

# PARTIAL REGULARITY OF SOLUTIONS TO THE SECOND BOUNDARY VALUE PROBLEM FOR GENERATED JACOBIAN EQUATIONS

YASH JHAVERI<sup>†</sup>

ABSTRACT. We prove that outside of a closed singular set of measure zero solutions to the second boundary value problem for generated Jacobian equations are smooth.

## 1. INTRODUCTION

In this paper, we begin the development of a partial regularity theory for solutions to the second boundary value problem for a class of prescribed Jacobian equations called generated Jacobian equations. A *prescribed Jacobian equation* (PJE) takes the form

$$(1.1) \quad \det(\nabla_x[\mathbf{T}(x, \mathbf{u}(x), \nabla \mathbf{u}(x))]) = \psi(x, \mathbf{u}(x), \nabla \mathbf{u}(x))$$

where  $\mathbf{T} = \mathbf{T}(x, u, p) : \text{dom } \mathbf{T} \subset \Omega \times \mathbb{R} \times \mathbb{R}^n \rightarrow \mathbb{R}^n$  and  $\psi = \psi(x, u, p) : \Omega \times \mathbb{R} \times \mathbb{R}^n \rightarrow \mathbb{R}$ ; and the *second boundary value problem* (SBVP) asks that

$$(1.2) \quad \mathbf{T}_{\mathbf{u}}(\Omega) = \Upsilon$$

for some given  $\Upsilon \subset \mathbb{R}^n$ . Here,  $\Omega$  and  $\Upsilon$  are open sets and  $\mathbf{T}_{\mathbf{u}}(x) := \mathbf{T}(x, \mathbf{u}(x), \nabla \mathbf{u}(x))$ . We consider the specific case when this prescription is given through the push-forward condition

$$(\mathbf{T}_{\mathbf{u}})_{\#} f = g$$

for two probability densities  $f$  and  $g$  supported in  $\Omega$  and  $\Upsilon$  respectively. This corresponds to

$$(1.3) \quad \psi(x, \mathbf{u}(x), \nabla \mathbf{u}(x)) = \frac{f(x)}{g(\mathbf{T}(x, \mathbf{u}(x), \nabla \mathbf{u}(x)))}.$$

When the map  $\mathbf{T}$  is generated by a function  $\mathbf{G} : \text{dom } \mathbf{G} \subset \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R} \rightarrow \mathbb{R}$ , we find ourselves in the world of *generated Jacobian equations* (GJEs), and we call  $\mathbf{T}_{\mathbf{u}}$  the *transport map associated to  $\mathbf{G}$  and  $\mathbf{u}$* .

PJEs, in particular, GJEs, encompass many problems in analysis, economics, and geometry (see [11] for a discussion of some of these problems as well as the many references therein). The simplest PJE, when  $\mathbf{T} = \mathbf{T}(p) = p$ , is the Monge–Ampère equation, and the SBVP corresponds to the optimal transportation problem for quadratic cost. Here, smooth data does not ensure the existence of a smooth solution. Rather, the problem requires an additional, strong geometric condition on the support of the target density  $g$  for such a statement to hold. Specifically, we need  $\Upsilon$  to be convex, as Caffarelli showed in [2]. (See [12] for an investigation of how important, quantitatively, the convexity of  $\Upsilon$  is in guaranteeing the regularity of  $\nabla \mathbf{u}$ .) When  $\mathbf{T}$  is highly non-linear in its variables, the complexity of the problem is compounded. If  $\mathbf{T} = \mathbf{T}(x, p)$ , as it does in the optimal transport problem for general cost, not only do we need to place geometric restrictions on  $\Upsilon$ , but the associated generator  $\mathbf{G}$  of  $\mathbf{T}$  must obey certain structural conditions to first ensure the production of  $\mathbf{T}$  and second guarantee the regularity of solutions. In the most complex situations, wherein  $\mathbf{T} = \mathbf{T}(x, u, p)$ , analogous geometric conditions on the target domain and structure conditions on  $\mathbf{G}$  are still insufficient to yield regular solutions. This phenomenon is exhibited, for instance, in the reflector shape design problem: Karakhanyan and Wang, in [13], showed that smooth data may produce distinct solutions with vastly different regularity.<sup>1</sup>

The distinguishing feature of a general GJE from the optimal transport case is the map  $\mathbf{T}$  may depend on the values of the potential  $\mathbf{u}$ . This feature will be the source of the challenges faced in this work. The third coordinate of  $\mathbf{G}$  in the optimal transport case for cost  $c$  is a simple height parameter. Changes in this variable translate to vertical shifts in the graph of  $\mathbf{G}(x, y, v) = -c(x, y) - v$ . In general, changes in the third variable of an arbitrary generating function affect the shape of the graph  $\mathbf{G}$  (see, e.g., [11, 13, 15]).

<sup>†</sup> Supported in part by the ERC grant “Regularity and Stability in Partial Differential Equations (RSPDE)”.

<sup>1</sup> In the optimal transport problem, potentials  $\mathbf{u}$  are unique up to the addition of a constant. In the near-field reflector problem, however, we find that solutions may not be unique in any natural sense.

As noted, the GJE setting is one in which the map  $\mathbf{T}$  is produced from another  $\mathbf{G}$ . In the optimal transport problem for cost  $c$ , the map  $\mathbf{T}$  is generated by the equation

$$D_x \mathbf{G}(x, \mathbf{T}) = -D_x c(x, \mathbf{T}) = p.$$

Generally, the map  $\mathbf{T}$  (along with another  $\mathbf{V}$ ) is generated through the system of equations

$$\begin{cases} D_x \mathbf{G}(x, \mathbf{T}, \mathbf{V}) = p \\ \mathbf{G}(x, \mathbf{T}, \mathbf{V}) = u. \end{cases}$$

As such,  $\mathbf{G}$  must satisfy a collection of basic structure conditions to produce  $\mathbf{T}$ .

**1.1. Structure of  $\mathbf{G}$ .** Our starting assumptions are three-fold: 1.  $\text{dom } \mathbf{G} = X \times Y \times I$  where  $X, Y \subset \mathbb{R}^n$  are open and  $I \subset \mathbb{R}$  is an open interval, 2.  $\mathbf{G}$  is of class  $C_{\text{loc}}^{2,\alpha}(X \times Y \times I)$  for some  $\alpha \in (0, 1)$ , and 3.

$$\text{(G-Mono)} \quad D_v \mathbf{G}(x, y, v) < 0.$$

Up to a change of variables, we let  $I = \mathbb{R}$ .

The remaining structure conditions on  $\mathbf{G}$  will hold on a subset of the domain of  $\mathbf{G}$ :

$$\mathfrak{g} := \{(x, y, v) : v \in V_{x,y}\};$$

for each pair  $(x, y) \in X \times Y$ , the set  $V_{x,y}$  is some open interval (possibly empty). We assume that  $\mathfrak{g}$  is open. In keeping with the nomenclature of [11], the final structure conditions we impose on  $\mathbf{G}$  are as follows. First, we ask that the map

$$\text{(G-Twist)} \quad (y, v) \mapsto (D_x \mathbf{G}(x, y, v), \mathbf{G}(x, y, v)) \text{ is injective on } \{(y, v) : (x, y, v) \in \mathfrak{g}\}.$$

Second, we assume that the map

$$\text{(G*-Twist)} \quad x \mapsto -\frac{D_y \mathbf{G}(x, y, v)}{D_v \mathbf{G}(x, y, v)} \text{ is injective on } \{x : (x, y, v) \in \mathfrak{g}\}.$$

Third, we suppose that

$$\text{(G-Nondeg)} \quad \det \left( D_{xy} \mathbf{G} - D_{xv} \mathbf{G} \otimes \frac{D_y \mathbf{G}}{D_v \mathbf{G}} \right) \neq 0 \text{ on } \mathfrak{g}.$$

We shall make some remarks on these conditions in Section 2.

**1.2. Statement of Main Result.** In [15], the local regularity of solutions to our SBVP is also studied. Under a pair of higher-order structural assumptions on the generating function  $\mathbf{G}$  and geometric restrictions on the open, bounded sets  $\Omega$  and  $\Upsilon$ , solutions are proved to be smooth (given smooth densities bounded away from zero and infinity in  $\Omega$  and  $\Upsilon$  respectively), and the transport associated to  $\mathbf{G}$  and  $\mathbf{u}$  is shown to be a diffeomorphism from  $\Omega$  onto  $\Upsilon$ . These assumptions are extensions of the MTW conditions on the cost  $c$  and the  $c$ -convexity and  $c^*$ -convexity requirements on the source and target domains in the optimal transport problem for general cost (see [14]). We refer the reader to [11] for other results on the regularity of solutions to general generated Jacobian equations under different, but related, additional conditions on the structure of  $\mathbf{G}$  and on the geometry of the domains of the equation.

The purpose of this paper is to show that solutions to (1.1) – (1.3) are smooth outside a singular set of measure zero without the presence of any additional structural or geometric conditions. Precisely, our main result is the following:

**Theorem 1.1.** *Let  $\mathbf{G}$  and  $\mathfrak{g}$  be as in Section 1.1 and  $\Omega \subset X$  and  $\Upsilon \subset Y$  be two open, bounded sets. Suppose  $f : \Omega \rightarrow \mathbb{R}^+$  and  $g : \Upsilon \rightarrow \mathbb{R}^+$  are two continuous probability densities bounded away from zero and infinity and  $\mathbf{u} : \Omega \rightarrow \mathbb{R}$  is a  $\mathbf{G}$ -convex function such that  $(\mathbf{T}_{\mathbf{u}})_{\#} f = g$ . Then, for every  $\beta < 1$ , there exist two relatively closed sets  $\mathcal{S}_{\Omega} \subset \Omega$  and  $\mathcal{S}_{\Upsilon} \subset \Upsilon$  of measure zero such that  $\mathbf{T}_{\mathbf{u}} : \Omega \setminus \mathcal{S}_{\Omega} \rightarrow \Upsilon \setminus \mathcal{S}_{\Upsilon}$  is a homeomorphism of class  $C_{\text{loc}}^{0,\beta}$ . If, in addition,  $\mathbf{G} \in C_{\text{loc}}^{k+2,\alpha}(\Omega \times \Upsilon \times \mathbb{R})$ ,  $f \in C_{\text{loc}}^{k,\alpha}(\Omega)$ , and  $g \in C_{\text{loc}}^{k,\alpha}(\Upsilon)$  for some  $k \geq 0$  and  $\alpha \in (0, 1)$ , then  $\mathbf{T}_{\mathbf{u}} : \Omega \setminus \mathcal{S}_{\Omega} \rightarrow \Upsilon \setminus \mathcal{S}_{\Upsilon}$  is a diffeomorphism of class  $C_{\text{loc}}^{k+1,\alpha}$ .*

Notice that when  $f$  and  $g$  are just assumed to be continuous, the regular sets depend on the value of  $\beta$ . In the higher regularity cases, the regular sets are independent of the values of  $k$  and  $\alpha$ . Recall that the SBVP for GJEs may, in general, have many solutions, all with potentially different regularity properties (again, see [13, 11]). Yet by Theorem 1.1, outside sets of measure zero, all of this variety is unseen.

The proof of Theorem 1.1 follows the global strategy of its optimal transport predecessor [5, Theorem 1.3]. However, new difficulties arise here coming from the additional non-linear nature of general generating functions over those that arise in optimal transportation and the non-existence of a Kantorovich formulation of the problem. In particular, the third component  $v$  of  $\mathbf{G}$  plays no role in [5], while its presence here is pervasive.

As far as we know, Theorem 1.1 is the first partial regularity result on general GJEs. That said, in the optimal transport setting, the first partial regularity result was proved by Figalli in two dimensions for quadratic cost in [8]. This two dimensional, quadratic cost result was subsequently pushed to arbitrary dimension by Figalli and Kim in [9] and then again by Goldman and Otto in [10]. In [5], De Philippis and Figalli extended these last quadratic cost results to general cost, while Chen and Figalli proved a partial Sobolev regularity result for general cost in [4]. Finally, we mention that the  $\varepsilon$ -regularity techniques developed by De Philippis and Figalli, in [5], and exploited here have been used to prove regularity results at the boundary for optimal transports in [3] and [12].

**1.3. Organization.** This paper has four additional sections. In Section 2, we introduce some more notation and some preliminary results. Section 3 is dedicated to the proof Theorem 1.1. Finally, in the last two sections, we prove the local regularity results around which the proof of our main result revolves.

## 2. PRELIMINARIES

In this section, we introduce some notation and preliminary results. We start with some remarks on  $\mathbf{G}$  and the structure conditions it obeys. Then, we visit the geometry of solutions to the SVBP for GJEs. Finally, we show that solutions to (1.1) – (1.3) satisfy a Monge–Ampère-type equation almost everywhere.

**2.1. Structure and Duality.** The assumption that  $\mathfrak{g}$  is open is mild. For instance, in the near-field reflector/reflector shape design problem [13], an important model setting for the SBVP for general GJEs — wherein we have non-uniqueness of solutions and varying regularity among solutions — the set  $\mathfrak{g}$  is open (see [11, Section 3.1]). More generally, as far as we know, the set  $\mathfrak{g}$  is open in all examples of GJEs.

Thanks to (G-Mono), there exists a unique function  $\mathbf{H}$  determined by the equation

$$\mathbf{G}(x, y, \mathbf{H}(x, y, u)) = u,$$

and  $\mathbf{H}(x, y, \cdot)$  is well-defined on the (non-empty) open interval  $\mathbf{G}(x, y, \mathbb{R})$ . We call  $\mathbf{H}$  the *dual* of  $\mathbf{G}$ . In the optimal transport case,  $\mathbf{H}(x, y, u) = -c(x, y) - u$ , and we see that  $\mathbf{G}(x, y, \mathbb{R}) = \mathbb{R}$  for all pairs  $(x, y)$ . Generally, however,  $\mathbf{G}(x, y, \mathbb{R})$  maybe not be  $\mathbb{R}$  for any pair  $(x, y)$ . In addition, we define the set

$$\mathfrak{h} := \{(x, y, u) : u \in U_{x,y}\}$$

with  $U_{x,y} := \mathbf{G}(x, y, V_{x,y})$ . As  $\mathfrak{g}$  is open, (G-Mono) and the continuity of  $D_v \mathbf{G}$  together imply that the set  $\mathfrak{h}$  is also open. Hence,  $\mathbf{H}$  is locally  $C^{2,\alpha}$  on  $\mathfrak{h}$  and

$$\text{(H-Mono)} \quad D_u \mathbf{H}(x, y, u) < 0.$$

We can see that the map

$$\text{(H-Twist)} \quad (x, u) \mapsto (D_y \mathbf{H}(x, y, u), \mathbf{H}(x, y, u)) \text{ is injective on } \{(x, u) : (x, y, u) \in \mathfrak{h}\},$$

the map

$$\text{(H* -Twist)} \quad y \mapsto -\frac{D_x \mathbf{H}(x, y, u)}{D_u \mathbf{H}(x, y, u)} \text{ is injective on } \{y : (x, y, u) \in \mathfrak{h}\},$$

and

$$\text{(H-Nondeg)} \quad \det \left( D_{yx} \mathbf{H} - D_{yu} \mathbf{H} \otimes \frac{D_x \mathbf{H}}{D_u \mathbf{H}} \right) \neq 0 \text{ on } \mathfrak{h}.$$

In particular, (G-Twist) and (H\* -Twist), (G\* -Twist) and (H-Twist), and (G-Nondeg) and (H-Nondeg) are respectively equivalent (see [11, Remark 9.5] and [15]).

