{2}} T_1 y_0) & \text{on } \partial\Omega_1, \end{cases} \quad (7.9)$$

for

$$f_1(x) := f(\mu^{1/2} T_1^{-1} x), \quad g_1(y) := g(\mu^{1/2} T_1^{-1} y)$$

and

$$c_1(x) := \frac{1}{\mu} c(\mu^{1/2} T_1^{-1} x) = \frac{1}{2} |x|^2 + \frac{1}{\lambda_1} E_1(x) \quad \text{with} \quad \frac{1}{\lambda_1} := \frac{\mu^{1/2}}{\lambda}.$$

The definition of c_1 implies that $c_1^*(z) = \frac{1}{\mu} c^*(\mu^{1/2} T_1^t z)$. It also becomes clear now from (7.9) that the boundary values of u_1 are controlled.

Proposition 5.2 establishes that u is a viscosity subsolution of $M_\lambda[u] = f$ in $S_\mu(u, x_0)$. Moreover by using (7.5) and the fact $Ch^{\hat{\alpha}} = \epsilon_0 = \tau_0\mu^3$, we obtain from Proposition 6.8 that $m_1 := \|\partial_{c_\lambda} u(S_\mu(u, x_0))\|_{L^\infty} \leq C\mu^{1/2}$. Hence (7.6), (7.7) and Corollary 5.4 yield that u is also a viscosity supersolution of $M_\lambda[u] = f$ in $S_\mu(u, x_0)^{\frac{C}{\sqrt{\lambda}}}$ relative to $S_\mu(u, x_0)$. When we rescale and using $\|A_1\| \leq \sqrt{C_2}$, we then know that u_1 is a viscosity subsolution in Ω_1 and viscosity supersolution solution in Ω_1^γ relative to Ω_1 of equation (7.9) for $\gamma := CC_2^{\frac{1}{2}}\mu^{\frac{-1}{2}}/\sqrt{\lambda} = CC_2^{\frac{1}{2}}\mu^{\frac{-3}{4}}/\sqrt{\lambda_1}$. Likewise u_1 is also an Alexandroff weak solution of (7.9) in Ω_1^γ relative to Ω_1 . Furthermore by using (7.7) and the same arguments leading to (7.4), we have

$$S_\rho(u_1, x_1, y_1) \subset \mu^{\frac{-1}{2}}T_1B_{C\sqrt{\rho\mu}}(\mu^{\frac{1}{2}}T_1^{-1}x_1) \subset B_{C\|A_1\|\sqrt{\rho}}(x_1) \quad (7.10)$$

for all $x_1 \in \Omega_1$, $y_1 \in \partial_{c_1} u_1(x_1)$ and $\rho > 0$ as long as $S_{\rho\mu}(u, \mu^{\frac{1}{2}}T_1^{-1}x_1, \mu^{\frac{1}{2}}T_1^{-1}y_1) \subset S_\mu(u, x_0)$.

Notice that we can apply Proposition 3.2 to say,

$$\text{dist}(x, \partial[S_\mu(u, x_0)]) \leq \frac{C}{\lambda} \quad \text{for all } x \in \partial S_\mu(u, x_0),$$

and after rescaling,

$$\text{dist}(x, \partial[\Omega_1]) \leq \mu^{\frac{-1}{2}}\|A_1\|\frac{C}{\lambda} \leq \frac{C\sqrt{C_2}\mu^{-1}}{\lambda_1} \quad \forall x \in \partial\Omega_1.$$

We have

$$|f_1(x) - 1| = |f(\mu^{\frac{1}{2}}T_1^{-1}x) - 1| \leq \tilde{\theta}\mu^{\frac{\alpha}{2}}|A_1^{-1}x|^\alpha \leq \tilde{\theta}(C_1^{-1}\mu)^{\frac{\alpha}{2}}|x|^\alpha = \theta_1|x|^\alpha \quad \forall x \in \Omega_1, \quad (7.11)$$

where $\theta_1 := \tilde{\theta}(C_1^{-1}\mu)^{\frac{\alpha}{2}} \leq \tilde{\theta}$. Also as $\mu^{\frac{1}{2}}T_1^{-1}\partial_{c_1}^{\text{loc}}u_1(\Omega_1) = \partial_c^{\text{loc}}u(S_\mu(u, x_0))$,

$$|g_1(y) - 1| = |g(\mu^{\frac{1}{2}}T_1^{-1}y) - 1| \leq \tilde{\theta}\mu^{\frac{\alpha}{2}}|A_1^{-1}y|^\alpha \leq \theta_1|y|^\alpha \quad \text{for all } y \in \partial_{c_1}^{\text{loc}}u_1(\Omega_1). \quad (7.12)$$

Therefore, we would like to apply Theorem 6.2 with $d = C\sqrt{C_2}\mu^{-1}$, $\lambda \rightsquigarrow \lambda_1$, $\theta \rightsquigarrow \theta_1$ and $\beta > 0$ satisfying $(\frac{d}{\lambda_1})^\beta = \theta_1^{\frac{n\hat{\alpha}}{\alpha}}$. Notice that $\beta \leq 1$ because our condition on $\hat{\alpha}$ yields

$$\frac{d}{\lambda_1} = \frac{C\sqrt{C_2}\mu^{1/2}\mu^{-1}}{\lambda} \leq C\sqrt{C_2}\mu^{1/2}\mu^{-1}h^{\frac{n\hat{\alpha}+2}{2}} \leq C\mu^{(\frac{3}{\alpha}-\frac{1+n\hat{\alpha}}{2})}\theta_1^{\frac{n\hat{\alpha}}{\alpha}} \leq \theta_1^{\frac{n\hat{\alpha}}{\alpha}} \quad (7.13)$$

and

$$\frac{1}{\lambda_1} = \frac{\mu^{\frac{1}{2}}}{\lambda} \leq Ch^{\frac{n\hat{\alpha}+1}{2}}(h\mu)^{\frac{(1-\alpha)}{2}[1-(n-2)\hat{\alpha}]} \theta_1^{\frac{\hat{\alpha}(n-2)(1-\alpha)}{\alpha}} \leq \theta_1^{\frac{\hat{\alpha}(n-2)(1-\alpha)}{\alpha}}. \quad (7.14)$$

