1-1-2015

On asymptotic behaviour and W 2, p regularity of potentials in optimal transportation

Jiakun Liu
University of Wollongong, jiakunl@uow.edu.au

Neil Trudinger
Australian National University, neilt@maths.anue.edu.au

Xu-Jia Wang
Australian National University

On asymptotic behaviour and W 2, p regularity of potentials in optimal transportation

Abstract

© 2014, Springer-Verlag Berlin Heidelberg. In this paper we study local properties of cost and potential functions in optimal transportation. We prove that in a proper normalization process, the cost function is uniformly smooth and converges locally smoothly to a quadratic cost x · y, while the potential function converges to a quadratic function. As applications we obtain the interior W2, p estimates and sharp C1, α estimates for the potentials, which satisfy a Monge–Ampère type equation. The W2, p estimate was previously proved by Caffarelli for the quadratic transport cost and the associated standard Monge–Ampère equation.

Keywords

potentials, optimal, w, transportation, regularity, 2, asymptotic, behaviour, p

Disciplines

Engineering | Science and Technology Studies

Publication Details

Liu, J., Trudinger, N. & Wang, X. (2015). On asymptotic behaviour and W 2, p regularity of potentials in optimal transportation. Archive for Rational Mechanics and Analysis, 215 (3), 867-905. Archive for Rational Mechanics and Analysis

ON ASYMPTOTIC BEHAVIOUR AND $W^{2,p}$ REGULARITY OF POTENTIALS IN OPTIMAL TRANSPORTATION

JIAKUN LIU, NEIL S. TRUDINGER, AND XU-JIA WANG

Abstract. In this paper we study local properties of cost and potential functions in optimal transportation. We prove that in a proper normalization process, the cost function is uniformly smooth and converges locally smoothly to a quadratic cost $x \cdot y$, while the potential function converges to a quadratic function. As applications we obtain the interior $W^{2,p}$ estimates and sharp $C^{1,\alpha}$ estimates for the potentials, which satisfy a Monge-Ampère type equation. The $W^{2,p}$ estimate was previously proved by Caffarelli for the quadratic transport cost and the associated standard Monge-Ampère equation.

1. Introduction

In this paper we study the local geometry of cost and potential functions in optimal transportation, and prove the $W^{2,p}$ estimates for potential functions, that is for generalized solutions u to the Monge-Ampère type equations of the form

$$\det [D^2u(x) - A(x, Du)] = f(x) \quad \text{in} \quad \Omega, \tag{1.1}$$

where Ω is a bounded domain in the Euclidean space $\mathbb{R}^n$, A is a matrix function given by

$$A(x, Du) = D_x^2 c(x, T_u), \tag{1.2}$$

$c(\cdot, \cdot)$ is the transport cost function on $\mathbb{R}^n \times \mathbb{R}^n$, T_u is the optimal mapping, which is uniquely determined by

$$Du(x) = D_x c(x, T_u(x)), \tag{1.3}$$

and the function f is determined by the initial and target densities.

When $c(x, y) = x \cdot y$ (or equivalently $c(x, y) = |x - y|^2$, with u replaced by $u - |x|^2$), then $A \equiv 0$ and (1.1) is the standard Monge-Ampère equation. In this case, the regularity has been studied by many researchers [28], and in particular the fundamental $W^{2,p}$ estimate was proved by Caffarelli [4], following his discovery in [2] of the

2000 *Mathematics Subject Classification.* 35J60, 35B45; 49Q20, 28C99.
Key words and phrases. Optimal transportation, Monge-Ampère equation, $W^{2,p}$ estimates.
This work is supported by the Australian Research Council.

corresponding estimate for uniformly elliptic equations. For general cost functions other than the quadratic one, Caffarelli [6] proposed to investigate the local geometry of potentials, which is essential for the study of the regularity of the potentials. As pointed out by Villani [32], the regularity of optimal transportation with general costs was a main open problem in the area ([32], §4.3 Open problems).

Recently it has been found that the regularity of optimal mappings relies on a sharp analytical condition of the cost function [24, 30, 23]. In this paper we introduce a proper normalization and prove that in the process of the normalization, cost functions satisfying the condition in [24] converge to the quadratic cost $c(x, y) = x \cdot y$ and the corresponding potential functions converge to the quadratic function $u(x) = |x|^2$. As applications, we prove interior $W^{2,p}$ estimates and sharp $C^{1,\alpha}$ estimates for potentials for general cost functions.

The relevant optimal transportation problem can be briefly stated as follows. Let Ω and Ω^* be two bounded domains in $\mathbb{R}^n$, and ρ, ρ^* be two probability densities supported in Ω and Ω^*, respectively. Let $c \in C^\infty(\mathbb{R}^n \times \mathbb{R}^n)$ be a transport cost function. The optimal transportation problem is to find a measure preserving mapping which maximizes (or minimizes) the cost functional,

$$\mathcal{C}(T) = \int_\Omega \rho(x) c(x, T(x)) dx, \tag{1.4}$$

among the set of measure preserving mappings T from Ω to Ω^*. A mapping T is called measure preserving if $\mu_\rho[T^{-1}(E)] = \mu_{\rho^*}[E]$ $\forall$ Borel set $E \subset \Omega^*$, where $\mu_\rho = \rho dx$.

The existence and uniqueness of optimal mappings was proved by various authors including Brenier, Caffarelli, and Gangbo-McCann [1, 5, 13], under appropriate conditions on the cost function. Their proof applies to cost functions satisfying [19, 24]:

(A1) For any $x, p \in \mathbb{R}^n$, there is a unique $y \in \mathbb{R}^n$ such that $D_x c(x, y) = p$; and for any $y, q \in \mathbb{R}^n$, there is a unique $x \in \mathbb{R}^n$ such that $D_y c(x, y) = q$.

(A2) For any $(x, y) \in \mathbb{R}^n \times \mathbb{R}^n$, $\det\{c_{i,j}(x, y)\} \neq 0$, where $c_{i,j} = \frac{\partial^2 c(x,y)}{\partial x_i \partial y_j}$.

Under these two conditions, the optimal mapping $T = T_u$ is uniquely determined by Kantorovich's potential function u through (1.3). If $u \in C^2(\Omega)$, one differentiates (1.3) to get the Monge-Ampère type equation (1.1), with the matrix A given by (1.2), and the right hand side $f = |\det D^2_{xy} c| \frac{\rho}{\rho^* \circ T_u}$ [24].

The remaining key theoretical issue is thus the regularity of optimal mappings. When the cost function $c(x, y) = x \cdot y$, equation (1.1) becomes the standard Monge-Ampère equation

$$\det [D^2 u] = f, \tag{1.5}$$

which has been extensively studied [28], in particular the interior $C^{2,\alpha}$ and $W^{2,p}$ estimates have been obtained by Caffarelli [4]. The regularity of optimal mappings with non-quadratic cost functions has been a focus of attention in the area in recent years [6, 32], as non-quadratic costs arise frequently in applications. By formula (1.3) and conditions (A1)–(A2), it suffices to study the regularity of the potential function u. Inspired by the work [35], the first breakthrough was made in [24], where the C^3 regularity of u was proved if $f > 0, \in C^2$, Ω^* is c^*-convex with respect to Ω, and the cost function c satisfies the following structure condition:

(A3) $\exists$ a constant $c_0 > 0$ such that $\forall\ x \in \Omega, y \in \Omega^*$, and $\xi, \eta \in \mathbb{R}^n$ with $\xi \perp \eta$,

$$\sum_{i,j,k,l,p,q,r,s} (c_{ij,rs} - c^{p,q} c_{ij,p} c_{q,rs}) c^{r,k} c^{s,l} \xi_i \xi_j \eta_k \eta_l \geq c_0 |\xi|^2 |\eta|^2, \tag{1.6}$$

where the subscripts of c before the comma mean derivatives in x, after the comma mean derivatives in y, and $c^{i,j}$ is the inverse of the matrix $c_{i,j}$.

Corresponding global regularity was subsequently proved in [30], under the weaker condition $c_0 \geq 0$ (called A3w). The c^*-convexity of Ω^* is necessary [24]. Surprisingly the conditions (A3) and (A3w) are also sharp. Loeper [23] showed that if there exist $x \in \Omega, y \in \Omega^*$, and vectors $\xi, \eta \in \mathbb{R}^n$ with $\xi \perp \eta$ such that the left hand side of (1.6) is negative, then there exists $f > 0, \in C^\infty$ such that the potential function u is not C^1 smooth.

To obtain more precise regularity, such as the $C^{2,\alpha}$ and $W^{2,p}$ estimates, for the optimal transportation, under weaker data assumptions, one needs a better understanding of the local geometry of potential functions, and the local geometry of cost functions satisfying (A3). The structure condition (A3) can be expressed in the equivalent form [24],

$$D^2_{p_k p_l} A_{ij}(x,p) \xi_i \xi_j \eta_k \eta_l \geq c_0 |\xi|^2 |\eta|^2, \qquad \forall\ \xi, \eta \in \mathbb{R}^n \text{ with } \xi \perp \eta, \tag{1.7}$$

where $A = (A_{ij})$ is the matrix in (1.2). From (1.7), Loeper [23] observed a convexity type property of cost functions satisfying (A3) and proved the $C^{1,\alpha}$ regularity of potentials when $f \in L^p(\Omega)$ with $p > n$. Based on Loeper's observation, we proved in [29] the strict c-convexity of potential functions, and proved in [20] the convexity of level sets, the $C^{1,\alpha}$ regularity when $f \in L^p(\Omega)$ with $p > \frac{n+1}{2}$, and the sharp $C^{1,\frac{1}{2n-1}}$ regularity when $f \in L^\infty(\Omega)$. In particular we mention an alternative approach to [29], also from [23], was found independently by Kim and McCann [18]. By a perturbation argument [37] and the local analysis in [20], the interior C^2 and $C^{2,\alpha}$ estimates were proved in [22], assuming respectively the Dini and Hölder continuity of f. Recently, the $W^{2,1+\varepsilon}_{loc}$ estimate for strictly c-convex potentials with the cost c satisfying (A3w) and the function f satisfying $C^{-1} < f < C$ has been established [8] extending the corresponding estimate for strictly convex solutions of (1.5) [9, 10, 26]. See also [18, 12, 34] for related works on local geometry of transport costs and potentials, and [27, 31, 32, 33] for more on optimal transportation.

To prove the $W^{2,p}$ estimate for the potential functions, in this paper we make a more detailed study of local geometry of potential functions and cost functions. Let us first introduce the following normalization, for details see §2. For any point $x_0 \in \Omega$, replace u by $u - u(x_0) - [c(x, y_0) - c(x_0, y_0)]$, denote $S^0_{h,u}(x_0) = \{x \in \Omega : u(x) < h\}$ the sub-level set of u, and make the coordinate transform $x \to D_y c(x, y_0)$ such that $S^0_{h,u}(x_0)$ is convex [20], where $y_0 = T_u(x_0)$. Choose proper coordinate axes such that $E = \left\{\sum_i x_i^2/r_i^2 < 1\right\}$ is the minimum ellipsoid of $S^0_{h,u}(x_0)$, see [17] and §2. Then we can normalize $S^0_{h,u}$ by the linear transform $\bar{x} = G_h x$, where $G_h = \mathrm{diag}(r_1^{-1}, \cdots, r_n^{-1})$ is a matrix, such that E becomes the unit ball. We also make the linear transform $\bar{y} = G_h^* y$ for the variable y, where $G_h^* = h^{-1}\mathrm{diag}(r_1, \cdots, r_n)$. Let

$$\begin{aligned} \bar{u}_h(\bar{x}) &= u(x)/h, \\ \bar{c}_h(\bar{x}, \bar{y}) &= \frac{1}{h} c(x, y). \end{aligned}$$

Then as $h \to 0$, we have the following asymptotic behaviour for the cost and potential functions.

Theorem 1.1. *Assume the cost function c satisfies (A1)–(A3), $f > 0$ and is continuous, and Ω^* is c^*-convex with respect to Ω. Let u be a generalized solution to* (1.1)*. Then under the above normalization, we have, as $h \to 0$,*

$$\bar{c}_h \to \bar{c}_0 \tag{1.8}$$

locally in $C^k(\mathbb{R}^n)$ for any $k \geq 0$, where $\bar{c}_0(x, y) = x \cdot y$, and

$$\bar{u}_h \to \bar{u}_0 \tag{1.9}$$

locally uniformly in $\mathbb{R}^n$, where $\bar{u}_0(x) = |x|^2$.

We remark that if f is not continuous but $C_1 \leq f \leq C_2$ for positive constants C_1, C_2, then (1.8) is still true but (1.9) is not in general. We have instead a weaker result, namely $\bar{u}_h$ converges to a strictly convex function defined in $\mathbb{R}^n$ of polynomial growth. We also note that the cost function in normalization satisfies (A1), (A2) in a locally uniform way, and satisfies a weak form of (A3). From Theorem 1.1 we also obtain a covering property of the sub-level sets. Combining the above results and Caffarelli's techniques [2], we are able to prove the following $W^{2,p}$ estimate.

Theorem 1.2. *Assume the cost function c satisfies (A1)–(A3), f is continuous, $C_1 \leq f \leq C_2$, and Ω^* is c^*-convex with respect to Ω. Let u be a generalized solution* *to* (1.1)*. Then* $D^2u \in L^p(\Omega')$ $\forall$ $p \geq 1$, $\Omega' \Subset \Omega$*, and we have the estimate*

$$\|u\|_{W^{2,p}(\Omega')} \leq C, \tag{1.10}$$

where C *depends on* n, p, C_1, C_2, $\Omega, \Omega', \Omega^*$*, and the modulus of continuity of* f.

Here we say u is a generalized solution to (1.1) if it is a potential function to the optimal transportation problem (1.4) [24]. From our argument, the c^*-convexity of Ω^* in Theorems 1.1 and 1.2 can be replaced by the assumption that u is strictly c-convex and is C^1 smooth. When Ω^* is c^*-convex, the strict c-convexity and C^1 regularity of u was proved in [29]. Also by an example in [36], the continuity of f is necessary for the $W^{2,p}$-estimate for all $p \geq 1$. Therefore conditions in Theorem 1.2 are sharp.

We also prove the following $C^{1,\alpha}$ regularity result.

Theorem 1.3. *Suppose the cost function* c *satisfies (A1)–(A3) and* Ω^* *is* c^**-convex with respect to* Ω*. Then there exists a constant* $\varepsilon > 0$ *such that if*

$$|f - 1| \leq \varepsilon, \tag{1.11}$$

then a generalized solution u *to* (1.1) *must be locally* $C^{1+\alpha}$ *for some* $\alpha \in (1 - C_1\varepsilon, 1]$*, and* $\forall$ $\Omega' \Subset \Omega$*, we have the estimate*

$$\|u\|_{C^{1+\alpha}(\Omega')} \leq C, \tag{1.12}$$

where the constants C *and* C_1 *depend on* $\varepsilon, n, \Omega, \Omega^*$*, and dist*$(\Omega', \partial\Omega)$*; and* C *also depends on* α*. Both* C *and* C_1 *are uniformly bounded for* $\varepsilon > 0$ *small.*

Note that in Theorem 1.3 we don't assume the continuity of f and we have the linear relation $\alpha \geq 1 - C_1\varepsilon$. Theorem 1.3 is a special case of Corollary 5.1 below. We remark that our proof of Theorem 1.2 also implies a related result. That is, $\forall$ $p < \infty$, $\exists$ $\varepsilon = \varepsilon(p)$ such that if f satisfies (1.11), then $u \in W^{2,p}_{loc}(\Omega)$. See Theorem 7.1 below.