Moreover, with  $\mathbf{H}$ , we can generate the map  $\mathbf{S}$  and look to solve the dual generated Jacobian equation

$$(2.1) \quad \det(\nabla_y [\mathbf{S}(y, \mathbf{v}(y), \nabla \mathbf{v}(y))]) = \frac{g(y)}{f(\mathbf{S}(y, \mathbf{v}(y), \nabla \mathbf{v}(y)))}.$$

We will often use the following first-order identities:

$$D_x \mathbf{H} = -\frac{D_x \mathbf{G}}{D_v \mathbf{G}}, \quad D_y \mathbf{H} = -\frac{D_y \mathbf{G}}{D_v \mathbf{G}}, \quad \text{and} \quad D_u \mathbf{H} = \frac{1}{D_v \mathbf{G}}$$

and the following second-order identities:

$$D_{xu} \mathbf{H} = -\frac{D_{xv} \mathbf{G} + D_v^2 \mathbf{G} D_x \mathbf{H}}{(D_v \mathbf{G})^2}, \quad D_{yu} \mathbf{H} = -\frac{D_{yv} \mathbf{G} + D_v^2 \mathbf{G} D_y \mathbf{H}}{(D_v \mathbf{G})^2}, \quad \text{and} \quad D_u^2 \mathbf{H} = -\frac{D_v^2 \mathbf{G}}{(D_v \mathbf{G})^3}.$$

Here, the derivatives of  $\mathbf{H}$  are taken at  $(x, y, u)$  and the derivatives of  $\mathbf{G}$  at  $(x, y, \mathbf{H}(x, y, u))$  provided, of course,  $u \in \mathbf{G}(x, y, \mathbb{R})$ . These identities are simple consequences of (G-Mono).

Now let  $\mathbf{E}$  be the  $n \times n$  matrix from (G-Nondeg):

$$(2.2) \quad \mathbf{E}(x, y, v) := \left[ D_{xy} \mathbf{G} - D_{xv} \mathbf{G} \otimes \frac{D_y \mathbf{G}}{D_v \mathbf{G}} \right] (x, y, v).$$

Notice that the Jacobian determinants of the maps in (G-Twist) and (G\*-Twist) are

$$|D_v \mathbf{G}(x, y, v)|^n |\det(\mathbf{E}(x, y, v))| \quad \text{and} \quad |D_v \mathbf{G}(x, y, v)|^{-n} |\det(\mathbf{E}(x, y, v))|$$

respectively.

**2.2. G-convexity.** Solutions to (1.1) – (1.3) are  $\mathbf{G}$ -convex functions. Let us recall the definition of  $\mathbf{G}$ -convexity and some related facts, definitions, and characteristics. We say that a function  $\mathbf{u} : X \rightarrow \mathbb{R}$  is  $\mathbf{G}$ -convex if for every  $x_0 \in X$ , there exists a *focus*  $(y_0, v_0) \in Y \times \mathbb{R}$  such that

$$(x_0, y_0, v_0) \in \mathfrak{g}$$

and

$$\mathbf{u}(x_0) = \mathbf{G}(x_0, y_0, v_0) \quad \text{and} \quad \mathbf{u}(x) \geq \mathbf{G}(x, y_0, v_0) \quad \forall x \in X.$$

Notice that if  $(y_0, v_0)$  and  $(y_0, v_1)$  are foci for a  $\mathbf{G}$ -convex function  $\mathbf{u}$  at the point  $x_0$ , then by (G-Mono),  $v_0 = v_1 = \mathbf{H}(x_0, y_0, \mathbf{u}(x_0))$ . So we can recast our definition and say that  $\mathbf{u} : X \rightarrow \mathbb{R}$  is  $\mathbf{G}$ -convex if for each  $x_0 \in X$ , there exists a point  $y_0 \in Y$  such that

$$(x_0, y_0, \mathbf{H}(x_0, y_0, \mathbf{u}(x_0))) \in \mathfrak{g}$$

and

$$\mathbf{u}(x_0) = \mathbf{G}(x_0, y_0, \mathbf{H}(x_0, y_0, \mathbf{u}(x_0))) \quad \text{and} \quad \mathbf{u}(x) \geq \mathbf{G}(x, y_0, \mathbf{H}(x_0, y_0, \mathbf{u}(x_0))) \quad \forall x \in X.$$

For a  $\mathbf{G}$ -convex function  $\mathbf{u} : X \rightarrow \mathbb{R}$ , we define its  $\mathbf{G}$ -subdifferential at  $x_0$  to be the (non-empty) set

$$(2.3) \quad \partial_{\mathbf{G}} \mathbf{u}(x_0) := \{y \in Y : \mathbf{u}(x) \geq \mathbf{G}(x, y, \mathbf{H}(x_0, y, \mathbf{u}(x_0))) \quad \forall x \in X\}$$

provided

$$(2.4) \quad (x_0, y, \mathbf{H}(x_0, y, \mathbf{u}(x_0))) \in \mathfrak{g}.$$

Given  $y \in \partial_{\mathbf{G}} \mathbf{u}(x_0)$ , we call

$$\mathcal{G}_{x_0, y, v}(\cdot) := \mathbf{G}(\cdot, y, v)$$

with  $v := \mathbf{H}(x_0, y, \mathbf{u}(x_0))$  a  $\mathbf{G}$ -support of  $\mathbf{u}$  at  $x_0$ . Recalling the *Fréchet subdifferential* of a function  $\mathbf{u}$  at  $x_0$ :

$$\partial^- \mathbf{u}(x_0) := \{p \in \mathbb{R}^n : \mathbf{u}(x) \geq \mathbf{u}(x_0) + p \cdot (x - x_0) + o(|x - x_0|)\},$$

with  $o(|x - x_0|) \rightarrow 0$  as  $x \rightarrow x_0$ , we see that

$$(2.5) \quad y \in \partial_{\mathbf{G}} \mathbf{u}(x_0) \quad \Rightarrow \quad D_x \mathbf{G}(x_0, y, \mathbf{H}(x_0, y, \mathbf{u}(x_0))) \in \partial^- \mathbf{u}(x_0).$$

For  $E \subset X$ , we set

$$\partial_{\mathbf{G}} \mathbf{u}(E) := \bigcup_{x \in E} \partial_{\mathbf{G}} \mathbf{u}(x) \quad \text{and} \quad \partial^- \mathbf{u}(E) := \bigcup_{x \in E} \partial^- \mathbf{u}(x).$$

**Remark 2.1.** In the optimal transport setting, the geometric condition (2.3) alone dictates whether or not a point  $y$  is in the  $\mathbf{G}$ -subdifferential of  $\mathbf{u}$  at  $x_0$ . The admissibility condition (2.4) always holds. Yet this is not the case in general. A simple but important consequence of this is that the  $\mathbf{G}$ -subdifferential may not be continuous in the way the  $c$ -subdifferential is for  $c$ -convex functions. For instance, in the quadratic cost case,  $c$ -convexity is convexity, and given  $y_k \in \partial^- \mathbf{u}(x_k)$  such that  $y_k \rightarrow y_0$  and  $x_k \rightarrow x_0$ , we know that  $y_0 \in \partial^- \mathbf{u}(x_0)$ . However, if we replace  $\partial^- \mathbf{u}$  with  $\partial_{\mathbf{G}} \mathbf{u}$ , this implication may not hold.

Akin to the Legendre transform, we define the  $\mathbf{G}$ -transform of  $\mathbf{u}$  to be the  $\mathbf{H}$ -convex function given by

$$(2.6) \quad \mathbf{u}_{\mathbf{G}}(y) := \sup_{x \in X} \mathbf{H}(x, y, \mathbf{u}(x)).$$

In actuality, the supremum here is taken over those  $x \in X$  such that  $\mathbf{H}(x, y, \mathbf{u}(x))$  is defined; the  $\mathbf{G}$ -convexity of  $\mathbf{u}$  implies that  $\mathbf{u}(x) \in U_{x,y}$  whenever  $y \in \partial_{\mathbf{G}}\mathbf{u}(x)$ , and so the supremum is over a non-empty set. Moreover, as noted in [15, Section 4]<sup>2</sup>,

$$(2.7) \quad \mathbf{u}_{\mathbf{GH}} = \mathbf{u}$$

where  $\mathbf{v}_{\mathbf{H}}(x) := \sup_{y \in Y} \mathbf{G}(x, y, \mathbf{v}(y))$  is the  $\mathbf{H}$ -transform of a given  $\mathbf{H}$ -convex function  $\mathbf{v} : Y \rightarrow \mathbb{R}$ , and

$$(2.8) \quad y \in \partial_{\mathbf{G}}\mathbf{u}(x) \iff x \in \partial_{\mathbf{H}}\mathbf{u}_{\mathbf{G}}(y).$$

For the  $\mathbf{H}$ -subdifferential, the analogue of the admissibility condition (2.4) is

$$(x, y_0, \mathbf{G}(x, y_0, \mathbf{v}(y_0))) \in \mathfrak{h},$$

for  $x \in \partial_{\mathbf{H}}\mathbf{v}(y_0)$ . Because we have assumed that  $\mathbf{G}$  is of class  $C_{\text{loc}}^2(X \times Y \times \mathbb{R})$ , we find that  $\mathbf{G}$ -convex functions are locally semiconvex. (The semiconvexity constant of  $\mathbf{u}$  in a set depends only on the  $C^0$ -norm of  $D_x^2\mathbf{G}$  in that set.) In particular,  $\mathbf{G}$ -convex functions are locally uniformly Lipschitz and twice differentiable at almost every point (see, e.g., [7]). This basic regularity will be the foundation of our analysis.

Since  $\mathbf{G}$  satisfies (G-Twist) and (G-Nondeg), we can generate the maps  $\mathbf{G}\text{-exp}_{x,u}(\cdot)$  and  $\mathbf{V}_x(\cdot, \cdot)$  from the pair of equations

$$\begin{cases} D_x\mathbf{G}(x, \mathbf{G}\text{-exp}_{x,u}(p), \mathbf{V}_x(u, p)) = p \\ \mathbf{G}(x, \mathbf{G}\text{-exp}_{x,u}(p), \mathbf{V}_x(u, p)) = u \end{cases} \quad \forall (p, u) \in (D_x\mathbf{G}, \mathbf{G})(\{(x, y, v) : (x, y, v) \in \mathfrak{g}\}).$$

Here,  $D_x\mathbf{G}$  is evaluated at the point  $(x, \mathbf{G}\text{-exp}_{x,u}(p), \mathbf{V}_x(u, p))$ . In other words,

$$\mathbf{G}\text{-exp}_{x,u}(p) = y \iff p = D_x\mathbf{G}(x, y, \mathbf{H}(x, y, u))$$

and

$$\mathbf{V}_x(u, p) = \mathbf{H}(x, \mathbf{G}\text{-exp}_{x,u}(p), u)$$

so long as  $(x, y, \mathbf{H}(x, y, u)) \in \mathfrak{g}$ . And so (2.5) can be rewritten as

$$(2.9) \quad \partial_{\mathbf{G}}\mathbf{u}(x_0) \subset \mathbf{G}\text{-exp}_{x_0, \mathbf{u}(x_0)}(\partial^-\mathbf{u}(x_0)).$$

When applying  $\mathbf{G}\text{-exp}_{x_0, \mathbf{u}(x_0)}(\cdot)$  to  $p \in \partial^-\mathbf{u}(x_0)$ , we only consider those  $p = D_x\mathbf{G}(x_0, y, \mathbf{H}(x_0, y, \mathbf{u}(x_0)))$  such that  $\mathbf{u}(x_0) \in U_{x_0, y}$ . Hence, we see that if  $\mathbf{u}$  is differentiable at  $x_0$ , then  $\partial_{\mathbf{G}}\mathbf{u}(x_0)$  is a singleton  $\{y_0\}$  and

$$(2.10) \quad \nabla\mathbf{u}(x_0) = D_x\mathbf{G}(x_0, y_0, v_0),$$

and if  $\mathbf{u}$  is twice differentiable at  $x_0$ , then

$$(2.11) \quad D^2\mathbf{u}(x_0) \geq D_x^2\mathbf{G}(x_0, y_0, v_0).$$

In (2.10) and (2.11),  $v_0 := \mathbf{H}(x_0, y_0, \mathbf{u}(x_0))$ . Finally, given a  $\mathbf{G}$ -convex function  $\mathbf{u} : X \rightarrow \mathbb{R}$ , let us define the map (at almost every  $x \in X$ )  $\mathbf{T}_{\mathbf{u}}$  by

$$\mathbf{T}_{\mathbf{u}}(x) := \mathbf{G}\text{-exp}_{x, \mathbf{u}(x)}(\nabla\mathbf{u}(x)).$$

Even though  $\mathbf{T}_{\mathbf{u}}$  depends on  $\mathbf{u}$  and  $\mathbf{G}$ , we shall often suppress the second dependence for notational simplicity.