Observe also that condition (6.7) is satisfied since

$$\begin{aligned} & C_n \left(\frac{d}{\lambda_1}\right)^{-\beta((n-1)^2 + \frac{n-2}{n}\alpha)} \left[\theta_1 + \frac{1}{\lambda_1} \left(\frac{d}{\lambda_1}\right)^{\frac{n-2}{n}(\alpha-1)\beta} \right] \\ & \leq C_n \theta_1^{-\frac{n\hat{\alpha}}{\alpha}((n-1)^2 + \frac{n-2}{n}\alpha)} \left[\theta_1 + \theta_1 \theta_1^{\frac{\hat{\alpha}(n-2)(1-\alpha)}{\alpha}} \theta_1^{\frac{\hat{\alpha}(n-2)(\alpha-1)}{\alpha}} \right] \\ & \leq 2C_n \theta_1^{1-\frac{n\hat{\alpha}}{\alpha}((n-1)^2 + \frac{n-2}{n}\alpha)} \leq 2C_n \tilde{\theta}^{1-\frac{n\hat{\alpha}}{\alpha}((n-1)^2 + \frac{n-2}{n}\alpha)} \leq 1. \end{aligned}$$

We have checked all the hypotheses necessary to apply the comparison principle from Theorem 6.2. Thus as $\|\partial_{c_1} u_1(\Omega_1)\|_{L^\infty} = \|\mu^{-1/2} T_1 \partial_c u(S_\mu(u, x_0))\|_{L^\infty} \leq C$ by Proposition 6.8 and $\gamma = CC_2^{\frac{1}{2}} \mu^{\frac{-3}{4}} / \sqrt{\lambda_1}$, we conclude from Theorem 6.2 that

$$\|u_1 - w_1\|_{L^\infty(\Omega_1)} \leq C \max \left\{ \left(\frac{d}{\lambda_1} \right)^{\frac{2\beta}{n}}, \frac{C_2^{\frac{1}{2}} \mu^{\frac{-3}{4}}}{\sqrt{\lambda_1}} \right\} = C \theta_1^{\frac{2\hat{\alpha}}{\alpha}} =: \epsilon_1, \quad (7.15)$$

where w_1 is the c_q -convex solution of

$$\begin{cases} \det(I + D^2 w_1) = 1 & \text{in } V := \mathcal{N}_{\theta_1^{\frac{n\hat{\alpha}}{\alpha}}}[\Omega_1], \\ w_1 = -\frac{1}{2} |x - \mu^{\frac{-1}{2}} T_1 y_0|^2 & \text{on } \partial V. \end{cases}$$

Then as $\epsilon_1 \leq \epsilon_0 \leq \tau_0 \mu^2$, Lemma 6.7 (now centered at $x_0 = 0$) shows the existence of a positive definite matrix $M = A^t A$ and a vector $p \in \mathbb{R}^n$ satisfying

$$\det M = 1, \quad (1 - C\sigma_1)I \leq M \leq (1 + C\sigma_1)I, \quad |p| \leq C\sigma_1$$

such that

$$B_{(1-\sigma_2)\sqrt{2}} \subset \mu^{-1/2} A S_\mu(u_1, 0) \subset B_{(1+\sigma_2)\sqrt{2}}, \quad \text{with } \sigma_2 := C(\sigma_1 \mu^{1/2} + \mu^{-1} \epsilon_1^{1/2}). \quad (7.16)$$

Moreover

$$|\bar{u}_1(x) - (\bar{u}_1(0) + \langle p, x \rangle + \frac{1}{2} \langle Mx, x \rangle)| \leq C(\sigma_1 \mu^{3/2} + \epsilon_1) \quad \text{in } S_\mu(u_1, 0), \quad (7.17)$$

where

$$\bar{u}_1(x) := u_1(x) + c_1(x - \mu^{\frac{-1}{2}} T_1 y_0) = \frac{1}{\mu} \left[\bar{u}(\mu^{1/2} T_1^{-1} x) - u(x_0) - c(x_0 - y_0) - \mu \right].$$

Let $M_2 := A_1^t M A_1$, A_1 is computed in the previous step. Since $M_2 - M_1 = A_1^t (M - I) A_1$, we have that $\|M_2 - M_1\| \leq \|A_1\|^2 \|M - I\| \leq C\sigma_1$. Also, because $M = A^t A$, we have that $(1 - C\sigma_1) |x|^2 \leq |Ax|^2 \leq (1 + C\sigma_1) |x|^2$. If we let $A_2 := A A_1$, then $M_2 = A_2^t A_2$ and therefore,

$$C(1 - C\sigma_1) |x|^2 \leq (1 - C\sigma_1) |A_1 x|^2 \leq |A A_1 x|^2 \leq (1 + C\sigma_1) |A_1 x|^2 \leq C(1 + C\sigma_1) |x|^2,$$

that is $\|A_2\|^2 \leq C(1 + C\sigma_1)$ and $\|A_2^{-1}\|^2 \leq 1/C(1 - C\sigma_1)$.

On the other hand, $S_\mu(u_1, 0) = \mu^{-1/2} T_1 S_{\mu^2}(u, x_0)$. It follows from (7.16) that

$$B_{(1-\sigma_2)\sqrt{2}} \subset \mu^{-1} T_2 S_{\mu^2}(u, x_0) \subset B_{(1+\sigma_2)\sqrt{2}}, \quad (7.18)$$

for $T_2 x := A_2(x - x_0)$. Moreover, from (7.17), letting $x = \mu^{-1/2} T_1 z$ we obtain

$$|\bar{u}(z) - (\bar{u}(x_0) + \langle p_2, z \rangle + \frac{1}{2} \langle M_2 z, z \rangle)| \leq C\mu \left(\sigma_1 \mu^{3/2} + \epsilon_1 \right) \quad \text{for all } z \in S_{\mu^2}(u, x_0),$$

where $p_2 := \mu^{1/2} A_1^t p$ and so $|p_2| \leq C\mu^{1/2} \sigma_1$.