The paper is organized as follows. In Section 2 we give some preliminary estimates and introduce the normalization used in Theorem 1.1. In Section 3 we prove (1.8), showing that the cost function converges to the function $x \cdot y$ in the normalization process. In Section 4, we prove (1.9), that is the potential function converges to a quadratic function in the normalization. We also derive a covering property of sub-level sets needed for the $W^{2,p}$ estimate. In Section 5 we estimate the eccentricity of minimum ellipsoids of sub-level sets of potential functions and prove Theorem 1.3. In Section 6 we give a density estimate for the set in which the second derivatives satisfy an upper bound. The $W^{2,p}$ estimate (Theorem 1.2) then follows from a polynomial decay estimate in Section 7.

Finally, in view of the asympotic limit of the cost function in Theorem 1.1 being degenerate with respect to the A3w condition, it is reasonable to consider the weakening of A3 to A3w in our hypotheses. For this we need to strengthen the condition on the target Ω^* to a "uniform" c-convexity as used in [11] and we plan to treat this in a future work.

2. Preliminaries

Through this section we assume that the cost function c satisfies (A1)-(A3), the function f is measurable and bounded, and the domain Ω^* is c-convex with respect to Ω. For some estimates, the condition $f \geq C > 0$ will also be specified.

2.1. Some terminologies. Let u be a continuous function in Ω. We say a function of the form $\varphi = c(\cdot, y_0) + a_0$ is a *c-support* of u at $x_0 \in \Omega$ if $u(x_0) = \varphi(x_0)$ and $u(x) \geq \varphi(x)$ for all $x \in \Omega$. Here the prefix c in c-support stands for the cost function. When $c(x, y) = x \cdot y$, a c-support coincides with the usual support function in the theory of convex bodies. We say u is *c-convex* if for any point $x_0 \in \Omega$, there exists a c-support at x_0 in Ω.

Let u be a c-convex function. The *c-normal mapping* of u is a set-valued mapping $T_u \colon \Omega \to \mathbb{R}^n$, which is given by the following: for any $x_0 \in \Omega$, $T_u(x_0)$ is the set of points y_0 such that $c(x, y_0) + a$ is a c-support of u at x_0, where $a = u(x_0) - c(x_0, y_0)$. For any subset $E \subset \Omega$, denote $T_u(E) = \bigcup_{x\in E} T_u(x)$.

By the c-normal mapping we introduce a measure $\mu = \mu_{u,\rho^*}$ in Ω such that for any Borel set $E \subset \Omega$,

$$\mu(E) = \int_{T_u(E)} \rho^*(y)dy, \tag{2.1}$$

where ρ^* is the density of mass distribution in Ω^*. It was proved that μ is a Radon measure if u is a potential function in optimal transportation [24].

Definition 2.1. *A c-convex function u is a (generalized) solution to* (1.1)*, in the sense of Aleksandrov, if $\mu_{u,\rho^*} = \rho dx$, that is for any Borel set $E \subset \Omega$,*

$$\int_E \rho(x)dx = \int_{T_u(E)} \rho^*(y)dy, \tag{2.2}$$

where $f = |detD^2_{xy}c|\frac{\rho}{\rho^*\circ T_u}$ is regarded as a single function. (In this paper we can always assume that $\rho^* = 1$.)

Similarly we can define sub-solution and super-solution, in the sense of Aleksandrov. That is u is a sub-solution to (1.1) if the equality " $=$ " in (2.2) is replaced by " $\leq$ ";

and u is a super-solution to (1.1) if " $=$ " is replaced by " $\geq$ ". Obviously a smooth sub-solution satisfies $\mathcal{M}(u) \geq f$ and a smooth super-solution satisfies $\mathcal{M}(u) \leq f$, where $\mathcal{M}$ denotes the operator on the left hand side of equation (1.1).

We say a curve $\ell \subset \mathbb{R}^n$ is a *c-segment* with respect to a point $y_0 \in \mathbb{R}^n$ if $D_y c(\ell, y_0)$ is a line segment in $\mathbb{R}^n$; and a set U is *c-convex* with respect to another set V if the image $D_y c(U, y)$ is convex for each $y \in V$.

Similarly we can define c^*-segment, c^*-support and c^*-convexity by exchanging the variables x and y. The above notions were introduced in [24, 29], but note that in this paper we consider c-convex functions rather than c-concave ones in [24, 29].

2.2. Sub-level sets. From now on, we always assume that u is a strictly c-convex and C^1 smooth solution to (1.1).

For any given point $x_0 \in \Omega$ and any positive constant $h > 0$, denote

$$S^0_{h,u}(x_0) = \{x \in \Omega : u(x) < \varphi(x) + h\} \tag{2.3}$$

the *sub-level set* of u, where $\varphi = c(\cdot, y_0) + a_0$ is the c-support of u at x_0, $a_0 = u(x_0) - c(x_0, y_0)$, $y_0 = T_u(x_0)$. Note that the c-support is unique under the above assumptions. For simplicity, we write sometimes $S^0_{h,u}(x_0)$ as S^0_h when no confusion arises.

For a fixed $x_0 \in \Omega$, we make the changes

$$\begin{array}{rcl} c(x,y) & \longrightarrow & [c(x,y) - c(x,y_0)] - [c(x_0,y) - c(x_0,y_0)], \\ u(x) & \longrightarrow & [u(x) - u(x_0)] - [c(x,y_0) - c(x_0,y_0)], \\ v(y) & \longrightarrow & [v(y) - v(y_0)] - [c(x_0,y) - c(x_0,y_0)], \end{array} \tag{2.4}$$

where $y_0 = T_u(x_0)$ and v is the dual potential function ([24]). Then the cost function c satisfies

$$\begin{array}{ll} c(x, y_0) \equiv 0 & \forall\, x \in \Omega, \\ c(x_0, y) \equiv 0 & \forall\, y \in \Omega^*, \end{array} \tag{2.5}$$

and the potential functions u, v satisfy

$$\begin{array}{llll} u(x_0) = 0, & Du(x_0) = 0, & u \geq 0 & \text{in } \Omega, \\ v(y_0) = 0, & Dv(y_0) = 0, & v \geq 0 & \text{in } \Omega^*. \end{array} \tag{2.6}$$

Here we list some properties which will be used below.

- The sub-level set $S^0_{h,u}(x_0)$ is c-convex w.r.t. y_0 if $S^0_{h,u}(x_0) \Subset \Omega$. Moreover, by making the coordinate transform

$$\begin{array}{l} x \to D_y c(x, y_0), \\ y \to D_x c(x_0, y), \end{array} \tag{2.7}$$

the sub-level set $S^0_{h,u}(x_0)$ can be made *convex* [20]. We note that the second transform for y is not needed for the convexity of S^0_h but is needed in the normalization §2.4 below. Note that the c-convexity of sub-level sets was also invoked and used critically in [11].

• We also have

$$D_{p_k}A_{ij}(\cdot, p_0) \equiv 0, \quad \text{for all } i, j, k. \tag{2.8}$$

This formula was also proved in [20], see (2.2) and (2.11) there.

• It is well known that [17] for any convex set $D \subset \mathbb{R}^n$ of positive volume, there is a unique ellipsoid E, called the *minimum ellipsoid* of D, which attains the minimum volume among all ellipsoids containing the set D and satisfies

$$\frac{1}{n}E \subset D \subset E, \tag{2.9}$$

where $\frac{1}{n}E = \{z + \frac{1}{n}(x - z) :\ x \in E\}$ and z is the centre of E. We say D is *normalized* if E is a unit ball.

• By a rotation of the coordinates, we may assume that

$$E = \left\{\sum\nolimits_i (x_i - z_i)^2/r_i^2 < 1\right\}, \quad r_1 \geq \cdots \geq r_n \tag{2.10}$$

is the minimum ellipsoid of S^0_h. It was proved in [20, 22] that

$$C^{-1} \leq (r_1 \cdots r_n)^2/h^n \leq C, \tag{2.11}$$

where the constant C depends only on n, the constant c_0 in (1.7), and the upper and lower bounds of f, but is independent of h. See also [11] for loosing the dependency of c_0 in (2.11).

• We have furthermore the estimate [20, Lemma 4]

$$hr_1^2/r_n^2 \leq C \tag{2.12}$$

for some constant $C > 0$ depending only on n, the cost function c and $\sup f, \inf f$. We point out that C actually depends on the constant c_0 in condition (A3).

The following lemma shows that the volume of the sub-level set S^0_h is comparable to that of the Euclidean ball of radius $\sqrt{h}$.

Lemma 2.1. *For any small $h > 0$ such that $S^0_h \Subset \Omega$, we have*

$$C_1 h^{n/2} \leq |S^0_h| \leq C_2 h^{n/2}, \tag{2.13}$$

where $C_1, C_2 > 0$ depend only on $n, \inf_\Omega f, \sup_\Omega f$, and the cost function c.

Proof. Let E be the minimum ellipsoid of the sub-level set S^0_h, given in (2.10). By the relation (2.9), estimate (2.13) follows from (2.11). □

2.3. Some preliminary estimates. For any domain $E \subset \mathbb{R}^n$, denote $\mathcal{N}_\delta(E) = \{x \in \mathbb{R}^n : \text{dist}(x, E) < \delta\}$ the δ-neighborhood of E, and $\mathcal{N}_{-\delta}(E) = \{x \in E : \text{dist}(x, \partial E) > \delta\}$ the δ-subset of E. Denote $A_{ij,kl} = D^2_{p_k p_l} A_{ij}$. We have (see (1.6),(1.7)),

$$A_{ij,kl} = (c_{ij,rs} - c^{m,n} c_{ij,m} c_{n,rs}) c^{r,k} c^{s,l}, \tag{2.14}$$

which is uniformly bounded. We also denote

$$Mu = \{D_{ij} u - A_{ij}(\cdot, Du)\}. \tag{2.15}$$

In the following two lemmas, r_0 is a fixed, small constant. But r_0 can be any fixed constant if the matrix $A(x, Du)$ is small. In Section 3 below, we will show that for the normalized solution $\bar{u} = \bar{u}_h$ (see (2.19) below), $A(x, Du)$ can be as small as we want, provided h is small.

Lemma 2.2. *Let $\delta > \varepsilon > 0$ be small constants, and $v \in C^0, w \in C^2$ be two c-convex functions in $B_{r_0}(0) \Subset \Omega$ with $\|v - w\|_{L^\infty(B_{r_0})} \le \varepsilon$. Then*

$$T_v(B_{r_0 - C\varepsilon/\delta}) \subset \mathcal{N}_\delta\{T_w(B_{r_0})\}, \tag{2.16}$$

where $C = r_0^{-1}$, T_v, T_w are the c-normal mappings of v, w respectively. If $w \in C^2(\overline{B}_{r_0})$ and $\det Mw > 0$, then

$$\mathcal{N}_{-\delta}\{T_w(B_{r_0 - C\varepsilon/\delta})\} \subset T_v(B_{r_0}), \tag{2.17}$$

where C depends also on the C^2 norm of w.

Proof. Let $w^* = w + \varepsilon - \delta(r_0^2 - |x|^2)$. Then

$$\begin{aligned} D_i w^* &= D_i w + 2\delta x_i, \\ D_{ij} w^* &= D_{ij} w + 2\delta \delta_{ij}. \end{aligned}$$

By (2.8), we have

$$|A_{ij,p_k}(\cdot, p)| \le |A_{ij,kl}|\,|p| \le C$$

when p is the gradient of w. Hence for any i, j and $x \in B_{r_0}(0)$, we have

$$|A_{ij}(\cdot, Dw) - A_{ij}(\cdot, Dw^*)| = \left|\sum\nolimits_k A_{ij,p_k}(\cdot, p) D_k(w - w^*)\right| \le C r_0 \delta.$$

Therefore, for $r_0 > 0$ small such that $C r_0 \le 1/4$,

$$Mw^* \ge Mw + \delta I,$$

where I is the $n \times n$ unit matrix. It implies that w^* is c-convex as the matrix $\{D_{ij} w^* - A_{ij}(\cdot, Dw^*)\}$ is positive definite [29]. By the assumption $\|v - w\|_{L^\infty(B_{r_0})} \le \varepsilon$, we have obviously

$$\begin{aligned} w^* &\ge v \quad \text{on } \partial B_{r_0}, \\ w^* &\le v \quad \text{when } |x| \le r_0 - \tfrac{\varepsilon}{r_0 \delta}. \end{aligned}$$

By the monotonicity lemma (Lemma 5.2 in [24]), we obtain

$$T_v(B_{r_0-\varepsilon/r_0\delta}) \subset T_{w^*}(B_{r_0}).$$

By the definition of w^*,

$$\{\nabla w^*|_{B_{r_0}}\} \subset \mathcal{N}_{2\delta r_0}\{\nabla w|_{B_{r_0}}\}.$$

By (1.3) and the assumption (A2) we obtain

$$T_v(B_{r_0-\varepsilon/r_0\delta}) \subset \mathcal{N}_{C'\delta r_0}\{T_w(B_{r_0})\}.$$

When r_0 is small, $2C'r_0 < 1$. The first inclusion is proved.

By considering the auxiliary function $w_* = w - \varepsilon + \delta(r_0^2 - |x|^2)$, we can similarly obtain the second inclusion $\mathcal{N}_{-\delta}\{T_w(B_{r_0-C\varepsilon/\delta})\} \subset T_v(B_{r_0})$. Note that $w \in C^2(\overline{B}_{r_0})$, w_* is c-convex when $\delta > 0$ is small, depending on the C^2-norm of w. □

Note that the assumption $w \in C^2$ for (2.16) can be relaxed. The C^2 assumption is used only in the calculation of the matrix Mw^*.

Next we give a comparison principle, which shows that close data imply close solutions.

Lemma 2.3. *Let v and w be respectively solutions to det$[Mv] = f$ and det$[Mw] = 1$ in B_{r_0}, where M is the operator in* (2.15). *Suppose $|v - w| < \varepsilon$ on ∂B_{r_0} and $1 < f < 1 + \varepsilon$. Then*

$$\|v - w\|_{L^\infty(B_{r_0})} \le 2\varepsilon. \tag{2.18}$$

Proof. By the comparison principle [24], we have $v \le w + \varepsilon$. To prove $v \ge w - 2\varepsilon$, let $w^* = w + \varepsilon(|x|^2 - r_0^2) - \varepsilon$. From the proof of the previous lemma, and note that w is smooth, we have $Mw^* \ge Mw + \varepsilon I$. Hence

$$\det Mw^* \ge \det Mw + \varepsilon \sum M^{ii} \ge 1 + \varepsilon,$$

where $\{M^{ij}\}$ is the co-factor matrix of Mw. By the comparison principle again, we obtain $v \ge w^*$. □

2.4. Renormalization. The argument in this paper is built upon a normalization process which we introduce now. Let $S_h^0 = S_{h,u}^0(x_0)$ be a sub-level set of u, which is convex by the change (2.7). Let E be the minimum ellipsoid of S_h^0 as in (2.10). First we make the changes (2.4) such that (2.5), (2.6) hold. By a translation of coordinates we also assume

$$x_0 = 0, \quad y_0 = 0.$$

The normalization of u in S_h^0 is the rescaled solution

$$\bar{u}(\bar{x}) = \frac{1}{h} u(x), \tag{2.19}$$

together with the change of the cost function

$$\bar{c}(\bar{x}, \bar{y}) = \frac{1}{h} c(x, y), \tag{2.20}$$

where

$$\bar{x} = G_h x, \quad G_h = \mathrm{diag}(r_1^{-1}, \cdots, r_n^{-1}), \tag{2.21}$$

$$\bar{y} = G_h^* y, \quad G_h^* = h^{-1}\mathrm{diag}(r_1, \cdots, r_n). \tag{2.22}$$

Note that both $\bar{u} = \bar{u}_h$ and $\bar{c} = \bar{c}_h$ depend on h. We point out that the transforms G_h, G_h^* are different. By (2.11),

$$C^{-1}h^{-n/2} \le |G_h^*| \approx |G_h| \le Ch^{-n/2},$$

where $|G_h|$ denotes the determinant of the corresponding matrix, which is positive.