<sup>2</sup> While (2.7) and (2.8) are mentioned in [15, Section 4], they are not proved. For completeness, we prove them here. Let  $y_0 \in \partial_{\mathbf{G}}\mathbf{u}(x_0)$  and  $v_0 := \mathbf{H}(x_0, y_0, \mathbf{u}(x_0))$ . By definition,  $(x_0, y_0, \mathbf{u}(x_0)) \in \mathfrak{h}$ . If  $x_0$  is the only point at which  $\mathbf{H}(\cdot, y_0, \mathbf{u}(\cdot))$  is defined, then  $x_0 \in \partial_{\mathbf{H}}\mathbf{u}_{\mathbf{G}}(y_0)$  trivially. On the other hand, let  $x \in X$  be such that  $\mathbf{H}(x, y_0, \mathbf{u}(x))$  is well-defined. Since  $\mathbf{u}(x) \geq \mathbf{G}(x, y_0, v_0)$  for all  $x \in X$ , we see that

$$\mathbf{H}(x, y_0, \mathbf{u}(x)) \leq \mathbf{H}(x, y_0, \mathbf{G}(x, y_0, v_0)) = v_0 = \mathbf{H}(x_0, y_0, \mathbf{u}(x_0)).$$

By construction,  $\mathbf{H}(x, y_0, \cdot)$  can be evaluated at  $\mathbf{G}(x, y_0, v_0)$ . Hence, the supremum in (2.6) is achieved at  $(x_0, y_0, \mathbf{u}(x_0))$ . That is, if  $y_0 \in \partial_{\mathbf{G}}\mathbf{u}(x_0)$ , then  $x_0 \in \partial_{\mathbf{H}}\mathbf{u}_{\mathbf{G}}(y_0)$  and  $\mathbf{u}_{\mathbf{G}}(y_0) = \mathbf{H}(x_0, y_0, \mathbf{u}(x_0))$ . It then follows that

$$\mathbf{u}_{\mathbf{GH}}(x_0) \geq \mathbf{G}(x_0, y_0, \mathbf{u}_{\mathbf{G}}(y_0)) = \mathbf{G}(x_0, y_0, \mathbf{H}(x_0, y_0, \mathbf{u}(x_0))) = \mathbf{u}(x_0).$$

A symmetric argument, taking  $\mathbf{u}_{\mathbf{G}}$  in place of  $\mathbf{u}$  and  $\mathbf{u}_{\mathbf{GH}}$  in place of  $\mathbf{u}_{\mathbf{G}}$  above, yields that

$$x_0 \in \partial_{\mathbf{H}}\mathbf{u}_{\mathbf{G}}(y_0) \implies y_0 \in \partial_{\mathbf{G}}\mathbf{u}_{\mathbf{GH}}(x_0)$$

and  $\mathbf{u}(x_0) = \mathbf{G}(x_0, y_0, \mathbf{H}(x_0, y_0, \mathbf{u}(x_0))) \geq \mathbf{G}(x_0, y_0, \mathbf{u}_{\mathbf{G}}(y_0)) = \mathbf{u}_{\mathbf{GH}}(x_0)$ .

**2.3. A Monge–Ampère-type Equation.** Set  $\mathbf{v} := \mathbf{u}_{\mathbf{G}}$ . Then,  $\mathbf{v}$  is a solution to the SBVP for (2.1), the dual equation (see [15, Lemma 4.1]). In particular, by [15, Lemma 4.1] and the remarks just before it,  $(\mathbf{S}_{\mathbf{v}})_{\#}g = f$ . Recall that  $\mathbf{u}$  and  $\mathbf{v}$  are twice differentiable almost everywhere; let  $\Omega_1$  and  $\Upsilon_1$  be the full (Lebesgue) measure subsets of  $\Omega$  and  $\Upsilon$  respectively on which  $\mathbf{u}$  and  $\mathbf{v}$  are respectively twice differentiable. By (2.8) and (2.9), we see  $\mathbf{T}_{\mathbf{u}}$  and  $\mathbf{S}_{\mathbf{v}}$  are inverses of one another in the sense that

$$(2.12) \quad \mathbf{S}_{\mathbf{v}}(\mathbf{T}_{\mathbf{u}}(x)) = x \quad \forall x \in \Omega_1 \quad \text{and} \quad \mathbf{T}_{\mathbf{u}}(\mathbf{S}_{\mathbf{v}}(y)) = y \quad \forall y \in \Upsilon_1.$$

Here, of course,  $\mathbf{S}_{\mathbf{v}}(y) := \mathbf{H}\text{-exp}_{y, \mathbf{v}(y)}(\nabla \mathbf{v}(y))$ . Since  $(\mathbf{T}_{\mathbf{u}})_{\#}f = g$ , we can apply [16, Theorem 11.1] to deduce that

$$|\det(\nabla \mathbf{T}_{\mathbf{u}}(x))| = \frac{f(x)}{g(\mathbf{T}_{\mathbf{u}}(x))} \quad \forall x \in \Omega_1.$$

Then, (2.10) and (2.11) imply that

$$(2.13) \quad \begin{aligned} & \det(D^2 \mathbf{u}(x) - D_x^2 \mathbf{G}(x, \mathbf{T}_{\mathbf{u}}(x), \mathbf{H}(x, \mathbf{T}_{\mathbf{u}}(x), \mathbf{u}(x))) \\ &= |\det(\mathbf{E}(x, \mathbf{T}_{\mathbf{u}}(x), \mathbf{H}(x, \mathbf{T}_{\mathbf{u}}(x), \mathbf{u}(x)))| \frac{f(x)}{g(\mathbf{T}_{\mathbf{u}}(x))} \quad \text{a.e.} \end{aligned}$$

Since by (G-Nondeg),  $\mathbf{E}$  has non-zero determinant, the nondegeneracy of the right-hand side of our Monge–Ampère-type equation (2.13) is preserved. (See [15] for more details.) In conclusion, a solution  $\mathbf{u}$  to (1.1) – (1.3) satisfies a Monge–Ampère-type equation almost everywhere.

### 3. PROOF OF THEOREM 1.1

Set  $\mathbf{v}$  to be the  $\mathbf{G}$ -transform of  $\mathbf{u}$  and let  $\Omega_1$ ,  $\Upsilon_1$ , and  $\mathbf{S}_{\mathbf{v}}$  be as in (2.12). Consider the set

$$\Omega_2 := \Omega_1 \cap \mathbf{T}_{\mathbf{u}}^{-1}(\Upsilon_1) \subset \Omega,$$

and observe that  $|\Omega \setminus \Omega_2| = 0$  since  $(\mathbf{T}_{\mathbf{u}})_{\#}f = g$  and the densities  $f$  and  $g$  are bounded away from zero and infinity. Recall that  $(\mathbf{S}_{\mathbf{v}})_{\#}g = f$ .

Fix  $x' \in \Omega_2$ . Since  $x'$  is a point of differentiability for  $\mathbf{u}$ , the  $\mathbf{G}$ -subdifferential of  $\mathbf{u}$  at  $x'$  is a singleton (see (2.9)):  $\partial_{\mathbf{G}} \mathbf{u}(x') = \{\mathbf{T}_{\mathbf{u}}(x')\}$ . Set  $y' := \mathbf{T}_{\mathbf{u}}(x')$  and  $v' := \mathbf{H}(x', \mathbf{T}_{\mathbf{u}}(x'), \mathbf{u}(x'))$ . Note that  $y' \in \Upsilon_1$ . Up to a translation, we can assume that  $(x', y', v') = (0, 0, 0)$ . Furthermore, up to subtracting  $\mathbf{G}(\cdot, 0, 0)$ , we can assume that  $\mathbf{u}(0) = 0$  and its  $\mathbf{G}$ -support at  $(0, 0, 0)$  is identically zero; that is,  $\mathcal{G}_{0,0,0}(x) = 0$  and  $\mathbf{u}(x) \geq 0$  for all  $x \in \Omega$ . In turn,

$$(3.1) \quad D_x \mathbf{G}(\cdot, 0, 0) = 0.$$

With these normalizations in hand, define

$$\hat{\mathbf{G}}(x, y, v) := \mathbf{G}(x, y, v + \mathbf{H}(0, y, 0)) \quad \text{and} \quad \hat{\mathbf{u}}(x) := \mathbf{u}(x).$$

Notice that  $\mathbf{H}(0, y, 0)$  may not be defined at all points  $y \in Y$  or even all points  $y \in \Upsilon$ . However, it is well-defined in  $B_\varepsilon \subset \Upsilon$  for some  $\varepsilon > 0$  by the implicit function theorem and (G-Mono). Furthermore, as  $\mathbf{u}$  is twice differentiable at the origin and  $\nabla \mathbf{u}(0) = 0$ , [16, Theorem 14.25] implies that

$$\partial^- \mathbf{u}(x) = D^2 \mathbf{u}(0)x + o(|x|).$$

So using (2.9), we can find an  $\varepsilon > 0$  such that

$$(3.2) \quad \partial_{\mathbf{G}} \mathbf{u}(B_\varepsilon) \subset B_\varepsilon \subset \Upsilon.$$

Here,  $\varepsilon$  and  $\epsilon$  are not necessarily equal. For each  $(x, y) \in B_\epsilon \times B_\epsilon$ , let  $\hat{V}_{x,y} := V_{x,y} - \mathbf{H}(0, y, 0)$ . Define

$$\hat{\mathbf{g}} := \{(x, y, v) : (x, y) \in B_\epsilon \times B_\epsilon \text{ and } v \in \hat{V}_{x,y}\}.$$

Observe that  $\hat{\mathbf{G}}$  satisfies (G-Twist), (G\*-Twist), and (G-Nondeg) on  $\hat{\mathbf{g}}$  (also (G-Mono)). In  $B_\epsilon$ , we find that  $\hat{\mathbf{u}}$  is  $\hat{\mathbf{G}}$ -convex. Indeed, let  $x_0 \in B_\epsilon$ ,  $y_0 \in \partial_{\mathbf{G}} \mathbf{u}(x_0)$ , and  $\hat{v}_0 := \mathbf{H}(x_0, y_0, \mathbf{u}(x_0)) - \mathbf{H}(0, y_0, 0)$ . Then,  $(x_0, y_0, \hat{v}_0) \in \hat{\mathbf{g}}$  and

$$\hat{\mathbf{G}}(x, y_0, \hat{v}_0) = \mathbf{G}(x, y_0, \mathbf{H}(x_0, y_0, \mathbf{u}(x_0))) \leq \mathbf{u}(x) = \hat{\mathbf{u}}(x) \quad \forall x \in B_\epsilon$$

with equality at  $x = x_0$ . In particular,

$$(3.3) \quad \partial_{\hat{\mathbf{G}}} \hat{\mathbf{u}}(x) = \partial_{\mathbf{G}} \mathbf{u}(x) \quad \forall x \in B_\epsilon.$$

Setting

$$(3.4) \quad \hat{f} := f\mathbf{1}_{B_\epsilon} \quad \text{and} \quad \hat{g} := g\mathbf{1}_{\partial_{\mathbf{G}}\hat{\mathbf{u}}(B_\epsilon)},$$

we claim that

$$(3.5) \quad (\mathbf{T}_{\hat{\mathbf{u}}})_{\#}\hat{f} = \hat{g}.$$

Note that the dual of  $\hat{\mathbf{G}}$  is

$$\hat{\mathbf{H}}(x, y, u) := \mathbf{H}(x, y, u) - \mathbf{H}(0, y, 0).$$

And so using (3.3) and recalling (2.9), we see that  $\mathbf{T}_{\mathbf{u}}|_{B_\epsilon} = \mathbf{T}_{\hat{\mathbf{u}}}$ . Thus, recalling (3.3), it suffices to show that  $\mathbf{T}_{\mathbf{u}}^{-1}(\partial_{\mathbf{G}}\mathbf{u}(B_\epsilon)) \setminus B_\epsilon$  has measure zero. To this end, observe that if  $x \in \mathbf{T}_{\mathbf{u}}^{-1}(\partial_{\mathbf{G}}\mathbf{u}(B_\epsilon)) \setminus B_\epsilon$ , then there exists an  $x_\epsilon \in B_\epsilon$  such that  $\partial_{\mathbf{G}}\mathbf{u}(x) \cap \partial_{\mathbf{G}}\mathbf{u}(x_\epsilon)$  is non-empty. Therefore, as  $x \neq x_\epsilon$  and recalling (2.8), we see that

$$\mathbf{T}_{\mathbf{u}}^{-1}(\partial_{\mathbf{G}}\mathbf{u}(B_\epsilon)) \setminus B_\epsilon \subset \mathbf{T}_{\mathbf{u}}^{-1}(\{\text{non-differentiability points of } \mathbf{v}\}).$$

Since

$$|\mathbf{T}_{\mathbf{u}}^{-1}(\{\text{non-differentiability points of } \mathbf{v}\})| = 0,$$

because  $\mathbf{v}$  is semiconvex and  $f$  is bounded away from zero, the claim holds.

By construction,  $\hat{\mathbf{G}}$  is such that

$$\hat{\mathbf{G}}(\cdot, 0, 0) = \hat{\mathbf{G}}(0, \cdot, 0) \equiv 0,$$

and  $\hat{\mathbf{u}}$  is such that

$$\hat{\mathbf{u}}(0) = \nabla\mathbf{u}(0) = 0.$$

Thus, Taylor expanding  $\hat{\mathbf{G}}$  and  $\hat{\mathbf{u}}$  around the origin yields

$$\begin{aligned} \hat{\mathbf{G}}(x, y, v) &= D_v\hat{\mathbf{G}}(0, 0, 0)v + D_{xy}\hat{\mathbf{G}}(0, 0, 0)x \cdot y \\ &\quad + D_{xv}\hat{\mathbf{G}}(0, 0, 0) \cdot vx + D_{yv}\hat{\mathbf{G}}(0, 0, 0) \cdot vy + \frac{1}{2}D_v^2\hat{\mathbf{G}}(0, 0, 0)v^2 + O(|x|^{2+\alpha} + |y|^{2+\alpha} + |v|^{2+\alpha}) \end{aligned}$$

and

$$\hat{\mathbf{u}}(x) = \frac{1}{2}D^2\hat{\mathbf{u}}(0)x \cdot x + o(|x|^2).$$

Let

$$a := -D_v\hat{\mathbf{G}}(0, 0, 0) > 0, \quad M := D_{xy}\hat{\mathbf{G}}(0, 0, 0), \quad \text{and} \quad P := D^2\hat{\mathbf{u}}(0).$$

As  $M = \mathbf{E}(0, 0, 0)$ , with  $\mathbf{E}$  defined in (2.2),  $\det(M) \neq 0$ . Hence, using (2.13) and (3.1), we find that  $\det(P) = \det(D^2\mathbf{u}(0)) > 0$ ; that is,  $P$  is positive definite and symmetric. Therefore, after the change of coordinates

$$(x, y, v) \mapsto (\tilde{x}, \tilde{y}, \tilde{v}) := (P^{1/2}x, P^{-1/2}M^t y, av),$$

we see that

$$(3.6) \quad \tilde{\mathbf{G}}(\tilde{x}, \tilde{y}, \tilde{v}) := \hat{\mathbf{G}}(x, y, v) = -\tilde{v} + \tilde{x} \cdot \tilde{y} + b_1 \cdot \tilde{v}\tilde{x} + b_2 \cdot \tilde{v}\tilde{y} + c_3\tilde{v}^2 + O(|\tilde{x}|^{2+\alpha} + |\tilde{y}|^{2+\alpha} + |\tilde{v}|^{2+\alpha}),$$

with

$$b_1 := \frac{1}{a}D_{xv}\hat{\mathbf{G}}(0, 0, 0)P^{-1/2}, \quad b_2 := \frac{1}{a}D_{yv}\hat{\mathbf{G}}(0, 0, 0)[M^t]^{-1}P^{1/2}, \quad \text{and} \quad c_3 := \frac{1}{2a^2}D_v^2\hat{\mathbf{G}}(0, 0, 0).$$

Also,

$$(3.7) \quad \tilde{\mathbf{u}}(\tilde{x}) := \hat{\mathbf{u}}(x) = \frac{1}{2}|\tilde{x}|^2 + o(|\tilde{x}|^2)$$

and is  $\tilde{\mathbf{G}}$ -convex in  $P^{1/2}B_\epsilon$ . In particular,

$$\partial_{\tilde{\mathbf{G}}}\tilde{\mathbf{u}}(\tilde{x}) = P^{-1/2}M^t\partial_{\mathbf{G}}\hat{\mathbf{u}}(x) \quad \text{with} \quad x = P^{-1/2}\tilde{x}.$$