We are now in good shape to iterate the induction process, that will be explained in the following paragraphs. The crucial observation is that problem (7.9) is better than the original one in the sense that the rescaled cost c_1 is closer to the quadratic one by a factor $\mu^{1/2}$.

Second induction step. We let

$$u_2(x) = \frac{1}{\mu^2} [u(\mu T_2^{-1}x) - u(x_0) - c(x_0 - y_0) - \mu^2], \quad x \in \Omega_2 := \mu^{-1}T_2 S_{\mu^2}(u, x_0).$$

Recall from (7.18) that $B_{(1-\sigma_2)\sqrt{2}} \subset \Omega_2 \subset B_{(1+\sigma_2)\sqrt{2}}$. We know that u_1 is a viscosity subsolution of $g_1(x - \nabla c_1^*(-\nabla u_1)) \det(I + D^2 c_1^*(-\nabla u_1) D^2 u_1) = f_1$ in $S_\mu(u_1, 0)$. Moreover as $m_2 := \|\partial_{c_1} u_1(S_\mu(u_1, 0))\|_{L^\infty} \leq C\sqrt{\mu + \epsilon_1} \leq C\sqrt{\mu}$ by the estimate (7.15) and Proposition 6.8, it follows from (7.10) and Corollary 5.4 that u_1 is also a viscosity supersolution in $S_\mu(u_1, 0)^{\frac{C\|A_1\|}{\sqrt{\lambda_1}}}$ relative to $S_\mu(u_1, 0)$ of the same equation. Also $u_2(x) = \frac{1}{\mu} u_1(\mu^{1/2} A^{-1}x)$, $\Omega_2 = \mu^{-1/2} A S_\mu(u_1, 0)$ and $\|A\| \leq \sqrt{C_2}$, where the matrix A is given in the previous step. Therefore by rescaling we see that u_2 is a viscosity subsolution in Ω_2 and viscosity supersolution solution in Ω_2^γ relative to Ω_2 with $\gamma := CC_2\mu^{-1/2}/\sqrt{\lambda_1}$ of the equation

$$\begin{cases} g_2(x - \nabla c_2^*(-\nabla u_2)) \det(I + D^2 c_2^*(-\nabla u_2) D^2 u_2) = f_2(x) & \text{in } \Omega_2, \\ u_2(x) = -c_2(x - \mu^{-1}T_2 y_0) & \text{on } \partial\Omega_2, \end{cases}$$

for

$$f_2(x) := f_1(\mu^{1/2} A^{-1}x) = f(\mu T_2^{-1}x), \quad g_2(y) := g_1(\mu^{1/2} A^{-1}y) = g(\mu T_2^{-1}y)$$

and

$$c_2(x) := \frac{1}{\mu} c_1(\mu^{1/2} A^{-1}x) = \frac{1}{\mu^2} c(\mu T_2^{-1}x) = \frac{1}{2} |x|^2 + \frac{1}{\lambda_2} E_2(x) \quad \text{with} \quad \frac{1}{\lambda_2} := \frac{\mu^{1/2}}{\lambda_1} = \frac{\mu}{\lambda}.$$

Likewise u_2 is also an Alexandroff weak solution of the equation in Ω_2^γ relative to Ω_2 . It also follows from (7.7) that

$$S_\rho(u_2, x_1, y_1) \subset \mu^{-1}T_2 B_{C\sqrt{\rho\mu^2}}(\mu T_2^{-1}x_1) \subset B_{C\|A_2\|\sqrt{\rho}}(x_1) \quad (7.19)$$

for all $x_1 \in \Omega_2$, $y_1 \in \partial_{c_2} u_2(x_1)$ and $\rho > 0$ as long as $S_{\rho\mu^2}(u, \mu T_2^{-1}x_1, \mu T_2^{-1}y_1) \subset S_{\mu^2}(u, x_0)$.

Moreover, Proposition 3.2 and the fact $\Omega_2 = \mu^{-1/2} A S_\mu(u_1, 0)$ give

$$\text{dist}(x, \partial[\Omega_2]) \leq \mu^{\frac{-1}{2}} \|A\| \frac{C}{\lambda_1} \leq \frac{C\sqrt{C_2}\mu^{-1}}{\lambda_2} \quad \forall x \in \partial\Omega_2.$$

We have from (7.11) that

$$|f_2(x) - 1| = |f_1(\mu^{\frac{1}{2}} A^{-1}x) - 1| \leq \theta_1 \mu^{\frac{\alpha}{2}} |A^{-1}x|^\alpha \leq \theta_2 |x|^\alpha \quad \text{for all } x \in \Omega_2,$$

where $\theta_2 := \theta_1(C_1^{-1}\mu)^{\frac{\alpha}{2}} = \tilde{\theta}(C_1^{-1}\mu)^\alpha \leq \theta_1$. Also as $\mu^{\frac{1}{2}} A^{-1} \partial_{c_2}^{\text{loc}} u_2(\Omega_2) = \partial_{c_1}^{\text{loc}} u_1(S_\mu(u_1, 0))$, (7.12) gives

$$|g_2(y) - 1| = |g_1(\mu^{\frac{1}{2}} A^{-1}y) - 1| \leq \theta_1 \mu^{\frac{\alpha}{2}} |A^{-1}y|^\alpha \leq \theta_2 |y|^\alpha \quad \text{for all } y \in \partial_{c_2}^{\text{loc}} u_2(\Omega_2).$$