Denote $U = \{x \in \mathbb{R}^n : \ (r_1 x_1, \cdots, r_n x_n) \in S_h^0\}$. Then $\inf_U \bar{u} = 0$ and $\bar{u} = 1$ on ∂U. After the normalization, equation (1.1) becomes

$$\det\left[D^2 \bar{u} - \bar{A}(\cdot, D\bar{u})\right] = \bar{f}, \tag{2.23}$$

where $\bar{f} = \frac{(r_1 \cdots r_n)^2}{h^n} f$, $\bar{A}_{ij} = \frac{r_i r_j}{h} A_{ij}$. By (2.8),

$$\bar{A}_{ij,k}(\cdot, 0) = 0. \tag{2.24}$$

By direct computation,

$$\bar{A}_{ij,kl} = h \frac{r_i r_j}{r_k r_l} A_{ij,kl}, \quad i, j, k, l = 1, \cdots, n, \tag{2.25}$$

which is uniformly bounded [20, Lemma 4].

3. Convergence of the normalized cost function

In this section we show that the cost function $\bar{c}$ is locally uniformly smooth and converges to the cost function $\bar{x} \cdot \bar{y}$ as $h \to 0$.

By making the linear transform $\hat{y}_k = c_{k,l}(0,0) y_l$ as in [29], such that $c_{x_i, \hat{y}_j} = c_{x_i, y_k} c^{k,j} = \delta_{ij}$, we may assume directly that $c_{i,j}(0,0) = \delta_{ij}$. Now we make the substitution (2.4) such that (2.5) and (2.6) hold. By a translation of coordinates we also assume $x_0 = y_0 = 0$ in (2.4). By the Taylor expansion we have

$$\begin{aligned} c(x,y) \ &= \ x \cdot y + c_{ij,k} x_i x_j y_k + c_{i,jk} x_i y_j y_k + c_{ij,kl} x_i x_j y_k y_l + \\ &\quad + c_{ijk,l} x_i x_j x_k y_l + c_{i,jkl} x_i y_j y_k y_l + \cdots \end{aligned} \tag{3.1}$$

for (x, y) near $(0, 0)$.

For simplicity we call a term $x_{i_1}\cdots x_{i_a}y_{j_1}\cdots y_{j_b}$ (which is homogeneous of degree a in x and degree b in y) an *(a,b)-term*. By the substitution (2.4), the coefficients of $(0,k),(k,0)$-terms are 0 for all $k\ge 0$, as shown in (2.5).

As in (2.7) we make the coordinate transform

$$\begin{aligned}\tilde x_i &= D_{y_i}c(x,0) = x_i + c_{jk,i}x_jx_k + c_{jkl,i}x_jx_kx_l + \cdots ,\\ \tilde y_i &= D_{x_i}c(0,y) = y_i + c_{i,jk}y_jy_k + c_{i,jkl}y_jy_ky_l + \cdots ,\end{aligned} \tag{3.2}$$

where $1\le i\le n$. Then

$$\begin{aligned}\tilde x\cdot\tilde y \;=\;& x\cdot y + c_{ij,k}x_ix_jy_k + c_{i,jk}x_iy_jy_k + c_{ij,m}c_{m,kl}x_ix_jy_ky_l +\\ &+c_{ijk,l}x_ix_jx_ky_l + c_{i,jkl}x_iy_jy_ky_l + \cdots .\end{aligned} \tag{3.3}$$

Hence in the $(\tilde x,\tilde y)$-coordinates, the cost function c becomes

$$\tilde c(\tilde x,\tilde y) = \tilde x\cdot\tilde y + (c_{ij,kl} - c_{ij,m}c_{m,kl})\tilde x_i\tilde x_j\tilde y_k\tilde y_l + \sum R_{a,b}(\tilde x,\tilde y) \tag{3.4}$$

for $(\tilde x,\tilde y)$ near $(0,0)$. The summation above is for $a,b\ge 2$ and $5\le a+b\le N$ (it was pointed out above that $R_{0,b}=R_{a,0}=0$, and we will show below that $R_{1,b}=R_{a,1}=0$), where $N\ge 5$ is an integer. When $a+b=N$, we also use the mean-value formula in the Taylor expansion. From (3.1) and (3.2),

$$\begin{aligned}\tilde c_{i,j} \;&=\; \delta_{ij},\\ \tilde c_{ij,k} \;&=\; \tilde c_{i,jk} = 0,\\ \tilde c_{ij,kl} \;&=\; c_{ij,kl} - c_{ij,m}c_{m,kl} = A_{ij,kl}.\end{aligned} \tag{3.5}$$

The second formula in (3.5) follows from (2.8), $D_{p_k}A_{ij} = c^{m,k}c_{ij,m}$. By inserting (3.2) into (3.3) and using (3.1), it is also easy to check that in (3.4), all coefficients of the $(1,k),(k,1)$-terms are 0 for $k\ge 2$. Therefore, if $R_{a,b}(\tilde x,\tilde y)$ is a nonzero remainder term, then either $a\ge 2$ or $b\ge 2$, and $a+b\ge 5$.

Lemma 3.1. *Let $\bar u,\bar c$ and $\bar x,\bar y$ be the normalization as defined in §2.4 (which depend on h). Then as $h\to 0$, the cost function $\bar c$ sub-converges locally in C^k for any $k\ge 0$ to the polynomial*

$$\bar c^0(\bar x,\bar y) = \bar x\cdot\bar y + a_{ij,kl}\bar x_i\bar x_j\bar y_k\bar y_l, \tag{3.6}$$

where $a_{ij,kl}$ are bounded constants.

Proof. The normalization (2.19)–(2.22) can be decomposed into two steps.

(i), $\hat u(\hat x) = \frac1h u(\tilde x)$, $\hat c(\hat x,\hat y) = \frac1h c(\tilde x,\tilde y)$, and the dilation of the coordinates L: $(\hat x,\hat y) = h^{-1/2}(\tilde x,\tilde y)$. In this step, the cost function (3.4) changes to

$$\hat c(\hat x,\hat y) = \hat x\cdot\hat y + \hat c_{ij,kl}\hat x_i\hat x_j\hat y_k\hat y_l + \sum\nolimits_{2\le a,b\le N} h^{\frac{a+b}{2}-1}R_{a,b}(\hat x,\hat y) \tag{3.7}$$

where $\hat c_{ij,kl} = h\tilde c_{ij,kl} = hA_{ij,kl}$, and the factor $h^{\frac{a+b}{2}-1}$ arises in the dilation.

(ii), $\bar u(\bar x)=\hat u(Q^{-1}\hat x)$, $\bar c(\bar x,\bar y)=\hat c(Q^{-1}\hat x,Q\hat y)$, and the coordinates transforms $\bar x=Q\hat x$, $\bar y=Q^{-1}\hat y$, where

$$Q=Q_h=h^{1/2}\mathrm{diag}(r_1^{-1},\cdots,r_n^{-1}). \tag{3.8}$$

In this step, the cost function (3.7) becomes

$$\bar c(\bar x,\bar y)=\bar x\cdot\bar y+h\frac{r_ir_j}{r_kr_l}A_{ij,kl}\bar x_i\bar x_j\bar y_k\bar y_l+\sum\nolimits_{2\le a,b\le N}\bar R_{a,b}(\bar x,\bar y). \tag{3.9}$$

By (2.11), $C^{-1}\le|Q_h|\le C$, here $|Q_h|$ denotes the determinant of the corresponding matrix.

As pointed in (2.25), the coefficients $h\frac{r_ir_j}{r_kr_l}A_{ij,kl}$ are uniformly bounded. The remainder terms $\bar R_{a,b}(\bar x,\bar y)$ satisfy

$$\begin{aligned}|\bar R_{a,b}(\bar x,\bar y)|&\le h^{\frac{a+b}{2}-1}\frac{r_1^a}{h^{a/2}}\frac{h^{b/2}}{r_n^b}|R_{a,b}(\bar x,\bar y)|\\&=h^{b-1}\frac{r_1^a}{r_n^b}|R_{a,b}(\bar x,\bar y)|\end{aligned} \tag{3.10}$$

for all (a,b) satisfying $a\ge2,b\ge2,a+b\ge5$. Since c is smooth, the coefficients of $R_{a,b}$ are bounded. Therefore to show that $\bar R_{a,b}$ is bounded, it suffices to estimate the factor $h^{b-1}\frac{r_1^a}{r_n^b}$. We claim that for all $a\ge2,b\ge2,a+b\ge5$,

$$h^{b-1}\frac{r_1^a}{r_n^b}\to0\ \text{ as }\ h\to0. \tag{3.11}$$

Indeed, by (2.12) we have

$$\begin{aligned}h^{b-1}\frac{r_1^a}{r_n^b}&=h^{b/2-1}r_1^{a-b}\cdot\left(h^{b/2}\frac{r_1^b}{r_n^b}\right)\\&\le Ch^{b/2-1}r_1^{a-b}.\end{aligned} \tag{3.12}$$

By the strict c-convexity of potential functions [29], we have $r_1\to0$ as $h\to0$. When $a\ge b$, we have either $\frac{b}{2}-1>0$ or $a-b>0$. Hence

$$h^{b/2-1}r_1^{a-b}\to0\ \text{ as }\ h\to0.$$

When $b\ge a\ge3$, by (2.11) we have $r_1\ge Ch^{1/2}$. Hence

$$h^{b/2-1}r_1^{a-b}\le h^{a/2-1}\to0.$$

When $b>a=2$, we go back to (3.11). By the $C^{1,\alpha}$ regularity of potential functions [20, 23], $h\le Cr_n^{1+\alpha}$. Hence by (2.12),

$$h^{b-1}\frac{r_1^a}{r_n^b}\le Ch^{b-2}/r_n^{b-2}\to0. \tag{3.13}$$

The claim is proved.

Therefore the coefficients in (3.9) are uniformly bounded, and $|\bar{R}_{a,b}| \to 0$ locally uniformly. Hence the cost function $\bar{c}$ converges locally uniformly to the function in (3.6).

To prove the local convergence in C^k, we write the Taylor expansion as follows

$$\bar{c}(\bar{x},\bar{y}) = \sum_{|\alpha|+|\beta|<N} \partial_{\bar{x}}^{\alpha}\partial_{\bar{y}}^{\beta} c(0,0)\bar{x}^{\alpha}\bar{y}^{\beta} + \sum_{|\alpha|+|\beta|=N} \partial_{\bar{x}}^{\alpha}\partial_{\bar{y}}^{\beta} c(\bar{x}',\bar{y}')\bar{x}^{\alpha}\bar{y}^{\beta} =: I + II,$$

where $(\bar{x}',\bar{y}') = t(\bar{x},\bar{y})$ for some $t \in [0,1]$, and $N > 1$ is chosen sufficiently large. From the above proof, the first part I converges locally uniformly. As I is a polynomial, hence it converges smoothly to a polynomial.

Therefore we need only to show that the second part II converges locally in C^k for any given k. It suffices to show that $II = o(|\bar{x}|^k + |\bar{y}|^k)$. By the expression (3.9), $II = \sum_{a+b=N} \bar{R}_{a,b}(\bar{x},\bar{y})$. Hence we need only to show that

$$\bar{R}_{a,b}(\bar{x},\bar{y}) = o(|\bar{x}|^k + |\bar{y}|^k). \tag{3.14}$$

Noticing that $h \le \min(|\bar{x}|,|\bar{y}|)$, similarly to (3.10), it suffices to show that

$$h^{b-k-1}\frac{r_1^a}{r_n^b} \to 0 \quad \text{as} \quad h \to 0. \tag{3.15}$$

If $a \ge b$, by (2.12) we have

$$h^{b-k-1}\frac{r_1^a}{r_n^b} \le h^{b/2-k-1} r_1^{a-b}.$$

Note that the strict c-convexity of the potential functions implies that $r_1 \le h^{\delta}$ for some $\delta > 0$. Hence

$$h^{b-k-1}\frac{r_1^a}{r_n^b} \le h^{b/2-k-1+\delta(a-b)} \to 0$$

provided $N = a+b$ is sufficiently large. If $a \le b$, by (2.12) we have

$$h^{b-k-1}\frac{r_1^a}{r_n^b} \le \frac{h^{b-a/2-k-1}}{r_n^{b-a}}.$$

By the $C^{1,\alpha}$ regularity, we have $h \le C r_n^{1+\alpha}$. Hence

$$\frac{h^{b-a/2-k-1}}{r_n^{b-a}} \le r_n^{\alpha(b-a)+a/2-k-1} \to 0$$

if $N = a+b$ is sufficiently large. We have thus proved (3.14). This completes the proof. □

Lemma 3.1 is critical for the rest part of this paper. The above proof relies on the condition (A3), from which we have (2.12), the strict c-convexity, $C^{1,\alpha}$ regularity, and the boundedness of the coefficients $h\frac{r_i r_j}{r_k r_l}A_{ij,kl}$. In the proof we have also used the assumption that the cost function $c \in C^{\infty}$, so that we can choose N sufficiently large.

But for the proof of Theorems 1.2 and 1.3, it suffices to have the C^4 convergence in Theorem 1.1 (or in Lemma 3.1). For the C^4 convergence, from the above proof, we need to assume the cost function $c \in C^m$ for a very large m, which depends on the α in the $C^{1,\alpha}$ regularity.

Note that
$$\frac{\partial^2 \bar c(\bar x, \bar y)}{\partial \bar x_i \partial \bar y_j} = r_i r_j^{-1} \frac{\partial^2 c(x,y)}{\partial x_i \partial y_j}.$$
Hence we have
$$\det\left(\frac{\partial^2 \bar c(\bar x, \bar y)}{\partial \bar x_i \partial \bar y_j}\right) = \det\left(\frac{\partial^2 c(x,y)}{\partial x_i \partial y_j}\right). \tag{3.16}$$
For any given constant $R > 0$ and any point $(\bar x, \bar y) \in B_R(0) \times B_R(0)$, by the strict c-convexity and C^1 regularity of potential functions [29],
$$G_h^{-1}(\bar x), (G_h^*)^{-1}(\bar y) \to 0 \quad \text{as} \quad h \to 0, \tag{3.17}$$
where G_h and G_h^* are given in (2.21) and (2.22). Hence when $h \to 0$, we have
$$\det\left(\frac{\partial^2 \bar c(\bar x, \bar y)}{\partial \bar x_i \partial \bar y_j}\right) \to 1 \tag{3.18}$$
locally uniformly. In particular the limit function $\bar c^0$ in (3.6) satisfies
$$\det\left(\frac{\partial^2 \bar c(\bar x, \bar y)}{\partial \bar x_i \partial \bar y_j}\right) \equiv 1. \tag{3.19}$$
Since $\bar c^0$ is a 4th order polynomial, it follows that
$$\bar c^0(\bar x, \bar y) \equiv \bar x \cdot \bar y. \tag{3.20}$$
It also implies that the coefficients $\bar A_{ij,kl} = h \frac{r_i r_j}{r_k r_l} A_{ij,kl}$ given in (2.25) converge to zero uniformly. Therefore Lemma 3.1 can be strengthened to

Theorem 3.1. *Let $\bar u, \bar c$ and $\bar x, \bar y$ be the normalization in §2.4. Then as $h \to 0$, the cost function $\bar c$ sub-converges in C^4 to the polynomial*
$$\bar c^0(\bar x, \bar y) = \bar x \cdot \bar y. \tag{3.21}$$

By Theorem 3.1, conditions (A1) and (A2) are satisfied by $\bar c\,(= \bar c_h)$, in a uniform way in h. But condition (A3), or (1.7), is degenerated to a weak version, namely
$$D^2_{p_k p_l} \bar A_{ij}(\bar x, p)\xi_i \xi_j \eta_k \eta_l \ge 0 \quad \forall\, \xi, \eta \in \mathbb{R}^n \text{ with } \ \xi \perp \eta. \tag{3.22}$$

Remark 3.1. (i) The proof of Theorem 3.1 uses the strict c-convexity and the $C^{1,\alpha}$ continuity of potential functions, but not the continuity of f. So it holds for solutions to (1.1) provided $C_1 \le f \le C_2$ for any positive constants C_1, C_2.
(ii) From now on, when we study a sub-level set $S^0_{h,u}$ of u, we always mean a small h

such that $S^0_{h,u} \Subset \Omega$, and the rescaled solution $\bar{u}$ satisfies (2.23) with small $\bar{A}$.