Now admissibility is with respect to

$$\tilde{\mathbf{g}} := \{(\tilde{x}, \tilde{y}, \tilde{v}) : (\tilde{x}, \tilde{y}) \in P^{1/2}B_\epsilon \times P^{-1/2}M^tB_\epsilon \text{ and } \tilde{v} \in a\hat{V}_{P^{-1/2}\tilde{x}, [M^t]^{-1}P^{1/2}\tilde{y}}\}.$$

Additionally, letting  $\tilde{\mathbf{H}}$  be the dual of  $\tilde{\mathbf{G}}$ , we see that

$$(3.8) \quad \tilde{\mathbf{H}}(\tilde{x}, \tilde{y}, \tilde{u}) = -\tilde{u} + \tilde{x} \cdot \tilde{y} - b_1 \cdot \tilde{u}\tilde{x} - b_2 \cdot \tilde{u}\tilde{y} - c_3\tilde{u}^2 + O(|\tilde{x}|^{2+\alpha} + |\tilde{y}|^{2+\alpha} + |\tilde{u}|^{2+\alpha}).$$

By construction,

$$\tilde{\mathbf{H}}(\cdot, 0, 0) = \tilde{\mathbf{H}}(0, \cdot, 0) \equiv 0.$$

Here,  $\tilde{u} := u$ . Furthermore, if we set

$$\tilde{f}(\tilde{x}) := \det(P^{-1/2})\hat{f}(P^{-1/2}\tilde{x}) \quad \text{and} \quad \tilde{g}(\tilde{y}) := |\det([M^t]^{-1}P^{1/2})|\hat{g}([M^t]^{-1}P^{1/2}\tilde{y}),$$

then from (3.4), (2.13), and (3.1), we deduce that

$$(3.9) \quad \frac{\tilde{f}(0)}{\tilde{g}(0)} = \frac{\det(P^{-1/2})}{|\det([M^t]^{-1}P^{1/2})|} \frac{f(0)}{g(0)} = \frac{\det(P^{-1})}{|\det(M^{-1})|} \frac{\det(D^2\mathbf{u}(0))}{|\det(\mathbf{E}(0,0,0))|} = 1.$$

Moreover,  $(\mathbf{T}_{\tilde{\mathbf{u}}})_{\#}\tilde{f} = \tilde{g}$  by (3.5) and construction.

Consider the rescalings<sup>3</sup>

$$\mathbf{G}_r(\tilde{x}, \tilde{y}, \tilde{v}) := \frac{\tilde{\mathbf{G}}(r\tilde{x}, r\tilde{y}, r^2\tilde{v})}{r^2}, \quad \mathbf{H}_r(\tilde{x}, \tilde{y}, \tilde{v}) := \frac{\tilde{\mathbf{H}}(r\tilde{x}, r\tilde{y}, r^2\tilde{v})}{r^2}, \quad \text{and} \quad \mathbf{u}_r(\tilde{x}) := \frac{\tilde{\mathbf{u}}(r\tilde{x})}{r^2}.$$

Since  $\mathfrak{g}$  and  $\mathfrak{h}$  are open and  $\mathbf{G}$  is of class  $C_{\text{loc}}^2$ , we have that  $B_8 \times B_8 \times (-64, 64) \subset \mathfrak{g}_r, \mathfrak{h}_r$  for all  $r$  sufficiently small. In addition to the openness of  $\mathfrak{g}$ , using (3.6) and the  $C^2$  regularity of  $\mathbf{G}$ , we can ensure that  $B_8 \times (-64, 64) \times B_8 \subset \text{dom } \mathbf{G}_r\text{-exp}$  for all  $r$  small enough.<sup>4</sup> Here,  $\mathfrak{g}_r$  is the set on which  $\mathbf{G}_r$  satisfies (G-Twist), (G\*-Twist), and (G-Nondeg); and  $\mathbf{H}_r$  and  $\mathfrak{h}_r$  are as expected. Also, define

$$f_r(\tilde{x}) := \frac{\tilde{f}(r\tilde{x})}{\tilde{f}(0)} \quad \text{and} \quad g_r(\tilde{y}) := \frac{\tilde{g}(r\tilde{y})}{\tilde{g}(0)}.$$

Since  $f$  and  $g$  are bounded away from zero and infinity, (3.9) implies that  $\tilde{f}(0) = \tilde{g}(0)$ . Therefore, from the continuity of  $f$  and  $g$ , we deduce that

$$\|f_r - 1\|_{L^\infty(B_4)} + \|g_r - 1\|_{L^\infty(B_4)} \leq \delta$$

with  $\delta = \delta(r) \rightarrow 0$  as  $r \rightarrow 0$ . Using the push-forward condition  $(\mathbf{T}_{\tilde{\mathbf{u}}})_{\#}\tilde{f} = \tilde{g}$ , we find that  $(\mathbf{T}_{\mathbf{u}_r})_{\#}f_r = g_r$ . Moreover, from (3.6) and (3.8), we determine that

$$(3.10) \quad \|\mathbf{G}_r - \tilde{x} \cdot \tilde{y} + \tilde{v}\|_{C^{2,\alpha}(B_8 \times B_8 \times (-64, 64))} + \|\mathbf{H}_r - \tilde{x} \cdot \tilde{y} + \tilde{u}\|_{C^{2,\alpha}(B_8 \times B_8 \times (-64, 64))} \leq \delta$$

where  $\delta = \delta(r) \rightarrow 0$  as  $r \rightarrow 0$ . Furthermore, from (3.7), we see that

$$\left\| \mathbf{u}_r - \frac{1}{2}|\tilde{x}|^2 \right\|_{C^0(B_4)} \leq \eta$$

where  $\eta = \eta(r) \rightarrow 0$  as  $r \rightarrow 0$ .

As we proved (3.2), we deduce that

$$(3.11) \quad \partial^- \mathbf{u}_r(\tilde{x}) \subset B_\varrho(\tilde{x}) \quad \forall \tilde{x} \in B_3$$

where  $\varrho = \varrho(r) \rightarrow 0$  as  $r \rightarrow 0$ . Additionally, from (3.10), we find that

$$\|\mathbf{G}_r\text{-exp}_{\tilde{x}, \tilde{u}}(\tilde{p}) - \tilde{p}\|_{C^1(B_8 \times (-64, 64) \times B_8)} \leq \delta.$$

Therefore, using (2.9) and (3.11), we obtain that

$$(3.12) \quad \partial_{\mathbf{G}_r} \mathbf{u}_r(\tilde{x}) \subset B_\rho(\tilde{x}) \quad \forall \tilde{x} \in B_2$$

where  $\rho = \rho(r) \rightarrow 0$  as  $r \rightarrow 0$ . Since  $\hat{\mathbf{u}}_{\tilde{\mathbf{G}}}$  is twice differentiable at 0, the  $\mathbf{G}_r$ -transform of  $\mathbf{u}_r$  is also twice differentiable at 0; let  $\mathbf{v}_r$  be the  $\mathbf{G}_r$ -transform of  $\mathbf{u}_r$ . In addition,  $D^2\mathbf{v}_r(0) = \text{Id}^5$ . So arguing as we did to

<sup>3</sup> The most basic generating function, coming from the optimal transport problem with quadratic cost, is invariant under parabolically quadratic rescalings, thinking of  $v$  as time: set

$$\mathbf{G}_r(x, y, v) := \frac{\mathbf{G}(rx, ry, r^2v)}{r^2};$$

if  $\mathbf{G}(x, y, v) = x \cdot y - v$ , then  $\mathbf{G}_r = \mathbf{G}$ . Rescaling in this way suggests that  $\hat{\mathbf{G}}$  need not account for  $D_{xv}\mathbf{G}(0,0,0)$ ,  $D_{yv}\mathbf{G}(0,0,0)$ , or  $D_v^2\mathbf{G}(0,0,0)$  being non-zero, a heuristic confirmation of our choice for  $\hat{\mathbf{G}}$ .

<sup>4</sup> By the  $C^2$  regularity of  $\mathbf{G}$ , the openness of  $\mathfrak{g}$ , and the inverse function theorem (recall (G-Twist)), we find that the family functions  $\{\mathbf{F}_x := (\mathbf{G}\text{-exp}_x, \mathbf{V}_x)\}_{x \in \mathcal{O}}$  varies in a  $C^1$  fashion (in  $x$ ) in some open set  $\mathcal{O}$ . Hence, by continuity and the openness of  $\mathfrak{g}$ , we can find an open subset of the origin in  $\mathbb{R}^{2n+1}$  on which  $\mathbf{G}\text{-exp}$  is well-defined. The rescaling  $(x, u, p) \mapsto (x/r, u/r^2, p/r)$ , therefore, permits this inclusion.

<sup>5</sup> We can see this by differentiating the equations

$$\nabla \mathbf{u}_r(\tilde{x}) = D_{\tilde{x}}\mathbf{G}_r(\tilde{x}, \mathbf{T}_{\mathbf{u}_r}(\tilde{x}), \mathbf{H}_r(\tilde{x}, \mathbf{T}_{\mathbf{u}_r}(\tilde{x}), \mathbf{u}_r(\tilde{x}))) \quad \text{and} \quad \nabla \mathbf{v}_r(\tilde{y}) = D_{\tilde{y}}\mathbf{H}_r(\mathbf{S}_{\mathbf{v}_r}(\tilde{y}), \tilde{y}, \mathbf{G}_r(\mathbf{S}_{\mathbf{v}_r}(\tilde{y}), \tilde{y}, \mathbf{v}_r(\tilde{y})))$$

at 0 and using (3.6), (3.8), and that  $[\nabla \mathbf{T}_{\mathbf{u}_r}(0)]^{-1} = \nabla \mathbf{S}_{\mathbf{v}_r}(0)$ .



prove (3.12), we find that

$$(3.13) \quad \partial_{\mathbf{H}_r \mathbf{v}_r}(\tilde{y}) \subset B_{\rho^*}(\tilde{y}) \quad \forall \tilde{y} \in B_2$$

where  $\rho^* = \rho^*(r) \rightarrow 0$  as  $r \rightarrow 0$ . Hence, if

$$\mathcal{C} := \overline{B_1} \quad \text{and} \quad \mathcal{K} := \partial_{\mathbf{G}_r \mathbf{u}_r}(\overline{B_1})$$

and  $r$  is sufficiently small, then we can force

$$(3.14) \quad B_{1/2} \subset \mathcal{K} \subset B_2$$

thanks to (3.12), (3.13), and duality. Observe that  $\mathcal{C}$  is convex by construction and  $\mathcal{K}$  is closed being the  $\mathbf{G}_r$ -subdifferential of a compact set and recalling (3.14), (3.10), and the inclusion  $B_8 \times B_8 \times (-64, 64) \subset \mathfrak{g}_r$ . Finally, recalling (3.9) and arguing as we did to prove (3.5), we have that

$$(\mathbf{T}_{\mathbf{u}_r})_{\#}(f_r \mathbf{1}_{\mathcal{C}}) = g_r \mathbf{1}_{\mathcal{K}}.$$

The remainder of the proof of Theorem 1.1 is identical to the optimal transport case after replacing [5, Theorem 4.3], [5, Theorem 5.3], and [5, Corollary 4.6] by Theorem 4.1, Theorem 5.5, and Corollary 4.7 respectively. We refer the reader to [5] rather than including the details.

#### 4. $C^{1,\beta}$ -REGULARITY AND STRICT $\mathbf{G}$ -CONVEXITY

In this section, we prove an  $\varepsilon$ -regularity result and exhibit some of its consequences. Before stating it, let us introduce some notation. Set

$$\mathcal{B}_R := B_{2R} \times B_{2R} \times (-R^3, R^3) \subset \mathbb{R}^{2n+1}.$$

Furthermore, we say a set  $E$  is  $C$ -semiconvex if every point on the boundary of  $E$  can be touched from outside by a ball of radius  $1/C$ .

**Theorem 4.1.** *Let  $\mathcal{C}$  be a closed,  $(C_2\delta)$ -semiconvex set and  $\mathcal{K}$  be a closed set such that*

$$(4.1) \quad B_{1/2} \subset \mathcal{C}, \mathcal{K} \subset B_2,$$

*$f$  and  $g$  be two densities supported on  $\mathcal{C}$  and  $\mathcal{K}$  respectively, and  $\mathbf{u}$  be a  $\mathbf{G}$ -convex function such that  $\partial_{\mathbf{G}\mathbf{u}}(\mathcal{C}) \subset B_2$  and  $(\mathbf{T}_{\mathbf{u}})_{\#}f = g$ . In addition, suppose that  $\mathcal{B}_4 \subset \mathfrak{g}, \mathfrak{h}$ . For every  $\beta \in (0, 1)$ , there exist constants  $\delta_0, \eta_0 \in (0, 1)$  such that the follow holds: if*

$$(4.2) \quad \|f - \mathbf{1}_{\mathcal{C}}\|_{L^\infty(B_4)} + \|g - \mathbf{1}_{\mathcal{K}}\|_{L^\infty(B_4)} \leq \delta_0,$$

$$(4.3) \quad \|\mathbf{G} - x \cdot y + v\|_{C^{2,\alpha}(B_4)} + \|\mathbf{H} - x \cdot y + u\|_{C^{2,\alpha}(B_4)} \leq \delta_0,$$

and

$$(4.4) \quad \left\| \mathbf{u} - \frac{1}{2}|x|^2 \right\|_{C^0(B_4)} \leq \eta_0,$$

then  $\mathbf{u} \in C^{1,\beta}(B_{1/6})$ .

Theorem 4.1 will follow from its pointwise version Proposition 4.2.

**Proposition 4.2.** *Let  $\mathcal{C}$  be a closed,  $(C_3\delta)$ -semiconvex set and  $\mathcal{K}$  be a closed set such that*

$$B_{1/3} \subset \mathcal{C}, \mathcal{K} \subset B_3,$$

*$f$  and  $g$  be two densities supported on  $\mathcal{C}$  and  $\mathcal{K}$  respectively, and  $\mathbf{u}$  be a  $\mathbf{G}$ -convex function such that  $\mathbf{u}(0) = 0$ ,  $\partial_{\mathbf{G}\mathbf{u}}(\mathcal{C}) \subset B_3$ , and  $(\mathbf{T}_{\mathbf{u}})_{\#}f = g$ . In addition, suppose that  $\mathcal{B}_3 \subset \mathfrak{g}, \mathfrak{h}$ ,*

$$(4.5) \quad \mathbf{G}(\cdot, 0, 0) = \mathbf{G}(0, \cdot, 0) = \mathbf{H}(\cdot, 0, 0) = \mathbf{H}(0, \cdot, 0) \equiv 0,$$

and

$$(4.6) \quad D_v \mathbf{G}(0, 0, 0) = D_u \mathbf{H}(0, 0, 0) = -1 \quad \text{and} \quad D_{xy} \mathbf{G}(0, 0, 0) = D_{xy} \mathbf{H}(0, 0, 0) = \text{Id}.$$

For every  $\beta \in (0, 1)$ , there exist constants  $\delta, \eta \in (0, 1)$  such that the follow holds: if

$$(4.7) \quad \|f - 1\|_{L^\infty(\mathcal{C})} + \|g - 1\|_{L^\infty(\mathcal{K})} \leq \delta,$$

$$(4.8) \quad \|\mathbf{G} - x \cdot y + v\|_{C^{2,\alpha}(B_3)} + \|\mathbf{H} - x \cdot y + u\|_{C^{2,\alpha}(B_3)} \leq \delta,$$

and

$$(4.9) \quad \left\| \mathbf{u} - \frac{1}{2}|x|^2 \right\|_{C^0(B_3)} \leq \eta,$$

then  $\mathbf{u} \in C^{1,\beta}(0)$ .