Therefore, we may apply Theorem 6.2 with $d = C\sqrt{C_2}\mu^{-1}$, $\lambda \rightsquigarrow \lambda_2$, $\theta \rightsquigarrow \theta_2$ and $\beta > 0$ satisfying $(\frac{d}{\lambda_2})^\beta = \theta_2^{\frac{n\hat{\alpha}}{\alpha}}$. Notice that $\beta \leq 1$ because it follows from (7.13) and (7.14) that

$$\frac{d}{\lambda_2} = \mu^{\frac{1}{2}} \frac{d}{\lambda_1} \leq \mu^{\frac{1}{2}} \theta_1^{\frac{n\hat{\alpha}}{\alpha}} = C_1^{\frac{n\hat{\alpha}}{2}} \mu^{\frac{1-n\hat{\alpha}}{2}} \theta_2^{\frac{n\hat{\alpha}}{\alpha}} \leq \theta_2^{\frac{n\hat{\alpha}}{\alpha}}$$

and

$$\begin{aligned} \frac{1}{\lambda_2} &= \frac{\mu^{\frac{1}{2}}}{\lambda_1} \leq \mu^{\frac{1}{2}} \theta_1 \theta_1^{\frac{\hat{\alpha}}{\alpha}(n-2)(1-\alpha)} = C_1^{\frac{1}{2}(\alpha + \hat{\alpha}(n-2)(1-\alpha))} \mu^{\frac{(1-\alpha)}{2}[1-(n-2)\hat{\alpha}]} \theta_2 \theta_2^{\frac{\hat{\alpha}}{\alpha}(n-2)(1-\alpha)} \\ &\leq \theta_2 \theta_2^{\frac{\hat{\alpha}}{\alpha}(n-2)(1-\alpha)}. \end{aligned}$$

These also imply that condition (6.7) is satisfied since

$$\begin{aligned} &C_n \left(\frac{d}{\lambda_2} \right)^{-\beta((n-1)^2 + \frac{n-2}{n}\alpha)} \left[\theta_2 + \frac{1}{\lambda_2} \left(\frac{d}{\lambda_2} \right)^{\frac{n-2}{n}(\alpha-1)\beta} \right] \\ &\leq 2C_n \theta_2^{1 - \frac{n\hat{\alpha}}{\alpha}((n-1)^2 + \frac{n-2}{n}\alpha)} \leq 2C_n \theta_1^{1 - \frac{n\hat{\alpha}}{\alpha}((n-1)^2 + \frac{n-2}{n}\alpha)} \leq 1. \end{aligned}$$

Thus as $\|\partial_{c_2} u_2(\Omega_2)\|_{L^\infty} = \|\mu^{-1/2} A \partial_{c_1} u_1(S_\mu(u_1, 0))\|_{L^\infty} \leq C$ by Proposition 6.8 and $\gamma = CC_2 \mu^{-3/4} / \sqrt{\lambda_2}$, we conclude from Theorem 6.2 that

$$\|u_2 - w_2\|_{L^\infty(\Omega_2)} \leq C \max \left\{ \left(\frac{d}{\lambda_2} \right)^{\frac{2\beta}{n}}, \frac{C_2 \mu^{-\frac{3}{4}}}{\sqrt{\lambda_2}} \right\} = C \theta_2^{\frac{2\hat{\alpha}}{\alpha}} =: \epsilon_2,$$

where w_2 is the c_q -convex solution of

$$\begin{cases} \det(I + D^2 w_2) = 1 & \text{in } V := \mathcal{N}_{\theta_2^{\frac{n\hat{\alpha}}{\alpha}}}[\Omega_2], \\ w_2 = -\frac{1}{2}|x - \mu^{-1} T_2 y_0|^2 & \text{on } \partial V. \end{cases}$$

Then as $\epsilon_2 \leq \epsilon_1 \leq \epsilon_0 \leq \tau_0 \mu^2$, Lemma 6.7 shows the existence of a positive definite matrix $M = A^t A$ and a vector $p \in \mathbb{R}^n$ satisfying

$$\det M = 1, \quad (1 - C\sigma_2)I \leq M \leq (1 + C\sigma_2)I, \quad |p| \leq C\sigma_2$$

such that

$$B_{(1-\sigma_3)\sqrt{2}} \subset \mu^{-1/2} A S_\mu(u_2, 0) \subset B_{(1+\sigma_3)\sqrt{2}}, \quad \text{with } \sigma_3 := C(\sigma_2 \mu^{1/2} + \mu^{-1} \epsilon_2^{1/2}); \quad (7.20)$$

together with

$$|\bar{u}_2(x) - (\bar{u}_2(0) + \langle p, x \rangle + \frac{1}{2} \langle Mx, x \rangle)| \leq C(\sigma_2 \mu^{3/2} + \epsilon_2) \quad \text{in } S_\mu(u_2, 0), \quad (7.21)$$

where

$$\bar{u}_2(x) := u_2(x) + c_2(x - \mu^{-1} T_2 y_0) = \frac{1}{\mu^2} [\bar{u}(\mu T_2^{-1} x) - u(x_0) - c(x_0 - y_0) - \mu^2].$$

Let $M_3 := A_2^t M A_2$, where the matrix A_2 comes from the previous case. As before, $M_3 - M_2 = A_2^t (M - I) A_2$, so that $\|M_3 - M_2\| \leq \|A_2\|^2 \|M - I\| \leq C(1 + C\sigma_1)\sigma_2$, and also, $(1 - C\sigma_2)|x|^2 \leq |Ax|^2 \leq (1 + C\sigma_2)^2 |x|^2$. If we let $A_3 := A A_2$, then $M_3 = A_3^t A_3$ and therefore,

$$\begin{aligned} C(1 - C\sigma_1)(1 - C\sigma_2)|x|^2 &\leq (1 - C\sigma_2)|A_2 x|^2 \leq |A_3 x|^2 \\ &= |A A_2 x|^2 \leq (1 + C\sigma_2)|A_2 x|^2 \leq C(1 + C\sigma_1)(1 + C\sigma_2)|x|^2, \end{aligned}$$

that is

$$\|A_3\|^2 \leq C(1 + C\sigma_1)(1 + C\sigma_2) \quad \text{and} \quad \|A_3^{-1}\|^2 \leq \frac{1}{C(1 - C\sigma_1)(1 - C\sigma_2)}.$$

Since $S_\mu(u_2, 0) = \mu^{-1}T_2S_{\mu^3}(u, x_0)$, it follows from (7.20) that for $T_3x := A_3(x - x_0)$,

$$B_{(1-\sigma_3)\sqrt{2}} \subset \mu^{-\frac{3}{2}}T_3S_{\mu^3}(u, x_0) \subset B_{(1+\sigma_3)\sqrt{2}}.$$

Also by changing variables $x = \mu^{-1}T_2z$, we obtain from (7.21) that

$$|\bar{u}(z) - (\bar{u}(x_0) + \langle p_3, z \rangle + \frac{1}{2}\langle M_3z, z \rangle)| \leq C\mu^2 \left(\sigma_2\mu^{3/2} + \epsilon_2 \right) \quad \text{for all } z \in S_{\mu^3}(u, x_0),$$

where $p_3 := \mu A_2^t p$ and so $|p_3| \leq C\mu\sqrt{C(1+C\sigma_1)}\sigma_2$.