4. A Bernstein theorem

In the previous section we proved the smooth convergence of the normalized cost function to the cost $\bar{c}(x,y) = x \cdot y$, namely Theorem 3.1. In this section we consider the behavior of the potential function $\bar{u}$ as $h \to 0$.

4.1. A local gradient estimate. In our rescaling argument we need an interior gradient estimate for the rescaled solution $\bar{u} = \bar{u}_h$. First we prove that after normalization (2.19)–(2.22), the minimum point of $\bar{u}$ stays away from the boundary ∂U, uniformly as $h \to 0$. In the following we will write $\bar{x}$ as x for simplicity.

Lemma 4.1. *There exist two small constants $\lambda, \delta_0 > 0$, such that*

$$\text{dist}\left(S^0_{\lambda,\bar{u}}, \partial S^0_{1,\bar{u}}\right) \geq \delta_0 \tag{4.1}$$

for all $h > 0$ small.

Proof. If not, we assume that dist $\left(S^0_{\lambda,\bar{u}}, \partial S^0_{1,\bar{u}}\right) =: \delta \to 0$, as $h \to 0$. Before the normalization G_h, let $x_0 \in \partial S^0_{\lambda h,u}$ and $\hat{x} \in \partial S^0_{h,u}$ such that

$$|G_h(x_0) - G_h(\hat{x})| = \delta \to 0, \quad \text{as } h \to 0. \tag{4.2}$$

By choosing a proper coordinate system we assume that $x_0 = 0$ and $\hat{x}$ is on the positive x_n-axis. Denote $d = |x_0 - \hat{x}|$ and

$$S' = S^0_{h,u} \cap \{x_n > -2d\}.$$

Because the normalization does not change the ratio $\frac{|S'|}{|S^0_{h,u}|}$, then from (4.2),

$$\frac{|S'|}{|S^0_{h,u}|} = O(\delta) \to 0 \quad \text{as } \delta \to 0. \tag{4.3}$$

The boundary of S' consists of two parts, one lies on the hyperplane $\{x_n = -2d\}$ and the other one is the remaining part of $\partial S'$.

Let E' be the minimum ellipsoid of the convex set S'. Making a rotation of the coordinates $x' = Rx$, we assume that $E' = \left\{\left(\frac{x_i - x_i^*}{r'_i}\right)^2 < 1\right\}$ with $r'_1 \geq \cdots \geq r'_n$, where x^* is the centre of the ellipsoid. Then (4.3) implies that

$$\frac{r'_1 \cdots r'_n}{h^{n/2}} \to 0 \quad \text{as } \ \delta \to 0. \tag{4.4}$$

By (4.2) we also have $r'_n \leq Cd$. Next we make a coordinate transform $x'' = T'x'$, where

$$T' = \text{diag}(r'^{-1}_1, \cdots, r'^{-1}_n),$$

such that E' becomes the unit ball. Then $\bar{u}$ satisfies (in the x''-variables)

$$\det\,[D_{ij}\bar{u} - hA_{ij,kl}\frac{r'_i r'_j}{r'_k r'_l}D_k\bar{u}D_l\bar{u}] = \tilde{f} \quad \text{in } \; U', \tag{4.5}$$

where $\tilde{f} = \frac{(r'_1\cdots r'_n)^2}{h^n}f$ and $U' = T'(S')$. The RHS $\to 0$ as $\delta \to 0$. From the argument in [20], we have the estimate

$$\big\{hA_{ij,kl}\frac{r'_i r'_j}{r'_k r'_l}\big\} \le C, \tag{4.6}$$

uniformly in h and δ. See Remark 3.1 in [20].

Note that the boundary $\partial U'$ has two parts, $\partial U' = \Gamma_1 \cup \Gamma_2$, where

$$\begin{aligned} \Gamma_1 :=& \;\; T'R\left(S^0_{h,u} \cap \{x_n = -2d\}\right), \\ \Gamma_2 :=& \;\; T'R\left(\partial S^0_{h,u} \cap \{x_n > -2d\}\right), \end{aligned}$$

and R is the rotation introduced above. We make another rotation of coordinates $x''' = R'x''$ such that Γ_1 is parallel to $\{x_n = 0\}$ and $R'(U')$ lies on the upper side of Γ_1. For simplicity we write x''' as x. After the rotation, we have the equation

$$\det\,[D_{ij}\bar{u} - a_{ij,kl}D_k\bar{u}D_l\bar{u}] = g \tag{4.7}$$

with $|a_{ij,kl}| < C$ and $g \to 0$ as $\delta \to 0$.

After the above coordinate changes we have $U' \Subset B_2(0)$. In particular, after the rotation R' we have

$$\Gamma_1 = \{x_n = -c_1\}$$

for some constant $c_1 > 0$. We claim that $\frac{1}{n} \le c_1 \le 4$. Indeed, since S' is contained in the slab $\{-2d < x_n < d\}$, after the above linear coordinate changes, $U' \subset \{-c_1 < x_n < \frac{1}{2}c_1\}$. But U' is normalized. So $\frac{1}{n} \le c_1 \le 4$.

Now we construct the auxiliary function

$$w(x) = \frac{\varepsilon}{2}|x|^2 + \frac{4\varepsilon}{c_1}x_n + \lambda + \varepsilon,$$

where $\varepsilon > 0$ is a relatively small constant. By direct calculation,

$$\begin{aligned} &|w_{x_i}| \le C\varepsilon \;\text{ in } U', \\ &w_{x_i x_j} = \varepsilon\delta_{ij}. \end{aligned}$$

Hence

$$\big\{D_{ij}w - \sum\nolimits_{k,l} a_{ij,kl}D_k w D_l w\big\} \ge (\varepsilon - C\varepsilon^2)I$$

as a matrix. Hence if $\varepsilon > 0$ is sufficiently small, we have $\det Mw \ge (\varepsilon/2)^n$, where Mw is the matrix on the left hand side of the above inequality.

Next we estimate w on the boundary $\partial U'$. Recalling that $U' \Subset B_2(0)$, hence on the part Γ_1 we have

$$\begin{aligned} w(x) &= \frac{\varepsilon}{2}|x|^2 - 4\varepsilon + \lambda + \varepsilon \\ &\le 2\varepsilon - 4\varepsilon + \lambda + \varepsilon \\ &= \lambda - \varepsilon < 0. \qquad (\text{if } \lambda < \varepsilon) \end{aligned}$$

That is,

$$w < 0 \le \bar{u} \quad \text{on} \quad \Gamma_1.$$

Noting that $|x| \le 2, |x_n| \le c_1$, we have

$$w(x) \le 7\varepsilon + \lambda \le 8\varepsilon < 1 \quad \text{on } \Gamma_2.$$

Hence

$$w < 1 = \bar{u} \quad \text{on } \Gamma_2.$$

Now we let $\lambda = \frac{1}{2}\varepsilon$ and fix a small $\varepsilon > 0$ such that the above inequalities hold.

If δ is sufficiently small such that $g < (\varepsilon/2)^n$, we find that w is a subsolution. So by the comparison principle we have $w \le \bar{u}$. Hence

$$\bar{u}(0) \ge w(0) = \lambda + \varepsilon = \frac{3}{2}\varepsilon.$$

But on the other hand, we have $\bar{u}(0) = \lambda$. The contradiction gives a positive lower bound for δ. Hence the lemma is proved. □

Remark 4.1. From Lemma 4.1 we obtain

$$u(\theta x) \ge \lambda u(x) \quad \text{for} \quad x \in \partial S^0_{h,u} \tag{4.8}$$

for any $h > 0$ small, where $\theta = 1 - \frac{1}{2}\delta_0$. It follows that

$$u(x) \ge \lambda^k u(\theta^{-k} x)$$

for any x near the origin. Hence

$$u(x) \ge C|x|^{1+\beta}, \tag{4.9}$$

where β is determined by $\theta^{1+\beta} = \lambda$. This estimate was previously proved in [22] by a duality argument.

Remark 4.2. Applying Lemma 4.1 to $\bar{u}$ in the set $\{x \in \mathbb{R}^n : \ \bar{u}(x) < \lambda^{-1}\}$, for $\lambda \in (0,1)$, we have $\text{dist}(S^0_{1,\bar{u}}, \partial S^0_{\lambda^{-1},\bar{u}}) \ge \delta_0 > 0$. Hence by the interior gradient estimate (3.11) in [20], Lemma 4.1 can be strengthened to the following:

- $\forall\ \lambda \in (0,1)$, $\exists\ \delta_0 > 0$ such that (4.1) holds for all $h > 0$ small.

From this statement we in turn obtain the following interior gradient estimate.

Lemma 4.2. *$\forall\ \lambda \in (0,1)$, $\exists\ C_\lambda > 0$ such that*

$$|D\bar{u}(x)| \le C_\lambda, \quad \forall\ x \in S^0_{\lambda,\bar{u}} \tag{4.10}$$

for all $h > 0$ small.

Lemma 4.2 strengthens the gradient estimate in [20], where an estimate was obtained in the set $\{x \in S^0_{1,\bar{u}} : \mathrm{dist}(x, \partial S^0_{1,\bar{u}}) > \delta\}$ for $\delta > 0$. The estimate (4.10) was pointed out before Lemma 3.1 in [22], with its proof based on estimates in [20]. It is nontrivial and so we include the details of the proof here. Note that the proof of Lemma 4.1 can be simplified if one uses Theorem 3.1, which implies that the coefficients $\bar{A}_{ij,kl} = h\frac{r_i r_j}{r_k r_l} A_{ij,kl}$ given in (2.25) converge to zero uniformly. However the gradient estimate (4.10) was also used in [22] so we present our original proof which is independent of Theorem 3.1.

By the gradient estimate, $\bar{u} = \bar{u}_h$ converges along a subsequence $h_k \to 0$ to a locally Lipschitz continuous function u^0,

$$u^0 = \lim_{k\to\infty} u^k \tag{4.11}$$

where $u^k = \bar{u}_{h_k}$. By our normalization in §2, we have $u^0(0) = 0$ and $u^0 \ge 0$.

Corollary 4.1. *The function u^0 is convex and is defined on the entire space $\mathbb{R}^n$.*

Proof. Recall that $\bar{u}$ is $\bar{c}$-convex, namely at any point $x_0 \in G_h^{-1}(\Omega)$, there is a $\bar{c}$-support of $\bar{u}$ at x_0. Hence by Theorem 3.1, u^0 is convex.

Let D be the domain of definition of u^0. If $D \ne \mathbb{R}^n$, then $u^0(x) \to \infty$ as $x \to \partial D$. For any constant $m \ge 1$, denote $S^0_m = \{x \in D :\ u^0(x) < m\}$ the sub-level set. Let T_m be the linear coordinate transform which normalizes S^0_m. Let

$$\begin{aligned} u^0_m(x) &= \tfrac{1}{m} u^0(T_m^{-1}(x)),\\ u^k_m(x) &= \tfrac{1}{m} u^k(T_m^{-1}(x)). \end{aligned} \tag{4.12}$$

Then for any given m, $u^k_m \to u^0_m$ locally uniformly.

If $D \ne \mathbb{R}^n$, the interior gradient estimate (4.10) does not hold for u^0_m for sufficiently large m. Fix such an m. By the convergence $u^k_m \to u^0_m$, hence (4.10) does not hold for u^k_m for sufficiently large k. But note that

$$u^k_m = \bar{u}_{h_k m} \tag{4.13}$$

and for the fixed m, $h_k m \to 0$ as $k \to \infty$. By Lemma 4.2 we reach a contradiction. □

4.2. Bernstein Theorem. By the above interior gradient estimate and using the interior second derivative estimate of Pogorelov type [22, 21], the equation becomes uniformly elliptic and we have the following interior estimate [22, 21].

Lemma 4.3. *Let $w \in C^4(U)$ be a solution to the equation (2.23) with right hand side $\bar f \equiv 1$. Suppose U is normalized, $w = 1$ on ∂U. Then for any $t < 1$, we have the estimate*

$$\|w\|_{C^4(U_t)} \le C, \tag{4.14}$$

where $U_t = \{x \in U : \ w(x) < t\}$, C is a constant depending on n, t, the cost function c, and the upper bound in (4.10).

Remark 4.3.
(i) In Lemma 4.3, the estimate (4.14) still holds if U has a *good shape*, namely if there exists a constant C^* (independent of U) such that $B_r \subset U \subset B_R$ for constants $R > r > 0$ with $R/r \le C^*$. But in this case the constant C in (4.14) also depends on C^*.
(ii) For the rescaled function $\bar u$, Lemma 4.3 also holds at $t = 1$, or more generally at any fixed $t > 0$, provided h is sufficiently small.

Theorem 4.1. *When f is positive and continuous, the function u^0 in (4.11) is a quadratic function,*

$$u^0(x) = |x|^2. \tag{4.15}$$

Proof. By (2.23), u^k is a solution to

$$\det\,[D^2u^k - A^k(\cdot, Du^k)] = f^k \quad \text{in } G_k(\Omega),$$

where $G_k = G_{h_k}$, $A^k = \bar A_{h_k}$ and $f^k = \bar f_{h_k}$. Note that $\bar A$ and $\bar f$ in (2.23) depend on h.

For any given constant $m \ge 1$, let $w^k = w^k_m$ be the solution to

$$\begin{cases} \det\,[D^2w - A^k(\cdot, Dw)] = \inf_{S^0_m} f^k & \text{in } S^0_m, \\ w = u^k & \text{on } \partial S^0_m, \end{cases} \tag{4.16}$$

where $S^0_m = \{x \in \mathbb{R}^n : \ u^0(x) < m\}$ is the sub-level set. Since u^0 is convex (Corollary 4.1) and S^0_1 is normalized, there exists $R > 0$ such that $S^0_m \subset B_R(0)$. The existence of solutions to the Dirichlet problem (4.16) can be obtained by the Perron method [14]. By the a priori estimate (Lemma 4.3) and approximation [24], the solution to (4.16) is smooth and we have the estimate $\|w^k\|_{C^4(S^0_{m/2})} \le C$. Hence by choosing a subsequence, $w^k \to w^0$ locally smoothly in $S^0_{m/2}$ and w^0 is a smooth function.