The proof of Proposition 4.2 makes use of two lemmas. The first is a compactness result that allows us to approximate  $\mathbf{u}$  with a solution to an optimal transport problem with quadratic cost. The second is an estimate on the  $\mathbf{G}$ -subdifferential of  $\mathbf{u}$  in terms of the gradient map of the convex potential that approximates  $\mathbf{u}$  found in the first lemma.

**Lemma 4.3.** *Let  $\mathcal{C}$  be a closed,  $(C_R\delta)$ -semiconvex set and  $\mathcal{K}$  be a closed set such that*

$$(4.10) \quad B_{1/R} \subset \mathcal{C}, \mathcal{K} \subset B_R$$

for some  $R \geq 3$ ,  $f$  and  $g$  be two densities supported on  $\mathcal{C}$  and  $\mathcal{K}$  respectively, and  $\mathbf{u} : B_R \rightarrow (-R^2, R^2)$  be a  $\mathbf{G}$ -convex function such that  $\partial_{\mathbf{G}}\mathbf{u}(\mathcal{C}) \subset B_R$  and  $(\mathbf{T}_{\mathbf{u}})_{\#}f = g$ . In addition, suppose that  $\mathcal{B}_R \times \subset \mathfrak{g}$ . Also, let  $\rho > 0$  be such that  $|\mathcal{C}| = |\rho\mathcal{K}|$  and  $\mathbf{w}$  be a convex function such that  $(\nabla\mathbf{w})_{\#}\mathbf{1}_{\mathcal{C}} = \mathbf{1}_{\rho\mathcal{K}}$  with  $\mathbf{w}(0) = \mathbf{u}(0)$ . Then, there exists an increasing function  $\omega : \mathbb{R}^+ \rightarrow \mathbb{R}^+$ , depending only on  $R$ , satisfying  $\omega(\delta) \geq \delta$  and  $\omega(0^+) = 0$  such that if

$$(4.11) \quad \|f - 1\|_{L^\infty(\mathcal{C})} + \|g - 1\|_{L^\infty(\mathcal{K})} \leq \delta$$

and

$$(4.12) \quad \|\mathbf{G} - x \cdot y + v\|_{C^2(\mathcal{B}_R)} + \|\mathbf{H} - x \cdot y + u\|_{C^2(\mathcal{B}_R)} \leq \delta,$$

then

$$(4.13) \quad \|\mathbf{u} - \mathbf{w}\|_{C^0(B_{1/R})} \leq \omega(\delta).$$

*Proof.* Suppose, to the contrary, that the lemma is false. Then, there exists an  $\varepsilon_0 > 0$  and sequences of closed sets  $\mathcal{C}_j$  and  $\mathcal{K}_j$  satisfying (4.10) with  $\mathcal{C}_j$  being  $(C_R/j)$ -semiconvex, functions  $f_j$  and  $g_j$  satisfying (4.11) with  $\delta = 1/j$ , and generating functions  $\mathbf{G}_j$  satisfying (4.12) also with  $\delta = 1/j$  such that

$$(4.14) \quad \mathbf{u}_j(0) = \mathbf{w}_j(0) = 0 \quad \text{and} \quad \|\mathbf{u}_j - \mathbf{w}_j\|_{C^0(B_{1/R})} \geq \varepsilon_0$$

where  $\mathbf{u}_j$  and  $\mathbf{w}_j$  are as in the statement of the lemma. Moreover,  $\mathcal{B}_R \subset \mathfrak{g}_j$ .

Using the push-forward condition  $(\mathbf{T}_{\mathbf{u}_j})_{\#}f_j = g_j$ , (4.10), and (4.11), we find that

$$(4.15) \quad \rho_j = (|\mathcal{C}_j|/|\mathcal{K}_j|)^{1/n} \rightarrow 1$$

as  $j \rightarrow \infty$ . Now let us extend each  $\mathbf{w}_j|_{\mathcal{C}_j}$  to a convex function on  $B_R$ , setting

$$\mathbf{w}_j(x) := \sup_{z \in \mathcal{C}_j, p \in \partial^- \mathbf{w}_j(z)} \mathbf{w}_j(z) + p \cdot (x - z).$$

From (4.15), we see that

$$\partial^- \mathbf{w}_j(B_R) \subset B_{\rho_j R} \subset B_{2R}$$

for  $j \gg 1$ . Hence, the family  $\mathbf{w}_j$  is uniformly Lipschitz (recall the equality in (4.14)), and so, up to a subsequence,  $\mathbf{w}_j$  converges uniformly in  $B_R$  to some convex function  $\mathbf{w}_\infty$ . Similarly, let us extend  $\mathbf{u}_j|_{\mathcal{C}_j}$  to  $B_R$ :

$$\mathbf{u}_j(x) := \sup_{z \in \mathcal{C}_j, y \in \partial_{\mathbf{G}_j} \mathbf{u}_j(z)} \mathbf{G}_j(x, y, \mathbf{H}_j(z, y, \mathbf{u}_j(z))).$$

Given  $x_0 \in B_R$ , let  $(z_0, y_0)$  be a pair at which the above supremum is attained. Then, using (4.12), we see that  $v_0 := \mathbf{H}_j(z_0, y_0, \mathbf{u}_j(z_0)) \in (-R^3, R^3)$ , and it follows that  $(x_0, y_0, v_0) \in \mathfrak{g}_j$ . Consequently, these extensions are  $\mathbf{G}_j$ -convex in  $B_R$ . So from (4.12), in particular, since the  $C^1$ -norms of  $\mathbf{G}_j$  are uniformly bounded, and as  $\partial_{\mathbf{G}_j} \mathbf{u}_j(B_R) \subset B_R$ , taking  $j \gg 1$ , we determine that the collection  $\mathbf{u}_j$  is uniformly  $(R+1)$ -Lipschitz (again, recall that  $\mathbf{u}_j(0) = 0$ ). Thus, up to a subsequence,  $\mathbf{u}_j$  converges uniformly in  $B_R$  to some convex function  $\mathbf{u}_\infty$ . Moreover, by (4.14),

$$(4.16) \quad \mathbf{u}_\infty(0) = \mathbf{w}_\infty(0) = 0 \quad \text{and} \quad \|\mathbf{u}_\infty - \mathbf{w}_\infty\|_{C^0(B_{1/R})} \geq \varepsilon_0.$$

Up to subsequences, the sets  $\mathcal{C}_j$  converge in the Hausdorff sense to some

$$B_{1/R} \subset \mathcal{C}_\infty \subset B_R.$$

Also, since each  $\mathcal{C}_j$  is  $(C_R/j)$ -semiconvex, it follows that  $\mathcal{C}_\infty$  is convex and

$$\text{dist}(\partial\mathcal{C}_j, \partial\mathcal{C}_\infty) + \text{dist}(\partial\mathcal{C}_\infty, \partial\mathcal{C}_j) \rightarrow 0.$$

Then, arguing exactly as is [1, Theorem], we find that  $|\mathcal{C}_j\Delta\mathcal{C}_\infty| \rightarrow 0$ , or, equivalently, that  $\mathbf{1}_{\mathcal{C}_j}$  converges in  $L^1$  to  $\mathbf{1}_{\mathcal{C}_\infty}$ . Furthermore, using (4.11) and as  $\rho_j \rightarrow 1$ , we see that  $f_j$  and  $\mathbf{1}_{\rho_j\mathcal{C}_j}$  converge in  $L^1$  to  $\mathbf{1}_{\mathcal{C}_\infty}$ . In addition, up to subsequences,  $g_j$  and  $\mathbf{1}_{\rho_j\mathcal{K}_j}$  converge weakly-\* in  $L^\infty$  to a density  $g_\infty$ .

By [16, Theorem 5.20] and the uniqueness of optimal transports, we see that  $\nabla\mathbf{w}_\infty$  is the optimal transport for the quadratic cost  $-x \cdot y$  taking  $\mathbf{1}_{\mathcal{C}_\infty}$  to  $g_\infty$ . If  $\nabla\mathbf{u}_\infty = \nabla\mathbf{w}_\infty$  almost everywhere in  $\mathcal{C}_\infty$ , then the equality in (4.16) implies that  $\mathbf{u}_\infty = \mathbf{w}_\infty$ , contradicting the inequality in (4.16). It then follows that there exists an increasing function  $\omega_R : \mathbb{R}^+ \rightarrow \mathbb{R}^+$ , depending only on  $R$ , such that  $\omega_R(0^+) = 0$  and (4.13) holds. Taking  $\omega(\delta) := \max\{\omega_R(\delta), \delta\}$  concludes the proof.

Define  $\pi_j := (\text{Id}, \mathbf{T}_{\mathbf{u}_j})\#f_j$ . By construction, this family of measures is tight and

$$\text{spt } \pi_j \subset \bigcup_{x \in \mathcal{C}_j} \{(x, y) : y \in \partial_{\mathbf{G}_j} \mathbf{u}_j(x)\}.$$

So  $\pi_j$  converges weakly to some measure  $\pi_\infty$  whose marginals are  $\mathbf{1}_{\mathcal{C}_\infty}$  and  $g_\infty$ . Furthermore, for any  $\{(x_k, y_k)\}_{k=1}^N \subset \text{spt } \pi_\infty$ , there exist sequences  $\{(x_k^j, y_k^j)\}_{k=1}^N \subset \text{spt } \pi_j$  such that  $(x_k^j, y_k^j) \rightarrow (x_k, y_k)$  for each  $k = 1, \dots, N$  and

$$\sum_{k=1}^N \mathbf{u}_j(x_{k+1}^j) \geq \sum \mathbf{G}_j(x_{k+1}^j, y_k^j, \mathbf{H}_j(x_k^j, y_k^j, \mathbf{u}_j(x_k^j)))$$

with  $(x_{N+1}^j, y_{N+1}^j) = (x_1^j, y_1^j)$ . This is just the  $\mathbf{G}_j$ -convexity of  $\mathbf{u}_j$ . Recalling (4.12) and that  $\mathbf{u}_j$  converges uniformly to  $\mathbf{u}_\infty$ , taking the limit as  $j \rightarrow \infty$ , we deduce that

$$0 \geq \sum_{k=1}^N (x_{k+1} - x_k) \cdot y_k.$$

In other words, the support of  $\pi_\infty$  is  $c$ -cyclically monotone for the quadratic cost. Therefore,

$$\pi_\infty = (\text{Id}, \nabla\mathbf{u}_0)\#\mathbf{1}_{\mathcal{C}_\infty},$$

and  $\nabla\mathbf{u}_0$  is the optimal transport for the quadratic cost  $-x \cdot y$  taking  $\mathbf{1}_{\mathcal{C}_\infty}$  to  $g_\infty$ . In particular,  $\nabla\mathbf{u}_0 = \nabla\mathbf{w}_\infty$  almost everywhere in  $\mathcal{C}_\infty$ . Using that  $f_j$  converges in  $L^1$  to  $\mathbf{1}_{\mathcal{C}_\infty}$  and arguing as in the proof of [16, Corollary 5.23], we see that  $\mathbf{T}_{\mathbf{u}_j}$  converges to  $\nabla\mathbf{u}_0$  in measure in  $B_R$  and as distributions, up to a further subsequence. By (4.12) and since  $\mathbf{u}_j$  converges to  $\mathbf{u}_\infty$  uniformly, we have that  $\mathbf{T}_{\mathbf{u}_j}$  also converges to  $\nabla\mathbf{u}_\infty$  as distributions in  $B_R$ . Hence, by the local integrability of  $\nabla\mathbf{u}_0$  and  $\nabla\mathbf{u}_\infty$ , we determine that  $\nabla\mathbf{u}_\infty = \nabla\mathbf{u}_0$  almost everywhere in  $\mathcal{C}_\infty$ , as desired.  $\square$

**Remark 4.4** (A remark on the regularity of  $\mathcal{C}$  and the proof of Lemma 4.3). In the optimal transportation setting, De Philippis and Figalli appeal to the strong stability results available for solutions, thanks to the Kantorovich formulation of the problem, to show that  $\nabla\mathbf{u}_\infty = \nabla\mathbf{w}_\infty$  almost everywhere. Here, however, the set  $\mathcal{C}$  has to have some regularity to deduce  $L^1$  convergence of the contradiction sequence's source densities and, in turn, prove the same equality. An inspection of the proof of [1, Theorem] shows that  $|\mathcal{C}_j\Delta\mathcal{C}_\infty|$  goes to zero, i.e.,  $\mathbf{1}_{\mathcal{C}_j}$  converges in  $L^1$  to  $\mathbf{1}_{\mathcal{C}_\infty}$ , provided that the boundaries  $\partial\mathcal{C}_j$  converge uniformly to the boundary  $\partial\mathcal{C}_\infty$  and have ( $n$ -dimensional Lebesgue) measure zero. Therefore, Lemma 4.3 can be applied, by Arzelà-Ascoli, if  $\mathcal{C}$  is a Lipschitz set whose boundary's Lipschitz constant depends only on  $R$ , for example. So the regularity assumption on  $\mathcal{C}$  in Theorem 4.1 and the following lemmas, propositions, and theorems, that  $\mathcal{C}$  is  $(C\delta)$ -semiconvex, can be weakened. Indeed, in the course of the proof of Proposition 4.2, every application of Lemma 4.3 after the first will be to the  $(C_3\delta)$ -semiconvex sets  $\{\mathbf{u}_k \leq 1\}$ . With respect to our main theorem,  $\mathcal{C} = \overline{B}_1 - x_0$ , which is as nice as imaginable.

From this point forward, let  $\mathcal{N}_r(E)$  denote the  $r$ -neighborhood of a set  $E$ .