The general induction step. It is clear from above steps that how the induction process continues. The main point is the following: while the gradient of u_k deteriorates by a factor of $\mu^{k/2}$ (bad!) with respect to the gradient of the original u , on the other hand the cost c_{λ_k} is closer to quadratic by the same factor (good!). The strategy is to find the right compensation between the two facts.

Let us explain the main details of the k -th step given the $(k-1)$ -th step. From the $(k-1)$ -th step, we have

$$\|u_{k-1} - w_{k-1}\|_{L^\infty(\Omega_{k-1})} \leq C\theta_{k-1}^{\frac{2\hat{\alpha}}{\alpha}} =: \epsilon_{k-1}, \quad (7.22)$$

where $\Omega_{k-1} := \mu^{-\frac{k-1}{2}}T_{k-1}S_{\mu^{k-1}}(u, x_0)$,

$$u_{k-1}(x) := \frac{1}{\mu^{k-1}} \left[u(\mu^{\frac{k-1}{2}}T_{k-1}^{-1}x) - u(x_0) - c(x_0 - y_0) - \mu^{k-1} \right]$$

and w_{k-1} is the c_q -convex solution of

$$\begin{cases} \det(I + D^2w_{k-1}) = 1 & \text{in } V := \mathcal{N}_{\theta_{k-1}^{\frac{n\hat{\alpha}}{\alpha}}}[\Omega_{k-1}], \\ w_{k-1} = -\frac{1}{2}|x - \mu^{-\frac{k-1}{2}}T_{k-1}y_0|^2 & \text{on } \partial V. \end{cases}$$

Also there is a positive definite matrix $M_k = A_k^t A_k$ such that for $T_kx := A_k(x - x_0)$,

$$B_{(1-\sigma_k)\sqrt{2}} \subset \mu^{-\frac{k}{2}}T_kS_{\mu^k}(u, x_0) \subset B_{(1+\sigma_k)\sqrt{2}}.$$

To continue the process, we consider the rescaling function

$$u_k(x) := \frac{1}{\mu^k} \left[u(\mu^{\frac{k}{2}}T_k^{-1}x) - u(x_0) - c(x_0 - y_0) - \mu^k \right], \quad x \in \Omega_k := \mu^{-\frac{k}{2}}T_kS_{\mu^k}(u, x_0). \quad (7.23)$$

Note that the rescaled u_k satisfies a Monge-Ampère equation (in the sense discussed in the above two steps) with a new cost c_k , i.e.,

$$\begin{cases} g_k(x - \nabla c_k^*(-\nabla u_k)) \det(I + D^2c_k^*(-\nabla u_k)D^2u_k) = f_k(x) & \text{in } \Omega_k, \\ u_k(x) = -c_k(x - \mu^{-\frac{k}{2}}T_ky_0) & \text{on } \partial\Omega_k, \end{cases}$$

for

$$f_k(x) := f_{k-1}(\mu^{\frac{1}{2}}A^{-1}x) = f(\mu^{\frac{k}{2}}T_k^{-1}x), \quad g_k(y) := g_{k-1}(\mu^{\frac{1}{2}}A^{-1}y) = g(\mu^{\frac{k}{2}}T_k^{-1}y) \quad (7.24)$$

and

$$c_k(x) := \frac{1}{\mu} c_{k-1}(\mu^{\frac{1}{2}} A^{-1} x) = \frac{1}{\mu^k} c(\mu^{\frac{k}{2}} T_k^{-1} x) \quad \text{or equivalently} \quad c_k^*(z) = \frac{1}{\mu^k} c^* \left(\mu^{\frac{k}{2}} T_k^t z \right),$$

where the matrix A is given precisely in the $(k-1)$ -th step which in particular satisfy $A_k = AA_{k-1}$, $\|A\| \leq \sqrt{C_2}$ and $\|A^{-1}\| \leq C_1^{-1/2}$. Because $\|A^{-1}\|$ is universally bounded, we can write

$$c_k(x) = \frac{1}{2} |x|^2 + \frac{1}{\lambda_k} E_k(x) \quad \text{with} \quad \frac{1}{\lambda_k} := \frac{\mu^{1/2}}{\lambda_{k-1}} = \frac{\mu^{k/2}}{\lambda}$$

for some

$$|E_k(x)| \leq C |x|^3, \quad |\nabla E_k(x)| \leq C |x|^2, \quad |D^2 E_k(x)| \leq C |x|.$$

We write

$$M_{\lambda_k}[u_k] = f_k$$

for

$$M_{\lambda_k}[u_k] := g_k(x + \nabla u_k + G_k(\nabla u_k)) \det [I + D^2 u_k + (D^2 u_k) H_k(\nabla u_k)]$$

where H_k and G_k satisfy for some universal constant C ,

$$\|H_k(\nabla u_k)\|_{L^\infty} \leq \frac{C}{\lambda_k} \|\nabla u_k\|_{L^\infty(\Omega_k)}, \quad \|G_k(\nabla u_k)\|_{L^\infty} \leq \frac{C}{\lambda_k} \|\nabla u_k\|_{L^\infty(\Omega_k)}^2. \quad (7.25)$$