By (3.17) and since f is continuous, f^k converges locally uniformly to $f(0)$. Hence as $k \to \infty$,

$$f^k - \inf_{S^0_m} f^k \to 0 \ \text{ uniformly in } S^0_m.$$

Let $\tilde{w}^k = w^k + \delta(|x|^2 - R^2)$, from Theorem 3.1 $M\tilde{w}^k \geq Mw^k + \delta I$ when k is sufficiently large, where M denotes the matrix operator in (2.15). By the comparison principle [29], we then have

$$w^k \geq u^k \geq w^k - \delta R^2.$$

Therefore, $\forall \varepsilon > 0$ set $\delta = \varepsilon/R^2$, there exists $k_0 > 0$ such that when $k \geq k_0$,

$$|w^k - u^k| \leq \varepsilon. \tag{4.17}$$

Namely, $w^k - u^k \to 0$ uniformly in S^0_m. Hence we have $w^0 = u^0$. That is, $u^0 \in C^4$.

Since f is positive and continuous, we have

$$\lim_{k\to\infty} \inf_{S^0_m} f^k = f(0).$$

Observe that w^k is a solution to (4.16) and w^k converges smoothly to $w^0 = u^0$. We see that u^0 is a smooth convex solution to

$$\det D^2 u = f(0) \quad \text{in } \mathbb{R}^n. \tag{4.18}$$

By the Bernstein theorem for (4.18) [25], we conclude that u^0 is a quadratic function. Since the sub-level set S^0_{1,u^0} is normalized and $u(0) = \min u = 0$, we have $f(0) = 2^n$ and $u^0(x) = |x|^2$. □

Instead of applying the Bernstein theorem for (4.18), we can prove Theorem 4.1 directly, by an argument similar to that for the standard Monge-Ampère equation (4.18) [25]. It is a rescaling argument based on the interior gradient estimate (4.10) and the a priori estimate (4.14).

Remark 4.4. In Theorem 4.1, f is assumed to be continuous. But in Theorem 1.3 and in Theorem 7.1 below, f can be discontinuous. In the case when f is discontinuous and $|f-1| \leq \varepsilon$ for some small $\varepsilon > 0$, one easily verifies that u^0 satisfies

$$1 - \varepsilon \leq \det D^2 u^0 \leq 1 + \varepsilon \tag{4.19}$$

in the viscosity sense. Denote

$$\det D^2 u^0 =: \tilde{f} \ \text{ in } \ \mathbb{R}^n. \tag{4.20}$$

By the weak continuity of the measure μ_{u,ρ^*} in (2.1) (see Corollary 3.1, [24]), one sees that $\tilde{f}$ is a weak limit of f^k. By (4.19), we also have $\tilde{f} \in L^\infty(\Omega)$ and $|\tilde{f} - 1| \leq \varepsilon$. Therefore by [3, 7], u^0 is strictly convex and of polynomial growth at $|x| = \infty$. In particular the covering property in §4.3 below also holds.

4.3. Covering property. From the Bernstein property (Theorem 4.1), the potential function u satisfies some nice properties. As was pointed out at the beginning of §2.2, we always assume that u is a strictly c-convex, C^1 smooth solution to (1.1). First we have the separating property.

Lemma 4.4. *For any given $\delta > 0$, there exists a constant $K > 1$ with the following property. Let $S^0_\mu(x), S^0_\nu(y)$ be two of the sub-level sets with $x, y \in \Omega_\delta$ and $\mu \geq \nu$. If $y \notin S^0_\mu(x)$, we have*

$$S^0_{\mu/K}(x) \cap S^0_{\nu/K}(y) = \emptyset, \tag{4.21}$$

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

Proof. To prove (4.21) it suffices to prove $S^0_{\mu/K}(x) \cap S^0_{\mu/K}(y) = \emptyset$. Namely we may assume $\mu = \nu$. But it simply follows from the Bernstein theorem by a compactness argument.

More precisely, if there exist $x_k, y_k \in \Omega_\delta$ and a sequence $\mu_k \to 0$ such that $y_k \notin S^0_{\mu_k}(x_k)$ but $S^0_{\mu_k/k}(x_k) \cap S^0_{\mu_k/k}(y_k) \neq \emptyset$, by the strict c-convexity of potential functions [29], we must have $|x_k - y_k| \to 0$. Choose $h_k \to 0$ such that $y_k \in \partial S^0_{h_k,u}(x_k)$. Then by Theorem 4.1, the rescaled function u^k converges to the quadratic function $|x|^2$. We get a contradiction when $K > 4$. □

Remark 4.5.
(i) If instead f satisfies $|f - 1| \leq \varepsilon$, for some small $\varepsilon > 0$, Lemma 4.4 is also true for some $K > 0$ depending only on the strict convexity of u^0, where u^0 is the limit function in (4.11), which satisfies (4.19) such that $u^0(0) = 0, u^0 \geq 0$ and S^0_{1,u^0} is normalized.
(ii) By a similar compactness argument, we also have the *engulfing property* for u. Namely, there exists a constant $K > 1$ such that for any $x, y \in \Omega_\delta$, if $y \in \overline{S^0_h}(x)$, then $S^0_h(x) \subset S^0_{Kh}(y)$.
(iii) The strict c-convexity of potential functions to (1.1) was proved in [29]. By Theorem 4.1 or the strict convexity of (4.19), the rescaled solution $\bar{u}$ is strictly c-convex with respect to the rescaled cost function $\bar{c}$, uniformly in the rescaling.

The following covering theorem is important for the proof of the $W^{2,p}$ estimate.

Theorem 4.2. *Let $A, B \Subset \Omega_\delta$ be two subsets of Ω_δ. Let $\mathcal{F} = \{S^0_{h_x,u}(x) : \ x \in B\}$ be a set of sub-level sets such that $A \subset \bigcup_{x \in B} S^0_{h_x,u}(x)$. Then there exists a sequence $\{S_i, \ i = 1, 2, \cdots\}$, where $S_i = S^0_{h_{x_i},u}(x_i)$, such that*

- (P1) $A \subset \bigcup S_i$;
- (P2) $\exists$ *a constant K such that any two sub-level sets in $\{S^0_{h_i/K}(x_i)\}$ are disjoint.*

Proof. We pick up $S_i = S^0_{h_{x_i},u}(x_i)$ one by one as follows. Let $M := \sup\{h : \ S^0_h(x) \in \mathcal{F}\}$ and pick up a sub-level set $S_1 = S^0_{h_1}(x_1)$ such that $h_1 > M/2$. Let

$$\mathcal{F}_1 = \{S^0_h(x) \in \mathcal{F} \colon x \in A - S_1\}, \qquad M_1 = \sup\{h : S^0_h(x) \in \mathcal{F}_1\}.$$

Pick $S_2 = S^0_{h_2}(x_2) \in \mathcal{F}_1$ such that $h_2 > M_1/2$.

Assume $S_1, S_2, \cdots, S_m$ have been chosen. If for some $m \geq 1$, $A - \bigcup_{i=1}^m S_i = \emptyset$, then we are through. Otherwise, in $A - \bigcup_{i=1}^m S_i$ we can select $S_{m+1} = S^0_{h_{m+1}}(x_{m+1})$ as above. Therefore we get a sequence $\{S_i : \ i = 1, 2, \cdots\}$.

By Lemma 4.4, (P2) is satisfied. By (P2) we have $h_i \to 0$ as $i \to \infty$, which in turn implies (P1) is satisfied. □

5. Eccentricity estimate

In this section, we give an estimate for the eccentricity of the sub-level sets S^0_h, or more precisely, the eccentricity of the minimum ellipsoid of S^0_h. The estimate was proved in [16] for the standard Monge-Ampère equation (1.5). Let Q_h be the linear transform given in (3.8) and let $\lambda_{h,max}$ and $\lambda_{h,min}$ be the largest and the smallest eigenvalues of Q_h. We give an estimate for the growth rate of $\frac{\lambda_{h,max}}{\lambda_{h,min}}$ as $h \to 0$. As a consequence we obtain Theorem 1.3. We also obtain the log-Lipschitz continuity of Du if ω_f satisfies an integral condition; and an upper bound for the second derivatives D^2u when f is Dini continuous, where

$$\omega_f(r) = \sup\{|f(x) - f(y)| : |x - y| < r\}. \tag{5.1}$$

First we prove the following theorem.

Theorem 5.1. *There exists a small constant $\varepsilon_0 > 0$ such that if $|f - 1| < \varepsilon_0$, then a strictly c-convex solution u to* (1.1) *satisfies the estimate*

$$|Du(x) - Du(y)| \leq Cd[1 + e^{-2\theta\psi(d)}] \ \ \forall \ x, y \in \Omega', \tag{5.2}$$

where $d = |x - y|$, $\Omega' \Subset \Omega$, the constants C and θ depend on $n, \varepsilon_0, dist(\Omega', \partial\Omega)$, and the c-convexity of u (both C and θ are uniformly bounded for small ε_0), and

$$\psi(d) = -\int_d^1 \frac{\omega_f(r)}{r} dr. \tag{5.3}$$

From Theorem 5.1 it follows a number of interesting estimates (including Theorem 1.3), which will be stated in Corollaries 5.1 and 5.2 below. To prove Theorem 5.1, we first state two lemmas, which correspond to Lemmas 4.1 and 4.2 in [22].

Lemma 5.1. *Let u be a C^2 smooth solution to*

$$det[u_{ij} - A_{ij}(x, Du)] = 1,$$

in a bounded convex domain Ω. *Suppose* u *vanishes on* $\partial\Omega$, *and at some point* $x_0 \in \frac{3}{4}\Omega$, $|D^2u(x_0)| \le C$. *Then the domain* Ω *is of good shape.*

Proof. By making a proper dilation of the coordinates x and the solution u, which does not change the condition and conclusion of the lemma, we may assume that the volume $|\Omega| \approx 1$. If the lemma is not true, we normalize the domain Ω. Then after normalization, the largest eigenvalue of $D^2u(x_0)$ becomes very large but the matrix A_{ij} is uniformly bounded [20, 22, 21]. Hence after normalization, the largest eigenvalue of the matrix Mu at x_0 becomes very large, which is in contradiction with the a priori estimate (4.14). □

For a convex domain Ω, we denote by $\alpha\Omega$ the α-dilation of Ω with respect to the center x_c of its minimum ellipsoid, namely $\alpha\Omega = \{x_c + \alpha(x - x_c) : \ x \in \Omega\}$. See also Remark 4.3(i) for the definition of good shape.

Lemma 5.2. *Let* $u^{(m)}, m = 1, 2$, *be two generalized solutions to* (1.1) *in* Ω. *Suppose* Ω *is a normalized convex domain. Then if* $\|u^{(m)}\|_{C^4} \le C_0$, $|u^{(1)} - u^{(2)}| \le \delta$ *in* $3\Omega/4$ *for some constant* $\delta > 0$, *we have, for* $1 \le k \le 3$,

$$|D^k(u^{(1)} - u^{(2)})| \le C\delta \quad \text{in} \ \ \tfrac{1}{2}\Omega.$$

Proof. Since both $u^{(1)}$ and $u^{(2)}$ are solutions to (1.1), $u^{(1)} - u^{(2)}$ satisfies a linearized Monge-Ampère equation, see (4.4) in [22]. By the assumption $\|u^{(m)}\|_{C^4} \le C_0$, the linearized Monge-Ampère equation is uniformly elliptic with smooth coefficients. Hence Lemma 5.2 follows from the Schauder estimate. □

Remark 5.1.
(i) Lemma 5.2 can be applied to two rescaled solutions of equation (1.1) obtained by the same rescaling. If the two functions $u^{(1)}$ and $u^{(2)}$ are obtained from the solution u of (1.1) by different rescaling, then the matrix $\bar{A}$ in (2.23) are different and the linearized Monge-Ampère equation in the proof of Lemma 5.2 is not homogeneous.
(ii) Note also that A_{ij} are bounded. Hence the condition $|D^2u(x_0)| \le C$ in Lemma 5.1 can be replaced by $|Mu(x_0)| < C$, where Mu is the matrix in (2.15).

Proof of Theorem 5.1. To prove (5.2) we may assume that x and y are not far away, so that $x \in S^0_{h,u}(y)$ for some small $h > 0$. There is no loss of generality in assuming that $y = 0$. By the normalization in §2, we may assume that $u \ge u(0) = 0$, $Du(0) = 0$, and the sub-level set S^0_h is convex. By (2.19)–(2.22) we may assume that S^0_1 is normalized, and $|A_{ij}|$ are small by Theorem 3.1.

To prove Theorem 5.1 we need to estimate the ratio $\frac{\lambda_{h,max}}{\lambda_{h,min}}$ for a sequence $h_k = N^{-k} \to 0$ for some large constant $N > 1$ (for simplicity we choose $N = 4$), where $\lambda_{h,max}$ and $\lambda_{h,min}$ are the largest and the smallest eigenvalues of Q_h, and Q_h is the linear transform given in (3.8). Denote

$$\nu(t) = \sup_{z\in B_1} \{|f(x) - f(y)| : \quad x, y \in S^0_{t^2}(z)\}, \quad \text{and } \nu_k = \nu(2^{-k}). \tag{5.4}$$

Step 1. Let $u_k, k = 0, 1, \cdots$, be the solution of

$$\begin{aligned} \det [D^2_{ij}u_k - A_{ij}(x, Du_k)] &= f_k \quad \text{in } U_k, \\ u_k &= u \quad \text{on } \partial U_k, \end{aligned} \tag{5.5}$$

where $U_k = S^0_{4^{-k},u}$, and $f_k = \inf_{U_k} f$. We make a linear transform

$$x^{(k)} = \bar{Q}^{(k)}x \ \text{ such that } \ D^2_{x^{(k)}}u_k(0) = I, \tag{5.6}$$

where I denotes the unit matrix. We can decompose $\bar{Q}^{(k)} = \bar{Q}_k \cdot \bar{Q}_{k-1} \cdots \bar{Q}_1 \cdot \bar{Q}_0$ as follows.

By the comparison principle (Lemma 2.3), $|u - u_0| \le C\nu_0$ and $|u - u_1| \le C\nu_1$. Hence

$$|u_0 - u_1| \le C\nu_0 \ \text{ in } \ U_1.$$

Since U_0 is normalized, by Lemma 4.3, $\|u_0\|_{C^4(S^0_{3/4,u})} \le C$. First we make a linear transform

$$x^{(0)} = \bar{Q}_0 x \quad \text{ such that } \quad D^2_{x^{(0)}}u_0(0) = I.$$

Here and below we use D to denote derivatives in x, and $D_{x^{(i)}}$ to denote derivatives in the new coordinates $x^{(i)}$, $i = 0, 1, 2, \cdots$.

By Lemma 5.1, the sub-level set U_1 in the coordinates $x^{(0)}$ has a good shape. By Lemma 4.3, $\|u_1\|_{C^4(S^0_{3/16,u})} \le C$ in the coordinates $x^{(0)}$. By Lemma 5.2 we obtain

$$\begin{aligned} |D_{x^{(0)}}u_0(x) - D_{x^{(0)}}u_1(x)| &\le C\nu_0, \\ |D^2_{x^{(0)}}u_0(x) - D^2_{x^{(0)}}u_1(x)| &\le C\nu_0 \end{aligned}$$

in the domain $S^0_{4^{-2},u_1}$, in the coordinates $x^{(0)}$.

By induction we assume that $D^2_{x^{(k-1)}}u_{k-1}(0) = I$. Then by Lemma 5.1, the sub-level set U_k has a good shape in the coordinates $x^{(k-1)}$. Make a linear transform

$$x^{(k)} = \bar{Q}_k x^{(k-1)} \quad \text{ such that } \quad D^2_{x^{(k)}}u_k(0) = I.$$

After the transform $\bar{Q}_k$, the sub-level set U_{k+1} has a good shape. Therefore the a priori estimate in Lemma 4.3 holds for $\hat{u}_0 := 4^k u_k(\frac{x}{2^k})$ and $\hat{u}_1 := 4^k u_{k+1}(\frac{x}{2^k})$. Hence applying Lemma 5.2 to $\hat{u}_0$ and $\hat{u}_1$, and scaling back we obtain

$$|D_{x^{(k)}}u_k(x) - D_{x^{(k)}}u_{k+1}(x)| \le C2^{-k}\nu_k, \tag{5.7}$$

$$|D^2_{x^{(k)}}u_k(x) - D^2_{x^{(k)}}u_{k+1}(x)| \le C\nu_k, \tag{5.8}$$

in $S^0_{4^{-k-2},u_{k+1}}$ in the coordinates $x^{(k)}$, where 2^{-k} in (5.7) is the scaling constant.