**Lemma 4.5.** *Let  $R \geq 3$ ,  $\mathbf{u} : B_{1/R} \rightarrow (-R, R)$  be a  $\mathbf{G}$ -convex function such that  $\partial_{\mathbf{G}}\mathbf{u}(B_{1/R}) \subset B_R$ , and  $\mathbf{w} \in C^1(B_{1/R})$  be convex. Suppose that  $\mathcal{B}_R \subset \mathbf{g}$ . Fix  $A \in \mathbb{R}^{n \times n}$  to be a symmetric matrix such that*

$$(4.17) \quad \frac{1}{K} \text{Id} \leq A \leq K \text{Id}$$

for some  $K \geq 1$ . Define the ellipsoid

$$\mathcal{E}(x_0, h) := \left\{ x : \frac{1}{2}A(x - x_0) \cdot (x - x_0) \leq h \right\},$$

and assume that  $\mathcal{E}(x_0, h) \subset B_{1/R}$ . If

$$\|\mathbf{u} - \mathbf{w}\|_{C^0(\mathcal{E}(x_0, h))} \leq \varepsilon$$

and

$$\|\mathbf{G} - x \cdot y + v\|_{C^2(\mathcal{B}_R)} + \|\mathbf{H} - x \cdot y + u\|_{C^2(\mathcal{B}_R)} \leq \delta$$

for small constants  $\varepsilon, \delta > 0$ , then

$$\partial_{\mathbf{G}}\mathbf{u}(\mathcal{E}(x_0, h - \varepsilon^{1/2})) \subset \mathcal{N}_{\delta + K'(h\varepsilon)^{1/2}}(\nabla\mathbf{w}(\mathcal{E}(x_0, h))) \quad \forall 0 < \varepsilon < h^2 \ll 1$$

where  $K' = K'(K) > 0$ .

*Proof.* Up to a change of coordinates, we can assume that  $x_0 = 0$ . Let  $\mathcal{E}(h) = \mathcal{E}(0, h)$  and define

$$\bar{\mathbf{w}}(x) := \mathbf{w}(x) + \varepsilon + \varepsilon^{1/2}(Ax \cdot x - 2h).$$

By construction,  $\bar{\mathbf{w}} \geq \mathbf{u}$  outside  $\mathcal{E}(h)$  and  $\bar{\mathbf{w}} \leq \mathbf{u}$  inside  $\mathcal{E}(h - \varepsilon^{1/2})$ . Therefore, if  $\mathcal{G}_{x,y,v}$  is a  $\mathbf{G}$ -support for  $\mathbf{u}$  at  $x \in \mathcal{E}(h - \varepsilon^{1/2})$ , then  $\mathcal{G}_{x,y,\bar{v}}$  will touch  $\bar{\mathbf{w}}$  from below at a point  $\bar{x} \in \mathcal{E}(h)$  for some  $\bar{v} \geq v$ . Moreover,  $(\bar{x}, y, \bar{v}) \in \mathfrak{g}$  since  $\bar{v} = \mathbf{H}(\bar{x}, y, \bar{\mathbf{w}}(\bar{x})) \in (-R^2, R^2)$ . Hence,

$$(4.18) \quad \partial_{\mathbf{G}}\mathbf{u}(\mathcal{E}(h - \varepsilon^{1/2})) \subset \partial_{\mathbf{G}}\bar{\mathbf{w}}(\mathcal{E}(h)).$$

Note that even though  $\bar{\mathbf{w}}$  may not be  $\mathbf{G}$ -convex, we can still consider its  $\mathbf{G}$ -subdifferential; it just might be empty at some points. In particular, the equality  $\partial_{\mathbf{G}}\bar{\mathbf{w}}(x) = \mathbf{G}\text{-exp}_{x,\bar{\mathbf{w}}(x)}(\nabla\bar{\mathbf{w}}(x))$  still holds. Thus, since  $|\mathbf{G}\text{-exp}_{x,\bar{\mathbf{w}}(x)}(\nabla\bar{\mathbf{w}}(x)) - \nabla\bar{\mathbf{w}}(x)| \leq \delta$  by assumption, we find that

$$(4.19) \quad \partial_{\mathbf{G}}\bar{\mathbf{w}}(\mathcal{E}(h)) \subset \mathcal{N}_{\delta}(\nabla\bar{\mathbf{w}}(\mathcal{E}(h))).$$

From (4.17), we determine that  $\mathcal{E}(h) \subset B_{(2Kh)^{1/2}}$ . And so since

$$|\nabla\bar{\mathbf{w}}(x)| \leq |\nabla\mathbf{w}(x)| + 2\varepsilon^{1/2}K|x|,$$

recalling (4.17), it follows that

$$(4.20) \quad \mathcal{N}_{\delta}(\nabla\bar{\mathbf{w}}(\mathcal{E}(h))) \subset \mathcal{N}_{\delta + 4K(Kh\varepsilon)^{1/2}}(\nabla\mathbf{w}(\mathcal{E}(h))).$$

Finally, combining (4.18), (4.19), and (4.20), we deduce that

$$\partial_{\mathbf{G}}\mathbf{u}(\mathcal{E}(h - \varepsilon^{1/2})) \subset \mathcal{N}_{\delta + K'(h\varepsilon)^{1/2}}(\nabla\mathbf{w}(\mathcal{E}(h)))$$

with  $K' = 4K^{3/2}$ , as desired.  $\square$

*Proof of Proposition 4.2.* The proof will be done in four steps.

– *Step 1:  $\mathbf{u}$  and its  $\mathbf{G}$ -sections are close to a strictly convex solution of a Monge–Ampère equation and its sections.*

Using Lemma 4.3 and arguing exactly as in [5], we find the existence of a strictly convex function  $\mathbf{w}$  such that  $\mathbf{w}(0) = \mathbf{u}(0) = 0$ ,

$$(4.21) \quad \|\mathbf{u} - \mathbf{w}\|_{C^0(B_{1/3})} \leq \omega(\delta),$$

and

$$(4.22) \quad \det(D^2\mathbf{w}) = 1 \quad \text{in } B_{1/4}$$

in the Alexandrov sense. Furthermore, there exists a constant  $K_0 = K_0(n) > 0$  such that

$$(4.23) \quad \|\mathbf{w}\|_{C^3(B_{1/5})} \leq K_0 \quad \text{and} \quad \frac{1}{K_0} \text{Id} \leq D^2\mathbf{w} \leq K_0 \text{Id} \quad \text{in } B_{1/5}.$$

And so

$$(4.24) \quad S(\mathbf{w}, h) := \{x : \mathbf{w}(x) \leq \nabla\mathbf{w}(0) \cdot x + h\} \subset B_{(2K_0h)^{1/2}}.$$

(Precisely, this is Step 1 in the proof of [5, Theorem 4.3], which uses (4.9).) By (4.23) and as  $\mathbf{u}$  is semiconvex with a semiconvexity constant depending only on  $\|D_x^2 \mathbf{G}\|_{C^0(\mathcal{B}_3)}$ , i.e.,  $\delta$  ((4.8)), we have that  $\mathbf{u} - \mathbf{w}$  is semiconvex with a semiconvexity constant depending on dimension (recall that  $\delta \ll 1$ ). So using (4.21), we deduce that

$$(4.25) \quad |\nabla \mathbf{w}(0)| \leq K_1 \omega(\delta)^{1/2}$$

for some constant  $K_1 = K_1(n) > 0$  (cf. (4.39), noticing that  $-\nabla \mathbf{w}(0) \in \partial^-(\mathbf{u} - \mathbf{w} + c|\cdot|^2)(0)$  for some  $c > 0$  depending on  $n$ ).

Define

$$S_{\mathbf{G}}(\mathbf{u}, h) := \{x : \mathbf{u}(x) \leq h\}.$$

We claim that if  $\delta$  and  $h$  are sufficiently small, then

$$(4.26) \quad S(\mathbf{w}, h - K_2 \omega(\delta)^{1/2}) \subset S_{\mathbf{G}}(\mathbf{u}, h) \subset S(\mathbf{w}, h + K_2 \omega(\delta)^{1/2}) \Subset B_{1/6}$$

where  $K_2 = K_2(n) > 0$ . First, by (4.24), we can choose  $\delta$  and  $h$  sufficiently small so that the last inclusion holds. To conclude, let  $x \in S(\mathbf{w}, h - K_2 \omega(\delta)^{1/2})$ . Then, by (4.21), recalling that  $\mathbf{u}(0) = 0$ , and from (4.25), we deduce that

$$\begin{aligned} \mathbf{u}(x) &\leq \nabla \mathbf{w}(0) \cdot x + h - K_2 \omega(\delta)^{1/2} + \omega(\delta) \\ &\leq h + K_1 \omega(\delta)^{1/2} - K_2 \omega(\delta)^{1/2} + \omega(\delta) \\ &\leq h \end{aligned}$$

taking  $K_2 = K_1 + 1$ . This proves the first inclusion; the proof of the second is analogous.

– *Step 2: The  $\mathbf{G}$ -sections of  $\mathbf{u}$  and their images under  $\partial_{\mathbf{G}} \mathbf{u}$  are close to ellipsoids with controlled eccentricity and  $\mathbf{u}$  is close to a paraboloid at some small scale.*

We claim that for every small  $\eta > 0$ , there exist constants  $h_0 = h_0(\eta, n) > 0$  and  $\delta = \delta(h_0, \eta, n) > 0$  such that the following holds: there exists a symmetric matrix satisfying

$$(4.27) \quad \frac{1}{K_3} \text{Id} \leq A \leq K_3 \text{Id},$$

$$(4.28) \quad \det(A) = 1,$$

$$(4.29) \quad AB_{h_0^{1/2}/3} \subset S_{\mathbf{G}}(\mathbf{u}, h_0) \subset AB_{3h_0^{1/2}},$$

and

$$(4.30) \quad A^{-1}B_{h_0^{1/2}/3} \subset \partial_{\mathbf{G}} \mathbf{u}(S_{\mathbf{G}}(\mathbf{u}, h_0)) \subset A^{-1}B_{3h_0^{1/2}}.$$

Moreover,

$$(4.31) \quad \left\| \mathbf{u} - \frac{1}{2}|A^{-1}x|^2 \right\|_{C^0(AB_{3h_0^{1/2}})} \leq \eta h_0.$$

Here,  $K_3 = K_3(n) > 0$ .

Let

$$A := [D^2 \mathbf{w}(0)]^{-1/2}.$$

With  $A$  defined in this way, using (4.23) and (4.22), we see that (4.27) and (4.28) hold.

Notice that (4.29) is equivalent to

$$\mathcal{E}(h_0/18) \subset S_{\mathbf{G}}(\mathbf{u}, h_0) \subset \mathcal{E}(9h_0/2)$$

where

$$\mathcal{E}(h) := \left\{ x : \frac{1}{2} D^2 \mathbf{w}(0) x \cdot x \leq h \right\}.$$

Now from (4.23), we deduce

$$(4.32) \quad \mathcal{E}(h) \subset B_{(2K_0 h)^{1/2}}.$$

Consequently,

$$\mathcal{E}(h) \subset S(\mathbf{w}, h + K_0(2K_0 h)^{3/2}) \quad \text{and} \quad S(\mathbf{w}, h) \subset \mathcal{E}(h + K_0(2K_0 h)^{3/2}).$$

The second inclusion follows from (4.24). Thus, applying (4.26), we see that

$$\mathcal{E}(h_0/18) \subset S_{\mathbf{G}}(\mathbf{u}, h_0)$$

provided that  $h_0$  and  $\delta$  are sufficiently small depending only on  $n$ . On the other hand, applying (4.26), we see that

$$S_{\mathbf{G}}(\mathbf{u}, h_0) \subset \mathcal{E}(9h_0/2)$$

so long as  $\delta$  and  $h_0$  are sufficiently small, again, depending only on  $n$ . Whence, (4.29) holds, as desired. More generally, for every  $c < 1$  and  $C > 1$ , we can find  $\delta$  and  $h_0$  sufficiently small so that

$$(4.33) \quad \mathcal{E}(ch_0) \subset S_{\mathbf{G}}(\mathbf{u}, h_0) \subset \mathcal{E}(Ch_0).$$

Let us now prove (4.30). To do this, we consider the  $\mathbf{G}$ -transform of  $\mathbf{u}$  and the Legendre transform of  $\mathbf{w}$ . Specifically,

$$\mathbf{v}(y) := \sup_{x \in B_{1/5}} \mathbf{H}(x, y, \mathbf{u}(x)) \quad \text{and} \quad \mathbf{w}^*(y) := \sup_{x \in B_{1/5}} \{x \cdot y - \mathbf{w}(x)\}.$$

Notice that by (4.8) and (4.21),

$$\|\mathbf{v} - \mathbf{w}^*\|_{C^0(B_{1/3})} \leq \omega(\delta) + \delta \leq 2\omega(\delta).$$

Also, observe that

$$(4.34) \quad \nabla \mathbf{w}^* = [\nabla \mathbf{w}]^{-1} \quad \text{and} \quad D^2 \mathbf{w}^*(\nabla \mathbf{w}(x)) = [D^2 \mathbf{w}(x)]^{-1}.$$

Moreover, from (4.23),  $\mathbf{w}^*$  is uniformly convex and of class  $C^3$  in the open set  $\nabla \mathbf{w}(B_{1/5})$ . Let

$$\mathcal{E}^*(h) := \left\{ y : \frac{1}{2} D^2 \mathbf{w}^*(\nabla \mathbf{w}(0))(y - \nabla \mathbf{w}(0)) \cdot (y - \nabla \mathbf{w}(0)) \leq h \right\}.$$

Thanks to (4.23), for every  $c < 1$  and  $C > 1$ , we can find  $h_0$  sufficiently small so that

$$(4.35) \quad \nabla \mathbf{w}(\mathcal{E}(ch_0)) \subset \mathcal{E}^*(h_0) \subset \nabla \mathbf{w}(\mathcal{E}(Ch_0)).$$

Indeed, using (4.34) and (4.23), we find that

$$|D^2 \mathbf{w}^*(\nabla \mathbf{w}(0))(\nabla \mathbf{w}(x) - \nabla \mathbf{w}(0)) \cdot (\nabla \mathbf{w}(x) - \nabla \mathbf{w}(0)) - D^2 \mathbf{w}(0)x \cdot x| \leq 2K_0|x|^3 + K_0^3|x|^4$$

and that  $|x| \leq K_0(2K_0h)^{1/2}$  whenever  $\nabla \mathbf{w}(x) \in \mathcal{E}^*(h)$ . Combining these last two inequalities and (4.32) proves (4.35). Then, by (4.33) with  $C = 3/2$ , Lemma 4.5 applied with  $h = 3h_0/2 + \omega(\delta)^{1/2}$ , and (4.35), we determine that

$$\partial_{\mathbf{G}} \mathbf{u}(S_{\mathbf{G}}(\mathbf{u}, h_0)) \subset \mathcal{N}_{K_0' \omega(\delta)^{1/2}}(\nabla \mathbf{w}(\mathcal{E}(2h_0))) \subset \mathcal{E}^*(7h_0/2)$$

for  $\delta$  and  $h_0$  sufficiently small, depending only on dimension. Hence, recalling (4.25) and choosing  $\delta$  sufficiently small, we find that the second inclusion in (4.30) holds. In order to conclude, we must show that

$$\mathcal{E}^*(h_0/16) \subset \partial_{\mathbf{G}} \mathbf{u}(S_{\mathbf{G}}(\mathbf{u}, h_0)).$$

The first inclusion in (4.30) then follows, again, by (4.25) and choosing  $\delta$  sufficiently small. Since

$$E \subset \partial_{\mathbf{G}} \mathbf{u}(\partial_{\mathbf{H}} \mathbf{v}(E)) \quad \forall E,$$

considering (4.33), we see it suffices to show that

$$\partial_{\mathbf{H}} \mathbf{v}(\mathcal{E}^*(h_0/16)) \subset \mathcal{E}(h_0/3).$$

Applying Lemma 4.5 to  $\mathbf{v}$  and  $\mathbf{w}^*$ , provided that  $\delta$  is small enough, we determine that

$$\partial_{\mathbf{H}} \mathbf{v}(\mathcal{E}^*(h_0/16)) \subset \mathcal{N}_{K_0'' \omega(\delta)^{1/2}}(\nabla \mathbf{w}^*(\mathcal{E}^*(h_0/8))) \subset \mathcal{E}(h_0/3).$$

Here, we have used (4.35) and (4.34) for the second inclusion. Thus, (4.30) indeed holds after taking  $\delta$  and  $h_0$  sufficiently small.