First observe that Proposition 3.2 and the fact $\Omega_k = \mu^{-1/2} A S_\mu(u_{k-1}, 0)$ give

$$\text{dist}(x, \partial[\Omega_k]) \leq \mu^{\frac{-1}{2}} \|A\| \frac{C}{\lambda_{k-1}} \leq \frac{C \sqrt{C_2} \mu^{-1}}{\lambda_k} \quad \forall x \in \partial\Omega_k.$$

Next since $\mu^{\frac{1}{2}} A^{-1} \Omega_k = \Omega_{k-1}$, we obtain from the $(k-1)$ -th step that

$$|f_k(x) - 1| = |f_{k-1}(\mu^{\frac{1}{2}} A^{-1} x) - 1| \leq \theta_{k-1} \mu^{\frac{\alpha}{2}} |A^{-1} x|^\alpha \leq \theta_k |x|^\alpha \quad \text{for all } x \in \Omega_k,$$

where $\theta_k := \theta_{k-1} (C_1^{-1} \mu)^{\frac{\alpha}{2}} \leq \theta_{k-1}$. Similarly as $\mu^{\frac{1}{2}} A^{-1} \partial_{c_k}^{\text{loc}} u_k(\Omega_k) = \partial_{c_{k-1}}^{\text{loc}} u_{k-1}(S_\mu(u_{k-1}, 0))$,

$$|g_k(y) - 1| = |g_{k-1}(\mu^{\frac{1}{2}} A^{-1} y) - 1| \leq \theta_{k-1} \mu^{\frac{\alpha}{2}} |A^{-1} y|^\alpha \leq \theta_k |y|^\alpha \quad \text{for all } y \in \partial_{c_k}^{\text{loc}} u_k(\Omega_k).$$

We apply Theorem 6.2 with $d = C \sqrt{C_2} \mu^{-1}$, $\lambda \rightsquigarrow \lambda_k$, $\theta \rightsquigarrow \theta_k$ and $\beta > 0$ satisfying $\left(\frac{d}{\lambda_k}\right)^\beta = \theta_k^{\frac{n\hat{\alpha}}{\alpha}}$. Notice that the constant d is the same for all steps. Also $\beta \leq 1$ because it follows from the $(k-1)$ -th step that

$$\frac{d}{\lambda_k} = \mu^{\frac{1}{2}} \frac{d}{\lambda_{k-1}} \leq \mu^{\frac{1}{2}} \theta_{k-1}^{\frac{n\hat{\alpha}}{\alpha}} = C_1^{\frac{n\hat{\alpha}}{2}} \mu^{\frac{1-n\hat{\alpha}}{2}} \theta_k^{\frac{n\hat{\alpha}}{\alpha}} \leq \theta_k^{\frac{n\hat{\alpha}}{\alpha}}$$

and

$$\begin{aligned} \frac{1}{\lambda_k} &= \frac{\mu^{\frac{1}{2}}}{\lambda_{k-1}} \leq \mu^{\frac{1}{2}} \theta_{k-1} \theta_{k-1}^{\frac{\hat{\alpha}(n-2)(1-\alpha)}{\alpha}} = C_1^{\frac{1}{2}} (\alpha + \hat{\alpha}(n-2)(1-\alpha)) \mu^{\frac{(1-\alpha)}{2} [1 - (n-2)\hat{\alpha}]} \theta_k \theta_k^{\frac{\hat{\alpha}(n-2)(1-\alpha)}{\alpha}} \\ &\leq \theta_k \theta_k^{\frac{\hat{\alpha}(n-2)(1-\alpha)}{\alpha}}. \end{aligned}$$

These also imply that condition (6.7) is satisfied because

$$\begin{aligned} & C_n \left(\frac{d}{\lambda_k} \right)^{-\beta \left((n-1)^2 + \frac{n-2}{n} \alpha \right)} \left[\theta_k + \frac{1}{\lambda_k} \left(\frac{d}{\lambda_k} \right)^{\frac{n-2}{n} (\alpha-1) \beta} \right] \\ & \leq 2C_n \theta_k^{1 - \frac{n\hat{\alpha}}{\alpha} \left((n-1)^2 + \frac{n-2}{n} \alpha \right)} \leq 2C_n \theta_{k-1}^{1 - \frac{n\hat{\alpha}}{\alpha} \left((n-1)^2 + \frac{n-2}{n} \alpha \right)} \leq 1. \end{aligned}$$

Now use Theorem 6.2 with $\gamma = CC_2^{\frac{k}{2}} \mu^{-\frac{3}{4}} / \sqrt{\lambda_k}$ and continue the same procedure as done in the first and second steps. Observe that we always have

$$\epsilon_k := C \max \left\{ \left(\frac{d}{\lambda_k} \right)^{\frac{2\beta}{n}}, \frac{C_2^{\frac{k}{2}} \mu^{-\frac{3}{4}}}{\sqrt{\lambda_k}} \right\} = C \max \left\{ \theta_k^{\frac{2\hat{\alpha}}{\alpha}}, \frac{\mu^{-\frac{3}{4}}}{\sqrt{\lambda}} (C_2^2 \mu)^{\frac{k}{4}} \right\} = C \theta_k^{\frac{2\hat{\alpha}}{\alpha}}$$

because $\theta_k^{\frac{2\hat{\alpha}}{\alpha}} = \tilde{\theta}^{\frac{2\hat{\alpha}}{\alpha}} (C_1^{-1} \mu)^{k\hat{\alpha}}$ and $\hat{\alpha} < 1/4$.