We point out here that the linear transform $\bar{Q}^{(k)}$ may not be unimodular, namely $\det \bar{Q}^{(k)}$ may not be equal to 1. Also one should note that $\bar{Q}^{(k)}$ is different from Q_{h_k} in (3.8). The transform $\bar{Q}^{(k)}$ normalizes the Hessian matrix $D^2u_k(0)$, but by definition (3.8), Q_{h_k} makes the sub-level set $S^0_{h_k}$ comparable to a ball of radius $\sqrt{h}$. But by Lemma 5.1, $Q^{-1}_{h_k}\cdot\bar{Q}^{(k)}$ is a uniformly bounded linear transform.

Step 2. We want to estimate $\lambda_k = \lambda_{max}(\bar{Q}^{(k)})$, the largest eigenvalue of $\bar{Q}^{(k)}$. Assume that in the coordinates $x^{(k-1)}$, u_k has the Taylor expansion (after a rotation of axes such that $D^2_{x^{(k-1)}}u_k(0)$ is diagonal)

$$u_k(x) = u_k(0) + a_{k,i}x_i + \frac{1}{2}b_{k,i}x_i^2 + O(|x|^3).$$

By induction assumption, $D^2_{x^{(k-1)}}u_{k-1}(0) = I$. Hence the Taylor expansion of u_{k-1} in the coordinates $x^{(k-1)}$ has the form

$$u_{k-1}(x) = u_{k-1}(0) + a_{k-1,i}x_i + \frac{1}{2}x_i^2 + O(|x|^3).$$

By (5.7),

$$|a_{k,i} - a_{k-1,i}| \le C2^{-k}\nu_k, \quad i = 1,\cdots,n.$$

Hence

$$|a_{k,i}| \le \sum_{j=k}^{\infty}|a_{j,i} - a_{j+1,i}| \le C2^{-k}\nu_k. \tag{5.9}$$

By (5.8),

$$|b_{k,i} - 1| \le 2\theta\nu_k, \quad i = 1,\cdots,n, \tag{5.10}$$

where θ is a constant independent of k. From the above Taylor expansion, the transform $\bar{Q}_k$ is given by

$$\bar{Q}_k x = (b^{1/2}_{k,1}x_1,\cdots,b^{1/2}_{k,n}x_n),$$

and in the coordinates $x^{(k)}$,

$$u_k(x) = u_k(0) + \frac{a_{k,i}}{b^{1/2}_{k,i}}x_i + \frac{1}{2}x_i^2 + O(|x|^3).$$

So the largest eigenvalue of $\bar{Q}_k$ is

$$\lambda_{\max}(\bar{Q}_k) = \max_{1\le j\le n} b^{1/2}_{k,j}.$$

By (5.10) we obtain

$$\lambda_{\max}(\bar{Q}_k) \leq 1 + \theta\nu_k, \tag{5.11}$$

which implies that

$$\det \bar{Q}_k = 1 + O(\nu_k).$$

From (5.11), we obtain

$$\begin{aligned} \lambda_k &\leq \prod_{j=0}^{k} \lambda_{max}(\bar{Q}_j) \leq \prod_{j=0}^{k}(1+\theta\nu_j) \\ &= e^{\sum_{j=0}^{k}\log(1+\theta\nu_j)} \leq e^{\theta\sum_{j=0}^{k}\nu_j}. \end{aligned} \tag{5.12}$$

Step 3. Next we estimate $|Du(z) - Du(0)|$, where z is a point near 0. We have

$$\begin{aligned} |Du(z) - Du(0)| \leq & |Du_k(z) - Du_k(0)| \\ & + |Du_k(0) - Du(0)| + |Du(z) - Du_k(z)| \\ =: & I_1 + I_2 + I_3, \end{aligned}$$

where we choose $k = k_z \geq 1$ such that $4^{-k-4} \leq u(z) \leq 4^{-k-3}$.

To estimate I_2, we have

$$I_2 = |Du_k(0) - Du(0)| \leq \sum_{i=k}^{\infty} |Du_i(0) - Du_{i+1}(0)|.$$

Recall that

$$|Du_i(0) - Du_{i+1}(0)| \leq \lambda_i |D_{x^{(i)}}u_i(0) - D_{x^{(i)}}u_{i+1}(0)|.$$

By (5.8),

$$|D_{x^{(i)}}u_i(0) - D_{x^{(i)}}u_{i+1}(0)| \leq C2^{-i}\nu_i.$$

We obtain

$$|Du_i(0) - Du_{i+1}(0)| \leq C\lambda_i 2^{-i}\nu_i,$$

where C is independent of i. Hence we obtain

$$I_2 \leq C\sum_{i=k}^{\infty} \lambda_i 2^{-i}\nu_i. \tag{5.13}$$

Similarly we can estimate I_3, namely

$$I_3 \leq C\sum_{i=k}^{\infty} \lambda_i 2^{-i}\nu_i. \tag{5.14}$$

Note that to get (5.14), one needs to repeat that above argument at z. Hence λ_i and ν_i in (5.14) are constants of u and f at z. However our estimates below for λ_i and ν_i apply to all points in Ω_δ. Therefore we may regard λ_i and ν_i as constants independent of the point in Ω_δ.

To estimate I_1, by (5.8) we have

$$|D^2_{x^{(i)}} u_i(x) - D^2_{x^{(i)}} u_{i+1}(x)| \le C\nu_i$$

for any $i = 0, 1, \cdots, k$, and $x \in \bar{Q}^{(i)}(S^0_{4^{-i-2},u})$. Hence

$$|D^2 u_i(x) - D^2 u_{i+1}(x)| \le C\lambda_i^2 \nu_i$$

for any $x \in S^0_{4^{-i-2},u}$.

Denote $h_i = u_i - u_{i-1}$. We have

$$|Dh_i(z) - Dh_i(0)| \le |D^2 h_i||z| \le C\lambda_i^2\nu_i|z|.$$

Hence,

$$\begin{aligned} I_1 &\le |Du_0(z) - Du_0(0)| + \sum_{i=1}^{k} |Dh_i(z) - Dh_i(0)| \\ &\le C|z|\big[1 + \sum_{i=0}^{k} \lambda_i^2 \nu_i\big]. \end{aligned} \tag{5.15}$$

We thus obtain

$$|Du(z) - Du(0)| \le C\sum_{i=k}^{\infty} \lambda_i 2^{-i}\nu_i + C|z|\big[1 + \sum_{i=0}^{k} \lambda_i^2\nu_i\big]. \tag{5.16}$$

By (5.12) we therefore obtain (recall that $Du(0) = 0$)

$$\begin{aligned} |Du(z)| &\le C\sum_{i=k}^{\infty} 2^{-i}\nu_i e^{\theta\sum_{j=0}^{i}\nu_j} + C|z|\left[1 + \sum_{i=0}^{k} \nu_i e^{2\theta\sum_{j=0}^{i}\nu_j}\right] \\ &\le C\int_0^{2^{-k}} \nu(t) e^{\theta\int_t^1 \frac{\nu(s)}{s}ds} dt + C|z|\left[1 + \int_{2^{-k}}^{1} \frac{\nu(t)}{t} e^{2\theta\int_t^1\frac{\nu(s)}{s}ds} dt\right]. \end{aligned} \tag{5.17}$$

Step 4. We simplify (5.17) to get (5.2). Denote

$$\varphi(t) = -\int_t^1 \frac{\nu(s)}{s} ds. \tag{5.18}$$

Assume that ν is small. Since $r\varphi'(r) = \nu(r)$ is small, both $\int_0^r e^{-\theta\varphi(t)}$ and $re^{-\theta\varphi(r)} \to 0$ as $r \to 0$. Then the first integral on the right hand side of (5.17) is equal to

$$\begin{aligned} \int_0^r \nu(t) e^{\theta\int_t^1\frac{\nu(s)}{s}ds} dt &= \int_0^r t\varphi'(t) e^{-\theta\varphi(t)} dt \qquad (r = 2^{-k}) \\ &= \frac{-r}{\theta} e^{-\theta\varphi(r)} + \frac{1}{\theta}\int_0^r e^{-\theta\varphi(t)}, \end{aligned}$$

and the second one is equal to

$$\int_r^1 \varphi'(t) e^{-2\theta\varphi(t)} dt = \frac{1}{2\theta}[e^{-2\theta\varphi(r)} - e^{-2\theta\varphi(1)}].$$

Moreover,

$$\lim_{r\to 0}\frac{\int_0^r e^{-\theta\varphi(t)}}{re^{-\theta\varphi(r)}} = \lim_{r\to 0}\frac{(\int_0^r e^{-\theta\varphi(t)})'}{(re^{-\theta\varphi(r)})'} = \lim_{r\to 0}\frac{e^{-\theta\varphi(r)}}{e^{-\theta\varphi(r)}(1-\theta r\varphi'(r))} = 1,$$

which implies that

$$\int_0^r e^{-\theta\varphi(t)} = O(re^{-\theta\varphi(r)}) \quad as \quad r\to 0.$$

Therefore from (5.17),

$$|Du(z)| \le C2^{-k}[1+e^{-\theta\varphi(2^{-k})}] + C|z|[1+e^{-2\theta\varphi(2^{-k})}]. \tag{5.19}$$

By the transform $x^{(k)} = \bar{Q}^{(k)}x$, the sub-level set U_k has a good shape. Let $\bar{u}(\bar{x}) = 4^k u\left(\frac{1}{2^k}(\bar{Q}^{(k)})^{-1}\bar{x}\right)$, then $\bar{u}$ satisfies the equation (2.23). Recall that $4^{-k-4}\le u(z)\le 4^{-k-3}$. Thus

$$4^{-4}\le \bar{u}(\bar{z})\le 4^{-3}, \qquad \text{where } \bar{z} = 2^k\bar{Q}^{(k)}z.$$

By the gradient estimate Lemma 4.2, we have $|\bar{z}|\ge C$. Hence

$$|\bar{Q}^{(k)}z|\ge C2^{-k}.$$

From (5.12) we then have

$$|z|\ge C2^{-k}\|\bar{Q}^{(k)}\|^{-1}\ge C2^{-k}e^{\theta\varphi(2^{-k})}. \tag{5.20}$$

If $e^{\theta\varphi(2^{-k})}\le 1$, then from (5.19) and (5.20),

$$|Du(z)| \le C2^{-k}e^{-\theta\varphi(2^{-k})} + C|z|e^{-2\theta\varphi(2^{-k})} \le C|z|e^{-2\theta\varphi(2^{-k})}.$$

If $e^{\theta\varphi(2^{-k})}>1$, then from (5.20), $|z|\ge C2^{-k}$. It is easily seen that

$$2^{-k}[1+e^{-\theta\varphi(2^{-k})}]\le C|z|[1+e^{-2\theta\varphi(2^{-k})}].$$

We have therefore proved that

$$|Du(z)-Du(0)|\le C|z|[1+e^{-2\theta\varphi(2^{-k})}]. \tag{5.21}$$

To obtain (5.2) from (5.21), recall that $u\in C^{1,\alpha}$ for some α close to 1, as noted at the end of Step 3. Hence for a small $\sigma>0$, the sub-level set $S^0_{t^2,u}(0)$ is contained in the ball $B_{t^{1-\sigma}}(0)$ provided $t>0$ is sufficiently small. In particular, we have $2^{-k}\ge |z|^{1+\sigma}$

and $\nu(t) \le \omega(t^{1-\sigma})$. With $d = |z|$ we then obtain

$$\begin{aligned} |\varphi(2^{-k})| &\le \int_{d^{1+\sigma}}^1 \frac{\nu(t)}{t} dt \le \int_{d^{1+\sigma}}^1 \frac{\omega(t^{1-\sigma})}{t} dt \\ &= \frac{1}{1-\sigma} \int_{d^{1+\sigma}}^1 \frac{\omega(t^{1-\sigma})}{t^{1-\sigma}} dt^{1-\sigma} \le \frac{1}{1-\sigma} \int_{d^{1-\sigma^2}}^1 \frac{\omega(t)}{t} dt \\ &= \frac{1}{1-\sigma} |\psi(d^{1-\sigma^2})| \le \frac{1}{1-\sigma} |\psi(d)|, \end{aligned}$$

where ψ is defined in (5.3). Therefore we get the desired estimate (5.2) with the constant θ replaced by $\frac{\theta}{1-\sigma}$ for any $\sigma > 0$ small, but fixed.

Theorem 5.1 has some interesting consequences.

Corollary 5.1.
(i) If $|f-1| \le \varepsilon$ for some small $\varepsilon > 0$, then $u \in C^{1+\alpha}_{loc}(\Omega)$ ($\alpha = 1 - 2\theta\varepsilon$).
(ii) If f is continuous, then $u \in C^{1+\alpha}_{loc}(\Omega)$ for any $\alpha \in (0,1)$.
(iii) If $\omega_f(r) \le \frac{C}{|\log r|}$ for some constant $C \le 1/2\theta$, then Du is log-Lipschitz continuous, namely

$$|Du(x) - Du(y)| \le C|x-y|(1 + |\log|x-y||). \tag{5.22}$$

(iv) If f is Dini continuous, namely $\int_0^1 \frac{\omega_f(r)}{r} dr < \infty$, then $\|u\|_{C^{1,1}(\Omega_\delta)} \le C$.

Note that Corollary 5.1(i) corresponds to Theorem 1.3. In the following we also use $\|Q\|$ to denote the largest eigenvalue of the matrix Q.

Corollary 5.2. *Let u be a strictly generalized solution to* (1.1), *$\det[Mu] = f$, where $|f-1| < \varepsilon$. Then for $h > 0$ small,*

$$\|Q_h\| \le Ch^{-C\varepsilon} \tag{5.23}$$

where the constant C is independent of h.

Proof. Assume that $4^{-k} \le h \le 4^{-k+1}$ for some $k > 1$. As was pointed out before, $\|Q_h\| \le C\|\bar{Q}^{(k)}\|$. By (5.12),

$$\|\bar{Q}^{(k)}\| \le e^{\theta \sum_{i=0}^k \nu_i} \le e^{(k+1)\theta\varepsilon}.$$

On the other hand, from $h \le 4^{-k+1}$ one has

$$k \le \frac{\ln h^{-1}}{\ln 4} + 1.$$

Combining the above inequalities, we obtain

$$\begin{aligned} \|Q_h\| &\le Ce^{\theta\varepsilon \ln h^{-1}/\ln 4} \\ &\le Ch^{-\theta\varepsilon/\ln 4}, \end{aligned}$$

for $h > 0$ small, where the constant C is independent of h. □

In [4] Caffarelli used a different rescaling and proved that $\|Q_h\| \le h^{-\varepsilon^\tau}$ for some constant $\tau > 1$. See Corollary 2 in [4]. Our estimate (5.23) improves the power from ε^τ to $C\varepsilon$. We will use the estimate in Section 7 to prove the $W^{2,p}$ estimate. We would like to point out that instead of (5.23), it is sufficient to use the $C^{1,\alpha}$ regularity of potential functions. We also refer the reader to the proof of the $W^{2,p}$ estimate in [15] which uses the engulfing property instead of the estimate $\|Q_h\| \le h^{-\varepsilon^\tau}$.