Finally, from (4.21), (4.23), (4.25), and (4.27), we see that

$$\begin{aligned} \left\| \mathbf{u} - \frac{1}{2}|A^{-1}x|^2 \right\|_{C^0(AB_{3h_0^{1/2}})} &\leq \|\mathbf{u} - \mathbf{w}\|_{C^0(AB_{3h_0^{1/2}})} + \left\| \mathbf{w} - \frac{1}{2}|A^{-1}x|^2 \right\|_{C^0(AB_{3h_0^{1/2}})} \\ &\leq \omega(\delta) + \|\nabla \mathbf{w}(0) \cdot x\|_{C^0(AB_{3h_0^{1/2}})} + \|K_0|x|^3\|_{C^0(AB_{3h_0^{1/2}})} \\ &\leq \omega(\delta) + 3K_1K_3\omega(\delta)^{1/2}h_0^{1/2} + 27K_0K_3^3h_0^{3/2} \\ &\leq \eta h_0 \end{aligned}$$

where the last inequality follows after first choosing  $h_0$  sufficiently small and then choosing  $\delta$  even smaller.

– *Step 3: An iterative construction.*

Set

$$\tilde{x} := \frac{1}{h_0^{1/2}}A^{-1}x, \quad \tilde{y} := \frac{1}{h_0^{1/2}}Ay, \quad \tilde{v} := \frac{1}{h_0}v, \quad \text{and} \quad \tilde{u} := \frac{1}{h_0}u.$$

Define

$$\mathbf{G}_1(\tilde{x}, \tilde{y}, \tilde{v}) := \frac{\mathbf{G}(h_0^{1/2}A\tilde{x}, h_0^{1/2}A^{-1}\tilde{y}, h_0\tilde{v})}{h_0} \quad \text{and} \quad \mathbf{u}_1(\tilde{x}) := \frac{\mathbf{u}(h_0^{1/2}A\tilde{x})}{h_0}.$$

Similarly, let

$$\mathbf{H}_1(\tilde{x}, \tilde{y}, \tilde{u}) := \frac{\mathbf{H}(h_0^{1/2}A\tilde{x}, h_0^{1/2}A^{-1}\tilde{y}, h_0\tilde{u})}{h_0}.$$

Recalling (4.5), (4.6), (4.8), and that we are considering parabolically quadratic rescalings, we see that

$$\|\mathbf{G}_1 - \tilde{x} \cdot \tilde{y} + \tilde{v}\|_{C^{2,\alpha}(\mathcal{B}_3)} + \|\mathbf{H}_1 - \tilde{x} \cdot \tilde{y} + \tilde{u}\|_{C^{2,\alpha}(\mathcal{B}_3)} \leq \delta$$

provided  $h_0 = h_0(K_3, \alpha) \ll 1$ . In addition, thanks to (4.29) and (4.30), we have the inclusions

$$B_{1/3} \subset \mathcal{C}_1, \mathcal{K}_1 \subset \mathcal{B}_3$$

where, recalling that  $\mathbf{G}_1(\cdot, 0, 0) \equiv 0$ ,

$$\mathcal{C}_1 := S_{\mathbf{G}_1}(\mathbf{u}_1, 1) = \{\mathbf{u}_1 \leq 1\} \quad \text{and} \quad \mathcal{K}_1 := \partial_{\mathbf{G}_1} \mathbf{u}_1(S_{\mathbf{G}_1}(\mathbf{u}_1, 1)).$$

Arguing as in Proposition 5.3, we find that  $\mathcal{C}_1$  is a closed,  $(C_3\delta)$ -semiconvex set. Furthermore, rewriting (4.31) yields

$$\left\| \mathbf{u}_1 - \frac{1}{2}|\tilde{x}|^2 \right\|_{C^0(\mathcal{B}_3)} \leq \eta.$$

Now let

$$f_1(\tilde{x}) := f(h_0^{1/2}A\tilde{x})\mathbf{1}_{\mathcal{C}_1} \quad \text{and} \quad g_1(\tilde{y}) := g(h_0^{1/2}A^{-1}\tilde{y})\mathbf{1}_{\mathcal{K}_1}.$$

Recalling that  $\det(A) = 1$  and arguing as in the proof of Theorem 1.1, we determine that

$$(\mathbf{T}_{\mathbf{u}_1})_{\#} f_1 = g_1.$$

Finally, using (4.7), it follows that

$$\|f_1 - 1\|_{L^\infty(\mathcal{C}_1)} + \|g_1 - 1\|_{L^\infty(\mathcal{K}_1)} \leq \delta.$$

Hence, we can apply Steps 1 and 2 to  $\mathbf{u}_1$  and find a symmetric matrix  $A_1$  such that

$$\frac{1}{K_3} \text{Id} \leq A_1 \leq K_3 \text{Id},$$

$$\det(A_1) = 1,$$

$$A_1 B_{h_0^{1/2}/3} \subset S_{\mathbf{G}_1}(\mathbf{u}_1, h_0) \subset A_1 B_{3h_0^{1/2}},$$

$$A_1^{-1} B_{h_0^{1/2}/3} \subset \partial_{\mathbf{G}_1} \mathbf{u}(S_{\mathbf{G}_1}(\mathbf{u}_1, h_0)) \subset A_1^{-1} B_{3h_0^{1/2}},$$

and

$$\left\| \mathbf{u}_1 - \frac{1}{2}|A_1^{-1}\tilde{x}|^2 \right\|_{C^0(A_1 B_{3h_0^{1/2}})} \leq \eta h_0.$$

(The set  $S_{\mathbf{G}_1}(\mathbf{u}_1, h) = \{\mathbf{u}_1 \leq h\}$ .)

We can continue, iteratively constructing

$$\mathbf{G}_{k+1}(\tilde{x}, \tilde{y}, \tilde{v}) := \frac{\mathbf{G}_k(h_0^{1/2} A_k \tilde{x}, h_0^{1/2} A_k^{-1} \tilde{y}, h_0 \tilde{v})}{h_0} \quad \text{and} \quad \mathbf{u}_{k+1}(\tilde{x}) := \frac{\mathbf{u}_k(h_0^{1/2} A_k \tilde{x})}{h_0}$$

where  $A_k$  is the symmetric matrix constructed in the the  $k$ -th iteration. In turn, setting

$$M_k := A_1 \cdots A_k,$$

we have a sequence of symmetric matrices such that

$$(4.36) \quad \frac{1}{K_3^k} \text{Id} \leq M_k \leq K_3^k \text{Id},$$

$$\det(M_k) = 1,$$

and

$$(4.37) \quad M_k B_{h_0^{k/2}/3} \subset S_{\mathbf{G}_1}(\mathbf{u}_1, h_0^k) \subset M_k B_{3h_0^{k/2}}.$$

– *Step 4:  $C^{1,\beta}$ (0)-regularity.*

Let  $\beta \in (0, 1)$ . By (4.36) and (4.37), we find that

$$(4.38) \quad B_{(h_0^{1/2}/3K_3)^k} \subset S_{\mathbf{G}_1}(\mathbf{u}_1, h_0^k) \subset B_{(3K_3 h_0^{1/2})^k}.$$

Defining  $r_0 := h_0^{1/2}/3K_3$  and recalling that  $\mathbf{G}_1(\cdot, 0, 0) \equiv 0$ , it follows that

$$\|\mathbf{u}_1\|_{C^0(B_{r_0^k})} \leq h_0^k = (3K_3 r_0)^{2k} \leq r_0^{(1+\beta)k}$$

provided  $h_0$  (and so  $r_0$ ) is sufficiently small. In other words,  $\mathbf{u}_1$  and  $\mathbf{u}$  are  $C^{1,\beta}$  at the origin.  $\square$

With Proposition 4.2 in hand, let us now prove Theorem 4.1. The proof amounts to a change of variables.

*Proof of Theorem 4.1.* Let  $x_0 \in B_{1/6}$  and  $y_0 \in \partial_{\mathbf{G}} \mathbf{u}(x_0)$ . By (4.3), observe that

$$|x_0 - y_0| \leq |x_0 - p_0| + |p_0 - y_0| \leq |x_0 - p_0| + \delta_0$$

where

$$p_0 := D_x \mathbf{G}(x_0, y_0, \mathbf{H}(x_0, y_0, \mathbf{u}(x_0))).$$

As  $\mathbf{u}$  is semiconvex with a semiconvexity constant depending only on  $\|D_x^2 \mathbf{G}\|_{C^0(\mathcal{B}_4)} \leq \delta$  (recall (4.3)), there exists a  $c > 0$  such that

$$\mathbf{w}(x) := \mathbf{u}(x) - \frac{1}{2}|x|^2 + c|x - x_0|^2$$

is convex. By construction,  $p_0 - x_0 \in \partial^- \mathbf{w}(x_0)$ . And so using (4.4) and since  $x_0 + \eta_0^{1/2} \mathbf{e} \in \mathcal{C}$  provided that  $\eta_0 < 1/9$ , for example, we find that

$$(p_0 - x_0) \cdot \mathbf{e} \leq \frac{\mathbf{w}(x_0 + \eta_0^{1/2} \mathbf{e}) - \mathbf{w}(x_0)}{\eta_0^{1/2}} \leq (2 + c)\eta_0^{1/2} \quad \forall \mathbf{e} \in \mathbb{S}^{n-1}.$$

In turn,

$$(4.39) \quad |x_0 - y_0| \leq C(\delta_0 + \eta_0^{1/2}).$$

Set

$$u_0 := \mathbf{u}(x_0) \quad \text{and} \quad v_0 := \mathbf{H}(x_0, y_0, \mathbf{u}(x_0)).$$

From (4.4) and using (4.3) and (4.1), we deduce that

$$(4.40) \quad |u_0| \leq \frac{1}{72} + \eta_0 \quad \text{and} \quad |v_0| \leq \frac{1}{3} + \frac{1}{72} + \eta_0 + \delta_0.$$

Define

$$M_0 := \mathbf{E}(x_0, y_0, v_0) \quad \text{and} \quad a_0 := -D_v \mathbf{G}(x_0, y_0, v_0),$$

where  $\mathbf{E}$  is as defined in (2.2). Observe that (4.3) implies that

$$(4.41) \quad |a_0 - 1|, |a_0^{-1} - 1| \leq \delta_0 \quad \text{and} \quad |M_0 - \text{Id}|, |M_0^{-1} - \text{Id}| \leq 2\delta_0.$$



Now consider the change of variables

$$\bar{x} := x - x_0, \quad \bar{y} := M_0(y - y_0), \quad \bar{v} := a_0(v - v_0), \quad \text{and} \quad \bar{u} := u - u_0.$$

Define

$$\bar{\mathcal{C}} := \mathcal{C} - x_0 \quad \text{and} \quad \bar{\mathcal{K}} := M_0(\mathcal{K} - y_0);$$

and set

$$\bar{f}(\bar{x}) := f(\bar{x} + x_0) \quad \text{and} \quad \bar{g}(\bar{y}) := \det(M_0^{-1})g(M_0^{-1}\bar{y} + y_0).$$

Then, from (4.1), (4.39), and (4.41), we see that

$$B_{1/3} \subset \bar{\mathcal{C}}, \bar{\mathcal{K}} \subset B_3$$

if  $\delta_0$  and  $\eta_0$  are sufficiently small. From (4.41), we have that  $|\det(M_0) - 1| \leq (1 + 4n)\delta_0$  if  $\delta_0$  is sufficiently small. Thus,

$$\|\bar{f} - 1\|_{L^\infty(\bar{\mathcal{C}})} + \|\bar{g} - 1\|_{L^\infty(\bar{\mathcal{K}})} \leq 4(1 + n)\delta_0,$$

recalling (4.2). Let

$$\bar{\mathbf{G}}(\bar{x}, \bar{y}, \bar{v}) := \mathbf{G}(x, y, v - v_0 + \mathbf{H}(x_0, y, u_0)) - \mathbf{G}(x, y_0, v_0).$$

Notice that the dual of  $\bar{\mathbf{G}}$  is

$$\bar{\mathbf{H}}(\bar{x}, \bar{y}, \bar{u}) = a_0(\mathbf{H}(x, y, u - u_0 + \mathbf{G}(x, y_0, v_0)) - \mathbf{H}(x_0, y, u_0)),$$

and  $\bar{\mathbf{H}}$  is well-defined on  $\mathcal{B}_3$  by the assumption that  $\mathcal{B}_4 \subset \mathfrak{h}$  and our estimates on  $x_0, y_0, u_0, v_0, a_0$ , and  $M_0$ . Similarly, from (4.3), (4.39), (4.40), and (4.41), it follows that

$$\|\bar{\mathbf{G}} - \bar{x} \cdot \bar{y} + \bar{v}\|_{C^{2,\alpha}(\mathcal{B}_3)} + \|\bar{\mathbf{H}} - \bar{x} \cdot \bar{y} + \bar{u}\|_{C^{2,\alpha}(\mathcal{B}_3)} \leq C_0\delta_0.$$

In particular,

$$\bar{\mathbf{G}}(\cdot, 0, 0) = \bar{\mathbf{G}}(0, \cdot, 0) = \bar{\mathbf{H}}(\cdot, 0, 0) = \bar{\mathbf{H}}(0, \cdot, 0) \equiv 0.$$

Also, computations show that

$$D_{\bar{v}}\bar{\mathbf{G}}(0, 0, 0) = D_{\bar{u}}\bar{\mathbf{H}}(0, 0, 0) = -1 \quad \text{and} \quad D_{\bar{x}\bar{y}}\bar{\mathbf{G}}(0, 0, 0) = D_{\bar{x}\bar{y}}\bar{\mathbf{H}}(0, 0, 0) = \text{Id}.$$