Conclusion. The above induction process allows to construct a sequence of matrices $M_k = A_k^t A_k$, transformations T_k , vectors $p_k \in \mathbb{R}^n$, and polynomials $P_k(x)$, $k \geq 1$, such that

1. $\det M_k = 1$, $\|M_{k+1} - M_k\| \leq C(1 + C\sigma_1) \cdots (1 + C\sigma_{k-1})\sigma_k$,
2. $\|A_k\|^2 \leq C(1 + C\sigma_1) \cdots (1 + C\sigma_{k-1})$, $\|A_k^{-1}\|^2 \leq C[(1 - C\sigma_1) \cdots (1 - C\sigma_{k-1})]^{-1}$,
3. $B_{(1-\sigma_k)\sqrt{2}} \subset \mu^{-k/2} T_k S_{\mu^k}(u, x_0) \subset B_{(1+\sigma_k)\sqrt{2}}$,
4. $\|\bar{u}(x) - P_k(x)\|_{L^\infty(S_{\mu^k}(u, x_0))} \leq C\mu^{k-1}(\sigma_{k-1}\mu^{3/2} + \epsilon_{k-1})$,
5. $|p_k| \leq C\mu^{\frac{k-1}{2}} \sqrt{(1 + C\sigma_1) \cdots (1 + C\sigma_{k-2})} \sigma_{k-1}$,

where

1. $P_k(x) := \bar{u}(x_0) + \langle p_k, x - x_0 \rangle + \frac{1}{2} \langle M_k(x - x_0), x - x_0 \rangle$,
2. $\epsilon_0 = \tau_0 \mu^3$ and $\epsilon_k = C\theta_k^{\frac{2\hat{\alpha}}{\alpha}} = \epsilon_0 (C_1^{-1} \mu)^{k\hat{\alpha}}$ for $k \geq 1$,
3. $\sigma_0 = 1$ and $\sigma_k = C(\sigma_{k-1}\mu^{1/2} + \mu^{-1}\epsilon_{k-1}^{1/2})$ for $k \geq 1$.

We would like to estimate σ_k . First choose μ small enough such that $\sigma_1 = C(1 + \tau_0^{\frac{1}{2}})\mu^{\frac{1}{2}} < 1$. Then as the sequence $\{\epsilon_k\}_{k=1}^\infty$ is decreasing to 0, it is clear that the sequence $\{\sigma_k\}_{k=1}^\infty$ is decreasing to 0 as well.

By induction we see that for all $k \geq 1$

$$\begin{aligned} \sigma_k &= (C\sqrt{\mu})^k + \frac{C}{\mu} (C\sqrt{\mu})^{k-1} \sum_{i=0}^{k-1} (C\sqrt{\mu})^{-i} \epsilon_i^{1/2} \\ &= (C\sqrt{\mu})^k + \frac{C\epsilon_0^{1/2}}{\mu} (C\sqrt{\mu})^{k-1} \sum_{i=0}^{k-1} [C^{-1} C_1^{-\frac{\hat{\alpha}}{2}} \mu^{\frac{\hat{\alpha}-1}{2}}]^i \\ &\leq (C\sqrt{\mu})^k + \frac{C\epsilon_0^{1/2}}{\mu} (C\sqrt{\mu})^{k-1} \frac{C^{-k} C_1^{-\frac{k\hat{\alpha}}{2}} \mu^{\frac{k(\hat{\alpha}-1)}{2}}}{C^{-1} C_1^{-\frac{\hat{\alpha}}{2}} \mu^{\frac{\hat{\alpha}-1}{2}}} \leq \hat{C} (C_1^{-1} \mu)^{\frac{k\hat{\alpha}}{2}}. \end{aligned}$$

Thus $\sum_{i=1}^{\infty} \sigma_i$ is convergent and hence $\prod_{i=1}^{\infty} (1 + C\sigma_i)$ and $\prod_{i=1}^{\infty} (1 - C\sigma_i)$ converge to some positive numbers since $C\sigma_i \leq 1/2$ for all i large. Consequently, we obtain the following estimates for all $k \geq 1$:

$$\|A_k\| \leq C, \quad \|A_k^{-1}\| \leq C, \quad |p_k| \leq C \mu^{\frac{k-1}{2}} (C_1^{-1} \mu)^{\frac{(k-1)\hat{\alpha}}{2}} \quad \text{and} \quad \|M_{k+1} - M_k\| \leq C (C_1^{-1} \mu)^{\frac{k\hat{\alpha}}{2}}.$$

It follows that M_k is a Cauchy sequence and hence $M_k \rightarrow M$ for some matrix M and $p_k \rightarrow 0$ as $k \rightarrow \infty$. So if we let $P_0(x) := \bar{u}(x_0) + \langle M(x - x_0), (x - x_0) \rangle$, then $P_k(x) \rightarrow P_0(x)$ pointwise in \mathbb{R}^n . Recall also from Theorem 4.7 that $B_{C\mu^{k/2}}(x_0) \subset S_{\mu^k}(u, x_0)$ for all $k \geq 1$. Therefore, for any $l = 1, 2, \dots$ and for any $x \in B_{C\mu^{l/2}}(x_0) \subset S_{\mu^l}(u, x_0)$, we obtain

$$\begin{aligned} |\bar{u}(x) - P_0(x)| &\leq |\bar{u}(x) - P_l(x)| + |P_l(x) - P_0(x)| \\ &\leq C\mu^{l-1}(\sigma_{l-1}\mu^{3/2} + \epsilon_{l-1}) + \|p_l\| |x - x_0| + \sum_{k=l}^{\infty} |\langle (M_k - M_{k+1})(x - x_0), (x - x_0) \rangle| \\ &\leq C'\mu^{l-1}(C_1^{-1}\mu)^{\frac{(l-1)\hat{\alpha}}{2}} + C'\mu^l \sum_{k=l}^{\infty} (C_1^{-1}\mu)^{\frac{k\hat{\alpha}}{2}} \leq C''\mu^l (C_1^{-1}\mu)^{\frac{l\hat{\alpha}}{2}} \leq C''\mu^{\frac{l}{2}(2+\alpha')}, \end{aligned}$$

where the last inequality holds if we choose $0 < \alpha' < \hat{\alpha}$ (note that $C_1 < 1$). This yields (7.2) as desired and we conclude that u is $\mathcal{C}^{2,\alpha'}$ at x_0 for any $0 < \alpha' < \frac{\alpha}{n((n-1)^2 + \frac{n-2}{n}\alpha)}$. The proof is completed.