6. Density estimate for second derivatives

In this section we prove that when $|f-1| < \varepsilon$, then after normalization, the density of points where u has large second derivatives has small Lebesgue measure when ε is small. The argument in this and the next section was inspired by that in [4].

For a given small $\varepsilon > 0$, by Theorem 3.1 we may assume

$$|A_{ij}(\cdot, Du)|,\ |A_{ij,kl}(\cdot, Du)| \le \varepsilon, \tag{6.1}$$

$$|\mathrm{det} D^2_{xy} c - 1| \le \varepsilon. \tag{6.2}$$

The assumptions (6.1), (6.2) imply our $W^{2,p}$ estimate also depends on the rate of the convergence in Theorem 3.1, which in turn depends on the C^1 and the strict convexity estimates for potential functions obtained in [29].

Given $u \in C^0(\overline{\Omega})$, denote

$$\Gamma(u) = \sup v, \tag{6.3}$$

where the sup is taken in the set of all c-convex functions v satisfying $v \le u$ in Ω. Then $\Gamma(u)$ is c-convex, and is called the *c-convex envelope* of u in Ω. Apparently $\Gamma(u) \le u$.

The following lemma provides an estimate for the contact set

$$\mathcal{C} = \{x \in \Omega : \ \Gamma(u-v)(x) = (u-v)(x)\},$$

where $v \in C^{2,\alpha}$ is a c-convex function.

Lemma 6.1. *Let u be a solution to* (1.1) *and v be a c-convex function in Ω. Suppose $1 < f < 1+\varepsilon$, u is Lipschitz continuous, $v \in C^{2,\alpha}(\overline{\Omega})$. Then for any Borel set $E \subset \Omega$,*

$$|T_\Gamma(E)| \le \int_{\mathcal{C}\cap E} \left[(1+\varepsilon)^{1/n} + C\varepsilon - det^{1/n} Mv\right]^n \frac{dx}{|detD^2_{xy}c|}, \tag{6.4}$$

where $|\cdot|$ denotes the Lebesgue measure, T_Γ is the c-normal mapping of $\Gamma := \Gamma(u-v)$.

Proof. For any Borel set $E \subset \Omega$, denote

$$\mu(E) = |T_\Gamma(E)|. \tag{6.5}$$

Then μ is the measure defined in (2.1) (with $\rho^* = 1$) for the function Γ. It is a Radon measure supported on the contact set $\mathcal{C}$, namely it vanishes in the open set $\Omega - \mathcal{C}$. By the weak convergence of the measure μ [24], it suffices to prove Lemma 6.1 for smooth u (see Remark 6.1 below).

Recall that a measure μ can be decomposed as the sum of a (local integrable) regular part and a singular part, $\mu = \mu_r + \mu_s$. For the measure μ in (6.5), since u and v are smooth, the function Γ is $C^{1,1}$. Hence the singular part μ_s vanishes and the regular part is a L^∞ function. By the proof of Lemma 2.3 [28], one can verify that the regular part μ_r is given by

$$\mu_r = \frac{\det[\partial^2\Gamma - A(x, \partial\Gamma)]}{\det c_{x,y}},$$

where $\partial^2 u(x) = D^2 u(x)$ if u is twice differentiable at x, and $\partial^2 u(x) = 0$ otherwise.

Let $T_{(u-v)}$ be the mapping determined by $D(u-v)(x) = D_x c(x, T_{(u-v)}(x))$ (see (1.3)). We have

$$D^2(u-v)(x) = D_x^2 c(x, T_{(u-v)}(x)) + c_{x,y} \cdot DT_{(u-v)}.$$

Noting that for any point $x \in \mathcal{C}$, the function $u - v - \Gamma$ attains its local minimum 0 at x, we have

$$\frac{D^2\Gamma(x) - D_x^2 c(x, T_\Gamma(x))}{\det c_{x,y}} \le \frac{D^2(u-v)(x) - D_x^2 c(x, T_{(u-v)}(x))}{\det c_{x,y}} \tag{6.6}$$

at any twice differentiable point of Γ. Therefore for any Borel set $E \subset \Omega$,

$$\begin{aligned}\mu(E \cap \mathcal{C}) = \mu_r(E \cap \mathcal{C}) &= \int_{E\cap\mathcal{C}} \frac{\det[D^2\Gamma(x) - D_x^2 c(x, T_\Gamma(x))]}{\det c_{x,y}} \\ &\le \int_{E\cap\mathcal{C}} \frac{\det[D^2(u-v)(x) - D_x^2 c(x, T_{(u-v)}(x))]}{\det c_{x,y}} \\ &= \int_{E\cap\mathcal{C}} \frac{\det[M(u-v)(x)]}{\det c_{x,y}}.\end{aligned}$$

Write

$$\det[M(u-v)] = \det[Mu - Mv + N],$$

where Mu is the matrix in (2.15),

$$N_{ij} = A_{ij}(\cdot, Du) - A_{ij}(\cdot, Dv) - A_{ij}(\cdot, D(u-v)).$$

By (6.1), we have $|N_{ij}| \le C\varepsilon$ for all i, j. Therefore

$$\begin{aligned}\det{}^{1/n}[Mu - Mv + N] &\le \det{}^{1/n}[Mu] - \det{}^{1/n}[Mv - N] \\ &\le \det{}^{1/n}[Mu] - \det{}^{1/n}[Mv] + C\varepsilon \\ &\le (1+\varepsilon)^{1/n} + C\varepsilon - \det{}^{1/n}[Mv]\end{aligned}$$

where the first inequality is due to the concavity of $\det^{1/n}$, and the second is by the C^2-smoothness of v. $\square$

Remark 6.1. Note that under the assumption (6.1) and the assumptions in Lemma 6.1, a solution u to (1.1) can be approximated by smooth solutions. Note that to prove (6.4) it suffices to show that for any uniformly convex sub-domain $\tilde\Omega \Subset \Omega$, u can be approximated by smooth solutions in $\tilde\Omega$. Let $w \in C^2$ be a smooth, uniformly convex function, vanishing on $\partial\tilde\Omega$. When the constant ε in (6.1) is small, by direct computation we have

$$M(u+\sigma w) \geq M(u) \ \text{ in } \tilde\Omega$$

for $\sigma > 0$ small. Hence $u+\sigma w$ is a subsolution to (1.1) in $\tilde\Omega$. In particular, it implies that there is a smooth solution u_δ to (1.1) in $\tilde\Omega$ with f_δ, the mollification of f, such that $u_\delta = u$ on $\partial\tilde\Omega$. Next we give the estimate of the density of "good set" for the normalized solution.

We would also like to point out that, by approximation and the uniqueness of solutions to the Dirichlet problem [24], to prove Theorem 1.2 it suffices to prove it for smooth solutions.

Lemma 6.2. *Let u be a solution to* (1.1) *and w be a solution to $\det[Mw] = 1$ in Ω. Suppose $1 < f < 1+\varepsilon$, $w \in C^2(\Omega)$, $w - \varepsilon \leq u \leq w$. Let U be a unit ball compactly contained in Ω. Then*

$$\big|\{\Gamma(u - \frac{1}{2}w) = u - \frac{1}{2}w\} \cap U\big| \geq (1 - C\varepsilon^{1/2})|U|, \tag{6.7}$$

where the constant $C > 0$ depends on $\|w\|_{C^2(U)}$.

Proof. Set $v = \frac{1}{2}w$. We have

$$Mv = \frac{1}{2}Mw + \frac{1}{2}A(\cdot, Dw) - A(\cdot, \frac{1}{2}Dw).$$

Since $w \in C^2(\overline{U})$, by (6.1) we have

$$\det\left[D^2 v - A(\cdot, Dv)\right] = \frac{1}{2^n} + O(\varepsilon).$$

It also implies that when $\varepsilon > 0$ is small, $v = \frac{1}{2}w$ is uniformly c-convex.

By assumption,

$$\frac{1}{2}w - \varepsilon \leq u - \frac{1}{2}w \leq \frac{1}{2}w.$$

Since $\frac{1}{2}w$ is c-convex, we have

$$\frac{1}{2}w - \varepsilon \leq \Gamma(u - \frac{1}{2}w) \leq \frac{1}{2}w.$$

From Lemma 2.2,

$$N_{-\delta}\{T_{\frac{1}{2}w}(U_{C\varepsilon/\delta})\} \subset T_{\Gamma(u-\frac{1}{2}w)}(U),$$

where $U_\delta = \{x \in U : \ \mathrm{dist}(x, \partial U) > \delta\}$. Since $\frac{1}{2}w$ is uniformly c-convex, $T_{\frac{1}{2}w}$ and its inverse are smooth mappings. Hence

$$|N_{-\delta}\{T_{\frac{1}{2}w}(U_{C\varepsilon/\delta})\}| \geq |T_{\frac{1}{2}w}(U_{C\varepsilon/\delta})| - C\delta.$$

We obtain

$$|T_{\Gamma(u-\frac{1}{2}w)}(U)| \geq |T_{\frac{1}{2}w}(U_{C\varepsilon/\delta})| - C\delta. \tag{6.8}$$

Since $\frac{1}{2}w$ is c-convex and smooth,

$$\begin{aligned} |T_{\frac{1}{2}w}(U_{C\varepsilon/\delta})| &= \int_{U_{C\varepsilon/\delta}} \mathrm{det} DT_{\frac{1}{2}w} \\ &= \int_{U_{C\varepsilon/\delta}} \frac{\frac{1}{2^n} + O(\varepsilon)}{|\mathrm{det} D^2_{xy} c|} \\ &\geq \big(\frac{1}{2^n} - C\varepsilon\big)|U_{C\varepsilon/\delta}|. \end{aligned} \tag{6.9}$$

By Lemma 6.1 and assumption (6.2) we also have

$$\begin{aligned} |T_{\Gamma(u-\frac{1}{2}w)}(U)| &\leq \frac{\Big((1+\varepsilon)^{1/n} + C\varepsilon - \mathrm{det}^{1/n} Mv\Big)^n}{\inf |\mathrm{det} D^2_{xy} c|} |\mathcal{C} \cap U| \\ &\leq \big(\frac{1}{2^n} + C\varepsilon\big)|\mathcal{C} \cap U| \end{aligned} \tag{6.10}$$

where $\mathcal{C} = \{\Gamma(u - \frac{1}{2}w) = u - \frac{1}{2}w\}$ is the contact set. Combining (6.8)–(6.10) we obtain

$$\big(\frac{1}{2^n} - C\varepsilon\big)|U_{C\varepsilon/\delta}| \leq \big(\frac{1}{2^n} + C\varepsilon\big)|\mathcal{C} \cap U| + C\delta.$$

Let $\delta = \varepsilon^{1/2}$. We obtain

$$\frac{|\{\Gamma(u - \frac{1}{2}w) = u - \frac{1}{2}w\} \cap U|}{|U|} \geq 1 - C\varepsilon^{1/2}.$$

□

For the proof of the $W^{2,p}$ estimate, we choose the domain Ω in Lemma 6.2 the sub-level set $S^0_{2,u} = S^0_{2,u}(0)$, and the domain Ω in Lemma 6.1 any convex subset of $S^0_{3/2,u}$, for the normalized solution u in §2.4 (i.e. the function $\bar{u}$ in §2.4). To apply Lemma 6.2 we choose the domain $U := S^0_{1,u}(0)$ and assume U is normalized. Then by Section 5, U is close to a unit ball. Hence we have, instead of (6.7), we have

$$\big|\big\{\Gamma(u - \frac{1}{2}w) = u - \frac{1}{2}w\big\} \cap U\big| \geq (1 - \varepsilon')|U|, \tag{6.11}$$

where the constant $\varepsilon' \geq 0$ and $\varepsilon' \to 0$ as $\varepsilon \to 0$. We also choose w the solution to $\det[Mw] = 1$ in $S^0_{2,u}$ and $w = u$ on $\partial S^0_{2,u}$. Then $w \in C^2(S^0_{2,u})$ and $\|w\|_{C^4(S^0_{3/2,u})} \leq C$. By (6.1), w is uniformly convex in $S^0_{3/2,u}$ and satisfies

$$\det D^2 w = 1 + O(\varepsilon).$$

At any point $\bar{x} \in S^0_{1,u}$ in $\{\Gamma(u - \frac{1}{2}w) = u - \frac{1}{2}w\}$, there exists a c-support function φ of Γ, such that

$$\begin{cases} \varphi(x) \leq (u - \frac{1}{2}w)(x) & \text{near } \bar{x}, \\ \varphi(\bar{x}) = (u - \frac{1}{2}w)(\bar{x}). \end{cases}$$

Namely

$$\begin{cases} \varphi(x) + \frac{1}{2}w(x) \leq u(x) & \text{near } \bar{x}, \\ \varphi(\bar{x}) + \frac{1}{2}w(\bar{x}) = u(\bar{x}). \end{cases}$$

Note that by Theorem 3.1 or (6.2), the c-support function φ is sufficiently close to a linear function. But w is smooth and uniformly convex, as noted above. Hence we have

$$u(x) \geq \psi(x) + \frac{2}{N}|x - \bar{x}|^2 \ \text{ near } \bar{x}, \tag{6.12}$$

for some positive constant N, where ψ is the c-support of u at $\bar{x}$. From (6.12) we have

$$S^0_{h,u}(\bar{x}) \subset B_{\sqrt{Nh}}(\bar{x}) \quad h < h_x, \tag{6.13}$$

where $h_x > 0$ depends on x. By (6.1) (6.2) and from the equation (2.23), we also have

$$\frac{1}{N}|x - \bar{x}|^2 \leq u(x) - \ell(x) \leq N^{n-1}|x - \bar{x}|^2 \ \text{ near } \bar{x}, \tag{6.14}$$

where ℓ is the tangent plane of u at $\bar{x}$. From(6.11) we have

Lemma 6.3. *Let u be the normalized solution to* (2.23) *such that* (6.1) *and* (6.2) *hold. Assume that $1 < f < 1 + \varepsilon$. Then there is a set $E \subset S^0_1$ with*

$$|S^0_1 - E| \leq \varepsilon' \tag{6.15}$$

such that u satisfies (6.12) *at any point $\bar{x} \in E$, where $\varepsilon' \to 0$ as $\varepsilon \to 0$.*

7. $W^{2,p}$ ESTIMATE

We are now ready to prove the $W^{2,p}$ estimate. Let u be the normalized solution to (2.23) such that (6.1) holds. By the strict c-convexity of u (see Remark 4.5(iii)), there exists $h_0 > 0$ such that for any $x_0 \in S^0_{1,u}(0)$, the set $S^0_{h_0,u}(x_0) \subset S^0_{5/4,u}(0)$. Denote

$$D_k = \{x \in S^0_{1,u} : \ S^0_{h,u}(x) \subset B_{\sqrt{N^k h}}(x) \ \forall\, h < h_x\},$$
$$A_k = S^0_{1,u} - D_k, \quad k = 1, 2, \cdots.$$

We allow that the positive constant h_x depends on x. In other words, $x \in D_k$ if and only if $D^2u(x) > 2N^{-k}I$, where N is the constant in (6.12). Replacing N by a larger constant if necessary, we may also assume that

$$S^0_{h,u}(x) \subset B_{\sqrt{Nh_0}}(x) \ \ \forall\, x \in S^0_{1,u} \quad h \le h_0.$$

In the following lemma we show fast polynomial decay of the measure of A_k, which implies an L^p estimate of D^2u.