References

- [1] L. Ambrosio, N. Gigli, and G. Savaré. *Gradient flows in metric spaces and in the space of probability measures*. Lectures in Mathematics ETH Zürich. Birkhäuser Verlag, Basel, 2005.
- [2] Y. Brenier. Polar factorization and monotone rearrangement of vector-valued functions. *Comm. Pure Appl. Math.*, 44(4):375–417, 1991.
- [3] L. Caffarelli and Y. Y. Li. A Liouville theorem for solutions of the Monge-Ampère equation with periodic data. *Ann. Inst. H. Poincaré Anal. Non Linéaire*, 21(1):97–120, 2004.
- [4] L. A. Caffarelli. Interior a priori estimates for solutions of fully nonlinear equations. *Ann. of Math. (2)*, 130(1):189–213, 1989.
- [5] L. A. Caffarelli. Interior $W^{2,p}$ estimates for solutions of the Monge-Ampère equation. *Ann. of Math. (2)*, 131(1):135–150, 1990.
- [6] L. A. Caffarelli. A localization property of viscosity solutions to the Monge-Ampère equation and their strict convexity. *Ann. of Math. (2)*, 131(1):129–134, 1990.
- [7] L. A. Caffarelli. The regularity of mappings with a convex potential. *J. Amer. Math. Soc.*, 5(1):99–104, 1992.
- [8] L. A. Caffarelli. Boundary regularity of maps with convex potentials. *Comm. Pure Appl. Math.*, 45(9):1141–1151, 1992.

- [9] L. A. Caffarelli. Allocation maps with general cost functions. In *Partial differential equations and applications*, volume 177 of *Lecture Notes in Pure and Appl. Math.*, pages 29–35. Dekker, New York, 1996.
- [10] L. A. Caffarelli. Boundary regularity of maps with convex potentials. II. *Ann. of Math. (2)*, 144(3):453–496, 1996.
- [11] L. A. Caffarelli. A priori estimates and the geometry of the Monge Ampère equation. In *Nonlinear partial differential equations in differential geometry (Park City, UT, 1992)*, volume 2 of *IAS/Park City Math. Ser.*, pages 5–63. Amer. Math. Soc., Providence, RI, 1996.
- [12] L. A. Caffarelli and X. Cabré. *Fully nonlinear elliptic equations*, volume 43 of *American Mathematical Society Colloquium Publications*. American Mathematical Society, Providence, RI, 1995.
- [13] L. A. Caffarelli, C. E. Gutiérrez, and Q. Huang. On the regularity of reflector antennas. *Ann. of Math. (2)*, 167(1):299–323, 2008.
- [14] M. de Guzmán. Differentiation of integrals in \mathbf{R}^n . In *Measure theory (Proc. Conf., Oberwolfach, 1975)*, pages 181–185. Lecture Notes in Math., Vol. 541. Springer, Berlin, 1976.
- [15] P. Delanoë. Gradient rearrangement for diffeomorphisms of a compact manifold. *Differential Geom. Appl.*, 20(2):145–165, 2004.
- [16] P. Delanoë and Y. Ge. Regularity of optimal transport on compact, locally nearly spherical, manifolds. *J. Reine Angew. Math.*, 646:65–115, 2010.
- [17] A. Figalli, Y.-H. Kim, and R. J. McCann. Continuity and injectivity of optimal maps for non-negatively cross-curved costs. Preprint 2009.
- [18] A. Figalli and G. Loeper. C^1 regularity of solutions of the Monge-Ampère equation for optimal transport in dimension two. *Calc. Var. Partial Differential Equations*, 35(4):537–550, 2009.
- [19] W. Gangbo and R. J. McCann. Optimal maps in Monge’s mass transport problem. *C. R. Acad. Sci. Paris Sér. I Math.*, 321(12):1653–1658, 1995.
- [20] W. Gangbo and R. J. McCann. The geometry of optimal transportation. *Acta Math.*, 177(2):113–161, 1996.
- [21] D. Gilbarg and N.S. Trudinger. *Elliptic Partial Differential Equations of Second Order*. Springer-Verlag, New York, 2001.
- [22] C. E. Gutiérrez. *The Monge-Ampère equation*. Progress in Nonlinear Differential Equations and their Applications, 44. Birkhäuser Boston Inc., Boston, MA, 2001.
- [23] C. E. Gutiérrez, Q. Huang and T. V. Nguyen. Interior Hölder estimates for solutions of the Monge-Ampère equation. Preprint.
- [24] C. E. Gutiérrez and T. V. Nguyen. On Monge-Ampère type equations arising in optimal transportation problems. *Calc. Var. Partial Differential Equations*, 28(3):275–316, 2007.

- [25] G. Loeper. On the regularity of solutions of optimal transportation problems. *Acta Math.*, 202(2):241–283, 2009.
- [26] Y.-H. Kim and R. J. McCann. Continuity, curvature, and the general covariance of optimal transportation. *Oberwolfach Rep.* 4: 2060–2062, 2007.
- [27] X.-N. Ma, N. S. Trudinger, and X.-J. Wang. Regularity of potential functions of the optimal transportation problem. *Arch. Ration. Mech. Anal.*, 177(2):151–183, 2005.
- [28] R. Schneider. *Convex bodies: the Brunn-Minkowski theory*. Encyclopedia of Mathematics and its Applications, 44. Cambridge University Press, Cambridge, 1993.
- [29] N. S. Trudinger and X.-J. Wang. On strict convexity and continuous differentiability of potential functions in optimal transportation. *Arch. Ration. Mech. Anal.*, 192(3):403–418, 2009.
- [30] N. S. Trudinger and X.-J. Wang. On convexity notions in optimal transportation. Preprint, 2008.
- [31] N. S. Trudinger and X.-J. Wang. On the second boundary value problem for Monge-Ampère type equations and optimal transportation. *Ann. Sc. Norm. Super. Pisa Cl. Sci. (5)*, 8(1):143–174, 2009.
- [32] C. Villani. *Topics in optimal transportation*, volume 58 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2003.
- [33] C. Villani. *Optimal transport, Old and new*, volume 338 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, 2009.
- [34] M. Warren. Regularity for a log-concave to log-concave mass transfer problem with near Euclidean cost. *Comm. Anal. Geom.* 19, no. 1, 191-208, 2011.