Lemma 7.1. *Given $q < \infty$, there exists an $\varepsilon = \varepsilon(q)$ such that if $|f-1| < \varepsilon$, then*

$$|A_k \cap B_{r_k}(0)| \le N^{-qk}, \tag{7.1}$$

where $r_1 = \frac{1}{n}$, r_{k+1} is given by

$$r_{k+1} = r_k - N^{-\frac{k}{8nC\varepsilon}} \ge \frac{1}{2n},$$

and C is the constant in (5.23).

Proof. For clarity we divide the proof into three steps.

Step 1: For $y \in A_k$, let

$$\hat h_y = \inf\{h : \ S^0_{h,u}(y) \subset B_{\sqrt{N^k h}}(y)\}. \tag{7.2}$$

By our definition of A_k and the choice of N, we have $0 < \hat h_y < h_0$.

We normalize $S^0_{\hat h_y,u}(y)$ to $\bar S$ and u to $\bar u$ as in §2. By Lemma 6.3, after the normalization there exists a set $\bar E \subset \bar S$ with $|\bar S - \bar E| \le \varepsilon'$ such that $\bar u$ satisfies (6.12) at any point $x \in \bar E$. That is (6.13) holds for $\bar u$ in the coordinates after normalization.

Note that if $x_0 \in D_{k-n+1}$, then after normalization (by our choice of $\hat h_y$ in (7.2)), the sub-level set $S^0_{h,u}(x_0)$ cannot have a good shape,

$$S^0_{h,u}(x_0) \not\subset B_{\sqrt{N^2h}}(x_0) \tag{7.3}$$

when h is sufficiently small. To see (7.3), let $\mathcal{M}$ be the normalization (matrix) which sends $S^0_{\hat h_y,u}(y)$ to $\bar S$. Then the least eigenvalue of $\mathcal{M}$ is bounded by N^{-k}. Hence for $x_0 \in D_{k-n+1}$, the inner-radius of the set $\mathcal{M}S^0_{h,u}(x_0)$ is bounded by $N^{-n+1}\sqrt{h}$, which in turn implies that the circum-radius of $\mathcal{M}S^0_{h,u}(x_0)$ is greater than $N\sqrt{h}$.

Scaling back to u in the set $S^0_{\hat h_y,u}(y)$, we see that there is a set $E_y \subset S^0_{\hat h_y,u}(y)$ with

$$\frac{|S^0_{\hat h_y,u}(y) - E_y|}{|S^0_{\hat h_y,u}(y)|} \le \varepsilon' \tag{7.4}$$

such that for any $x \in E_y$,

$$S^0_{h,u}(x) \subset B_{\sqrt{N^{k+1}h}}(x) \ \ \forall\, h \le h_x, \tag{7.5}$$

for some $h_x > 0$. Hence

$$E_y \subset D_{k+1}, \quad i.e.\ E_y \cap A_{k+1} = \emptyset. \tag{7.6}$$

But by (7.3), if $x \in E_y$, then $x \notin D_{k-n+1}$. Hence we also have

$$E_y \subset A_{k-n+1}. \tag{7.7}$$

Step 2: Note that if a sub-level set $S^0_{h,u}(x)$ has the property that $\|Q_h\| \geq t$, where Q_h is the linear transform which makes the sub-level set $S^0_{h,u}$ comparable to a ball of radius $\sqrt{h}$, as defined in (3.8), then from Corollary 5.2

$$t \leq Ch^{-C\varepsilon}, \quad \text{i.e. } h \leq (t/C)^{-\frac{1}{C\varepsilon}}.$$

Recall that $Q_h(S^0_{h,u}(x)) \sim B_{h^{1/2}}$. We have

$$\begin{aligned} \mathrm{diam}(S^0_{h,u}(x)) &\leq Ch^{\frac{1}{2}}h^{-C\varepsilon} \leq h^{\frac{1}{4}} \\ &\leq t^{-\frac{1}{8C\varepsilon}}. \end{aligned}$$

For any $y \in A_{k+1}$, let $\hat{h}_y$ be given in (7.2) and let $Q_{\hat{h}_y}$ be the linear transform which normalizes $S^0_{\hat{h}_y,u}$, as defined in (3.8). Then by (7.2) one easily verifies that

$$\|Q_{\hat{h}_y}\| \geq N^{k/n}. \tag{7.8}$$

Hence we obtain

$$\mathrm{diam}(S^0_{\hat{h}_y,u}(y)) \leq N^{-\frac{k}{8Cn\varepsilon}}. \tag{7.9}$$

By our choice of r_k in Lemma 7.1, we see that if $y \in A_{k+1} \cap B_{r_{k+1}}(0)$, then

$$S^0_{\hat{h}_y,u}(y) \subset B_{r_k}(0). \tag{7.10}$$

Step 3: The set of all sub-level sets $S^0_{\hat{h}_y,u}(y)$, with $\hat{h}_y$ given by (7.2), is obviously a covering of $A_k \cap B_{r_k}(0)$. By Theorem 4.2, there exists a countable set $\{y_i\} \subset A_k \cap B_{r_k}(0)$, $i = 1, 2, \cdots$, such that

(P1), $A_k \cap B_{r_k}(0) \subset \bigcup S^0_{\hat{h}_{y_i},u}(y_i)$; and

(P2), any two sub-level sets in $\{S^0_{\hat{h}_i/K,u}(y_i)\}$ are disjoint,

where the constant K depends on n and ε but is independent of u.

Denote $S_i = S^0_{\hat{h}_{y_i},u}(y_i)$ and $S'_i = S^0_{\hat{h}_{y_i}/K}(y_i)$. From (7.4), there exists a set $E_{y_i} \subset S_i$ such that

$$\frac{|S_i - E_{y_i}|}{|S_i|} \leq \varepsilon', \tag{7.11}$$

and E_{y_i} satisfies (7.6), (7.7), from which we also have

$$
\begin{aligned}
&|S_i'| \le (1+\varepsilon')|S_i' \cap A_{k-n+1}|,\\
&|A_{k+1} \cap S_i| \le \varepsilon'|S_i|.
\end{aligned}
\tag{7.12}
$$

Since $A_{k+1} \subset A_k$, by (7.10) and (7.12) we can now estimate

$$
\begin{aligned}
|A_{k+1} \cap B_{r_{k+1}}(0)| &\le \sum_i |A_{k+1} \cap S_i| \\
&\le \sum \varepsilon'|S_i| \\
&\le \sum \varepsilon' K^n |S_i'| \\
&\le \sum \varepsilon' K^n |S_i' \cap A_{k-n+1}| \\
&\le \varepsilon' K^n |A_{k-n+1} \cap B_{r_{k-n+1}}(0)|.
\end{aligned}
$$

Therefore given $q < \infty$, we can choose $\varepsilon = \varepsilon(q)$ small enough such that $\bar{\varepsilon} := \varepsilon' K^n \le N^{-q}$. We obtain the desired estimate (7.1). □

Theorem 1.2 now follows from Lemma 7.1 easily.

Proof of Theorem 1.2. Let u be the generalized solution to (1.1). For a given $p \ge 1$, we choose a small $h_0 > 0$, such that for any point $x \in \Omega'$ (where $\Omega' \Subset \Omega$ is the set in Theorem 1.2), the equation (2.23) and the sub-level set $S^0_{h_0,u}(x)$, after normalization as in §2, satisfy (6.1) and (6.2). By the proof of Theorem 4.2, there exist finitely many sub-level sets $S^0_{h_0,u}(y_i), i = 1, 2, \cdots, m$, such that $\Omega' \subset \bigcup S^0_{h_0/K,u}(y_i)$; and any two sub-level sets in $\{S^0_{h_0/K^2,u}(y_i)\}$ are disjoint (we also choose the constant $K \ge 4n$). Therefore it suffices to prove Theorem 1.2 on these sub-level sets $S^0_{h,u}(y_i), i = 1, 2, \cdots, m$.

For a given sub-level set $S^0_{h_0,u}(y_i)$ where $1 \le i \le m$, we assume it is normalized and apply Lemma 7.1 in the set $S^0_{h_0,u}(y_i)$. Denote $A_k' = A_k - A_{k+1}$. Then in A_k', we have $|D^2u| \le CN^{nk}$. Hence

$$
\begin{aligned}
\int_{S^0_{h_0/K,u}(y_i)} |D^2u|^p dx &\le C \int_{B_{1/2n}(0)} |D^2u|^p dx \\
&\le C \sum N^{nkp} |A_k'| \\
&\le C \sum N^{nkp}\, N^{-kq}.
\end{aligned}
$$

Letting $q > np$ and summing over i from 1 to m, we obtain $u \in W^{2,p}(\Omega')$. □

For a fixed $p \geq 1$, our proof of $W^{2,p}$ estimate does not require the continuity of f. It suffices to assume $|f-1| < \varepsilon$ for a sufficiently small $\varepsilon > 0$. Therefore we have actually proved the following result.

Theorem 7.1. *Let u be a generalized solution to* (1.1). *Assume the cost function c satisfies (A1)–(A3), Ω^* is c^*-convex with respect to Ω. Then for any given $p \geq 1$, there exists a small $\varepsilon > 0$ such that if $|f-1| \leq \varepsilon$, $D^2u \in L^p(\Omega')$ $\forall$ $\Omega' \Subset \Omega$, and we have the estimate*

$$\|u\|_{W^{2,p}(\Omega')} \leq C, \tag{7.13}$$

where C depends on n, p, ε, $\Omega, \Omega', \Omega^*$.

References

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

[2] Caffarelli, L. A., Interior a priori estimates for solutions of fully nonlinear equations. *Ann. of Math.* (2) 130 (1989), 189–213.

[3] Caffarelli, L. A., A localization property of viscosity solutions to the Monge-Ampère equation and their strict convexity. *Ann. of Math.* (2) 131 (1990), 129–134.

[4] Caffarelli, L. A., Interior $W^{2,p}$ estimates for solutions of Monge-Ampère equations. *Ann. of Math.* 131 (1990), 135–150.

[5] Caffarelli, L. A., Allocation maps with general cost functions. in *Partial differential equations and applications*, 29–35, Lecture Notes in Pure and Appl. Math., 177, *Dekker, New York,* 1996.

[6] Caffarelli, L. A., Nonlinear elliptic theory and the Monge-Ampère equation. Proc. Inter. Congress Math. Vol.1, 179-187, 2002.

[7] Chou, K. S. and Wang, X-J., Entire solutions of the Monge-Ampère equation. *Comm. Pure Appl. Math.* 49 (1996), 529–539.

[8] De Philippis, G. and Figalli, A., Sobolev regularity for Monge-Ampère type equations. *SIAM J. Math. Anal.* 45 (2013), 1812–1824.

[9] De Philippis, G. and Figalli, A., $W^{2,1}$ regularity for solutions of the Monge-Ampère equation. *Invent. Math.* 192 (2013), 55–69.

[10] De Philippis, G., Figalli, A. and Savin, O., A note on interior $W^{2,1+\varepsilon}$ estimates for the Monge-Ampère equation. *Math. Ann.* 357 (2013), 11-22.

[11] Figalli, A., Kim, Y.-M. and McCann, R., Hölder continuity and injectivity of optimal maps. *Arch. Rat. Mech. Anal.* 209 (2013), 747–795.

[12] Figalli, A., Rifford, L. and Villani, C., On the Ma-Trudinger-Wang curvature on surfaces. *Calc. Var. PDE* 39 (2010), 307–332.

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

[14] Gilbarg, D. and Trudinger, N. S., *Elliptic partial differential equations of second order. Springer-Verlag, Berlin*, 1983.

[15] Gutiérrez, C. E., *The Monge-Ampère equation.* Progress in Nonlinear Differential Equations and their Applications, 44. *Birkhäuser Boston, Inc., Boston, MA,* 2001.

[16] Jian, H. Y. and Wang, X.-J., Continuity estimates for the Monge-Ampère equation. *SIAM J. Math. Anal.* 39 (2007), 608–626.

[17] John, F., Extremum problems with inequalities as subsidiary conditions. Studies and Essays Presented to R. Courant on his 60th Birthday, p.187–204, Interscience Publishers, New York, 1948.

[18] Kim, Y. H., and McCann, R., Continuity, curvature, and the general covariance of optimal transportation. *J. Euro. Math. Soc.* 12 (2010), 1009–1040.

[19] Levin, V., Abstract cyclical monotonicity and Monge solutions for the general Monge-Kantorovich problem, *Set-Valued Anal.* 7 (1999), 7–32.

[20] Liu, J., Hölder regularity of optimal mappings in optimal transportation. *Calc. Var. PDE.* 34 (2009), 435–451.

[21] Liu, J., Trudinger, N. S., On Pogorelov estimates for Monge-Ampere type equations, Discrete and Continuous Dynamical Systems, Vol. 28(2010), 1121-1135.

[22] Liu, J., Trudinger, N. S. and Wang, X.-J., Interior $C^{2,\alpha}$ regularity for potential functions in optimal transportation. *Comm. Partial Differential Equations*, 35 (2010), 165–184.

[23] Loeper, G., On the regularity of solutions of optimal transportation problems. *Acta Math.* 202 (2009), 241–283.

[24] Ma, X. N., Trudinger, N. S. and Wang, X.-J., Regularity of potential functions of the optimal transportation problem. *Arch. Rat. Mech. Anal.*, 177 (2005), 151–183.

[25] Pogorelov, A. V., *The Minkowski Multidimensional Problem.* Scripta Series in Mathematics. *New York: J.Wiley*, 1978.

[26] Schmidt, T., $W^{2,1+\varepsilon}$ estimates for the Monge-Ampère equation. *Adv. Math.* 240 (2013), 672–689.

[27] Trudinger, N. S., Recent developments in elliptic partial differential equations of Monge-Ampère type. *International Congress of Mathematicians. Vol. III*, 291–301, *Eur. Math. Soc., Zürich*, 2006.

[28] Trudinger, N. S. and Wang, X.-J., The Monge-Ampère equation and its geometric applications. *Handbook of geometric analysis. No. 1,* 467–524, Adv. Lect. Math. (ALM), 7, *Int. Press, Somerville, MA,* 2008.

[29] Trudinger, N. S. and Wang, X.-J., On strict convexity and continuous differentiability of potential functions in optimal transportation. *Arch. Rat. Mech. Anal.*, 192 (2009), 403–418.

[30] Trudinger, N. S. and Wang, X.-J., On the second boundary value problem for Monge-Ampère type equations and optimal transportation, Ann. Sc. Norm. Super. Pisa Cl. Sci. (5) 8 (2009), 143–174.

[31] Urbas, J., *Mass transfer problems*, Lecture Notes, Univ. of Bonn, 1998.

[32] Villani, C., *Topics in optimal transportation*, Grad. Stud. Math. 58, *Amer. Math. Soc.*, 2003.

[33] Villani, C., *Optimal transport. Old and new*, Grundlehren Math. Wiss. 338, *Springer-Verlag, Berlin*, 2009.

[34] Villani, C., Stability of a 4th-order curvature condition arising in optimal transport theory. *J. Funct. Anal.* 255 (2008), 2683–2708.

[35] Wang, X.-J., On the design of a reflector antenna, *Inverse Problems*, 12(1996), 351–375.

[36] Wang, X.-J., Some counterexamples to the regularity of Monge-Ampère equations. *Proc. Amer. Math. Soc.* 123 (1995), no. 3, 841–845.

[37] Wang, X.-J., Schauder estimate for elliptic and parabolic equations. *Chinese Ann. of Math.*, Series B, 27 (2006), 637–642.

Institute for Mathematics and its Applications, School of Mathematics and Applied Statistics, University of Wollongong, Wollongong, NSW 2522, Australia.

E-mail address: jiakunl@uow.edu.au

Centre for Mathematics and Its Applications, The Australian National University, Canberra, ACT 0200, Australia.

E-mail address: Neil.Trudinger@anu.edu.au, Xu-Jia.Wang@anu.edu.au