Set

$$\bar{\delta} := \min\{4(1 + n)\delta_0, C_0\delta_0\}.$$

Finally, define

$$\bar{\mathbf{u}}(\bar{x}) := \mathbf{u}(x) - \mathbf{G}(x, y_0, v_0).$$

Arguing as in the proof of Theorem 1.1, we see that  $\bar{\mathbf{u}}$  is  $\bar{\mathbf{G}}$ -convex and  $(\mathbf{T}_{\bar{u}})_{\#}\bar{f} = \bar{g}$ . Furthermore, from (4.4), (4.3), and (4.39), we determine that

$$\left\| \bar{\mathbf{u}} - \frac{1}{2}|\bar{x}|^2 \right\|_{C^0(\mathcal{B}_3)} \leq 2\eta_0 + 2\delta_0 + C(\delta_0 + \eta_0^{1/2}) =: \bar{\eta}.$$

Indeed, recalling that  $v_0 := \mathbf{H}(x_0, y_0, \mathbf{u}(x_0))$ ,

$$\begin{aligned} \left| \bar{\mathbf{u}} - \frac{1}{2}|\bar{x}|^2 \right| &= \left| \mathbf{u}(x) - \mathbf{G}(x, y_0, v_0) - \frac{1}{2}x \cdot x + x \cdot x_0 + \mathbf{u}(x_0) - \mathbf{u}(x_0) + \frac{1}{2}x_0 \cdot x_0 - x_0 \cdot x_0 \right| \\ &\leq 2\eta_0 + |-\mathbf{G}(x, y_0, v_0) + x \cdot x_0 + \mathbf{u}(x_0) - x_0 \cdot x_0| \\ &= 2\eta_0 + |x \cdot y_0 - v_0 - \mathbf{G}(x, y_0, v_0) - x \cdot y_0 + v_0 - x_0 \cdot y_0 + \mathbf{u}(x_0) \\ &\quad + x_0 \cdot y_0 + x \cdot x_0 - x_0 \cdot x_0| \\ &\leq 2\eta_0 + 2\delta_0 + (|x| + |x_0|)|x_0 - y_0| \\ &\leq 2\eta_0 + 2\delta_0 + C(\delta_0 + \eta_0^{1/2}). \end{aligned}$$

In summary, we see that  $\bar{\mathbf{u}}, \bar{\mathbf{G}}, \bar{\mathbf{H}}, \bar{f}, \bar{g}, \bar{\mathcal{C}},$  and  $\bar{\mathcal{K}}$  satisfy the hypotheses of Proposition 4.2. Hence, taking  $\delta_0$  and  $\eta_0$ , in turn,  $\bar{\delta}$  and  $\bar{\eta}$ , sufficiently small, we find that  $\mathbf{u} \in C^{1,\beta}(B_{1/6})$ , as desired.  $\square$

An important corollary of Theorem 4.1 is a strict  $\mathbf{G}$ -convexity estimate for  $\mathbf{u}$  in  $B_{1/6}$ .

**Corollary 4.6.** *Under the hypotheses of Theorem 4.1, we find that  $\mathbf{u}$  is strictly  $\mathbf{G}$ -convex in  $B_{1/6}$ . More precisely, for all  $\sigma > 2$ , there exist constants  $\eta_0, \delta_0 > 0$ , depending on  $\sigma$  and dimension, such that for all  $x_0 \in B_{1/6}$ , we have that*

$$(4.42) \quad \inf_{\partial B_d(x_0)} \{\mathbf{u} - \mathcal{G}_{x_0, y_0, v_0}\} \geq c_0 d^\sigma \quad \forall d \leq \text{dist}(x_0, \partial B_{1/6})$$

for some constant  $c_0 = c_0(\sigma, n) > 0$ . Here,  $y_0 \in \partial_{\mathbf{G}} \mathbf{u}(x_0)$  and  $v_0 := \mathbf{H}(x_0, y_0, \mathbf{u}(x_0))$ .

*Proof.* Let  $\mathbf{u}_1$  be as in Proposition 4.2 and set  $d_0 := 3K_3 h_0^{1/2}$ . Then, from (4.38), we deduce that

$$\inf_{\partial B_d} \mathbf{u}_1 \geq d_0^\sigma d^\sigma \quad \forall d \leq d_0$$

provided  $d_0$  is sufficiently small depending on  $\sigma$  and dimension. In turn,

$$\inf_{\partial B_d(x_0)} \{\mathbf{u} - \mathcal{G}_{x_0, y_0, v_0}\} \geq h_0 d_0^\sigma \left( \frac{h_0^{1/2} d}{K_3} \right)^\sigma \geq c_0 d^\sigma \quad \forall d \leq \text{dist}(x_0, \partial B_{1/6})$$

taking  $c_0 = c_0(\sigma, n) > 0$  sufficiently small, as desired.  $\square$

From the strict  $\mathbf{G}$ -convexity of  $\mathbf{u}$  in  $B_{1/6}$ , we deduce that  $\mathbf{T}_{\mathbf{u}}(B_{1/6})$  is open, a key fact used in the proof of Theorem 1.1.

**Corollary 4.7.** *Under the hypotheses of Theorem 4.1, we have that  $\mathbf{T}_{\mathbf{u}}(B_{1/6})$  is open.*

*Proof.* Since  $\mathbf{u}$  is differentiable in  $B_{1/6}$ , we have that  $\mathbf{T}_{\mathbf{u}}(B_{1/6}) = \partial_{\mathbf{G}} \mathbf{u}(B_{1/6})$ . We show that for each  $x_0 \in B_{1/6}$ , there exists an  $\varepsilon_0 > 0$  such that for any  $y \in B_{\varepsilon_0}(y_0)$ , the map

$$z \mapsto \mathbf{H}(z, y, \mathbf{u}(z))$$

has a local maximum at some point  $x \in B_{1/6}$ . Here,  $\{y_0\} := \partial_{\mathbf{G}} \mathbf{u}(x_0)$ . If so, then

$$\nabla \mathbf{u}(x) = D_x \mathbf{G}(x, y, \mathbf{H}(x, y, \mathbf{u}(x)));$$

that is,  $\{y\} = \partial_{\mathbf{G}} \mathbf{u}(x)$  and  $B_{\varepsilon_0}(y_0) \subset \mathbf{T}_{\mathbf{u}}(B_{1/6})$ , as desired. To this end, let  $d > 0$  be such that  $B_d(x_0) \subset B_{1/6}$  and

$$x \in \operatorname{argmax}_{z \in \overline{B}_d(x_0)} \mathbf{H}(z, y, \mathbf{u}(z)).$$

Since  $\mathbf{u}(x) = \mathbf{G}(x, y, \mathbf{H}(x, y, \mathbf{u}(x)))$  and  $\mathbf{G}$  is decreasing in  $v$ , we observe that

$$\begin{aligned} \mathbf{u}(x) - \mathcal{G}_{x_0, y_0, v_0}(x) &\leq \mathbf{G}(x, y, \mathbf{H}(x_0, y, \mathbf{u}(x_0))) - \mathbf{G}(x, y_0, \mathbf{H}(x_0, y_0, \mathbf{u}(x_0))) \\ &= \mathbf{G}(x, y, \mathbf{H}(x_0, y, \mathbf{u}(x_0))) - \mathbf{G}(x, y_0, \mathbf{H}(x_0, y, \mathbf{u}(x_0))) \\ &\quad + \mathbf{G}(x, y_0, \mathbf{H}(x_0, y, \mathbf{u}(x_0))) - \mathbf{G}(x, y_0, \mathbf{H}(x_0, y_0, \mathbf{u}(x_0))) \\ &\leq C \varepsilon_0 \end{aligned}$$

with  $C := \|\mathbf{G}\|_{C^1(\mathcal{B}_4)}(1 + 2\|\mathbf{H}\|_{C^1(\mathcal{B}_4)})$ . Hence, taking  $\varepsilon_0 < c_0 d^\sigma / C$ , we see that

$$\mathbf{u}(x) - \mathcal{G}_{x_0, y_0, v_0}(x) < c_0 d^\sigma,$$

which, recalling (4.42), implies that  $x$  lives inside  $B_d(x_0)$  and not on its boundary.  $\square$

## 5. HIGHER REGULARITY

Here, we prove a higher regularity version of Theorem 4.1. To do this, we will need a more refined comparison-type principle than the one established in Lemma 4.3. The comparison-type principle in this section makes use of a change of variables formula for the  $\mathbf{G}$ -exponential map, Lemma 5.1, and the coincidence of the  $\mathbf{G}$ -subdifferential of  $\mathbf{u}$  at  $x$  and the  $\mathbf{G}$ -exponential map at  $(x, \mathbf{u}(x), \nabla \mathbf{u}(x))$  when  $\mathbf{u}$  is differentiable, Remark 5.2.

Given a  $\mathbf{G}$ -convex function  $\mathbf{w}$  on an open set  $\mathcal{O}$ , we have defined  $\mathbf{T}_{\mathbf{w}}(x) := \mathbf{G}\text{-exp}_{x, \mathbf{w}(x)}(\nabla \mathbf{w}(x))$ . Yet even when  $\mathbf{w}$  is not  $\mathbf{G}$ -convex, we may still consider  $\mathbf{T}_{\mathbf{w}}(x)$  if  $\{(x, \mathbf{w}(x), \nabla \mathbf{w}(x)) : x \in \mathcal{O}\} \subset \text{dom } \mathbf{G}\text{-exp}$ .

**Lemma 5.1.** *Let  $\mathcal{O} \subset \mathbb{R}^n$  be open,  $\mathbf{w} \in C^2(\mathcal{O})$ , and  $\{(x, \mathbf{w}(x), \nabla \mathbf{w}(x)) : x \in \mathcal{O}\} \subset \text{dom } \mathbf{G}\text{-exp}$ . If*

$$D^2 \mathbf{w}(x) - D_x^2 \mathbf{G}(x, \mathbf{T}_{\mathbf{w}}(x), \mathbf{H}(x, \mathbf{T}_{\mathbf{w}}(x), \mathbf{w}(x))) \geq 0 \quad \forall x \in \mathcal{O},$$

*then for every Borel set  $E \subset \mathcal{O}$ ,*

$$|\mathbf{T}_{\mathbf{w}}(E)| \leq \int_E \frac{\det(D^2 \mathbf{w}(x) - D_x^2 \mathbf{G}(x, \mathbf{T}_{\mathbf{w}}(x), \mathbf{H}(x, \mathbf{T}_{\mathbf{w}}(x), \mathbf{w}(x))))}{|\det(\mathbf{E}(x, \mathbf{T}_{\mathbf{w}}(x), \mathbf{H}(x, \mathbf{T}_{\mathbf{w}}(x), \mathbf{w}(x))))|} dx.$$

*In addition, if the map  $\mathbf{T}_{\mathbf{w}}$  is injective, then equality holds.*

*Proof.* After differentiating the identity

$$\nabla \mathbf{w}(x) = D_x \mathbf{G}(x, \mathbf{T}_{\mathbf{w}}(x), \mathbf{H}(x, \mathbf{T}_{\mathbf{w}}(x), \mathbf{w}(x))),$$

we see that the Jacobian determinant of the  $C^1$  map  $x \mapsto \mathbf{T}_{\mathbf{w}}(x)$  is the integrand above. Thus, applying the Area Formula (see, e.g., [6]) concludes the proof.  $\square$

**Remark 5.2.** Recall, by (2.9), that if  $\mathbf{u}$  is differentiable at  $x$  and  $\mathbf{G}$ -convex, then

$$\partial_{\mathbf{G}} \mathbf{u}(x) = \{\mathbf{G}\text{-exp}_{x, \mathbf{u}(x)}(\nabla \mathbf{u}(x))\} = \{\mathbf{T}_{\mathbf{u}}(x)\}.$$

Let  $\text{co}[E]$  denote the convex hull of the set  $E$ . Also, recall that  $\mathcal{N}_r(E)$  denotes the  $r$ -neighborhood of a set  $E$ . The following comparison-type principle compares  $\mathbf{G}$ -convex functions of class  $C^1$  and smooth solutions of Monge–Ampère equations.

**Proposition 5.3.** *Let  $R \geq 3$  and  $\mathbf{u} \geq 0$  be a  $\mathbf{G}$ -convex function of class  $C^1$  such that  $\mathbf{u}(0) = 0$  and*

$$(5.1) \quad B_{1/R} \subset S := \{\mathbf{u} \leq 1\} \subset B_R.$$

*Assume  $\mathcal{B}(R, S) := B_{2R} \times \mathbf{T}_{\mathbf{u}}(S) \times (-R^3, R^3) \subset \mathfrak{g}, \mathfrak{h}$  and  $B_{2R} \times (-R^3, R^3) \times \nabla \mathbf{u}(S) \Subset \text{dom } \mathbf{G}\text{-exp}$ . Suppose that  $f$  and  $g$  are two densities such that  $(\mathbf{T}_{\mathbf{u}})_{\#} f = g$  and*

$$(5.2) \quad \left\| \frac{f}{\lambda_1} - 1 \right\|_{C^0(S)} + \left\| \frac{g}{\lambda_2} - 1 \right\|_{C^0(\mathbf{T}_{\mathbf{u}}(S))} \leq \varepsilon$$

*for some constants  $\lambda_1/\lambda_2 \in (1/2, 2)$  and  $\varepsilon \in (0, 1/4)$ . Furthermore, assume that*

$$(5.3) \quad \|\mathbf{G} - x \cdot y + v\|_{C^2(\mathcal{B}(R, S))} + \|\mathbf{H} - x \cdot y + u\|_{C^2(\mathcal{B}(R, S))} \leq \delta.$$

*Then, there exists constants  $\gamma = \gamma(n, R) \in (0, 1)$  and  $\delta_1 = \delta_1(n, R) > 0$  such that the following holds: if  $\mathbf{w}$  is convex and satisfies*

$$\begin{cases} \det(D^2 \mathbf{w}) = \lambda_1/\lambda_2 & \text{in } \mathcal{N}_{\delta^\gamma}(\text{co}[S]) \\ \mathbf{w} = 1 & \text{on } \partial \mathcal{N}_{\delta^\gamma}(\text{co}[S]), \end{cases}$$

*then*

$$(5.4) \quad \|\mathbf{u} - \mathbf{w}\|_{C^0(S)} \leq K(\varepsilon + \delta^{\gamma/n})$$

*provided  $\delta \leq \delta_1$ . Here,  $K = K(n, R) > 0$ .*

The proof of Proposition 5.3 follows the proof of [5, Proposition 5.2]. Yet because the map  $\mathbf{T}_{\mathbf{u}}$  depends on  $x$ ,  $\mathbf{u}$ , and  $\nabla \mathbf{u}$  and not just on  $x$  and  $\nabla \mathbf{u}$ , the argument is more delicate.

*Proof.* Recall that  $\mathbf{u} + \delta|x|^2$  is convex by (5.3). Thus, as  $\mathbf{u}(0) = 0$ ,  $\mathbf{u} = 1$  on  $\partial S$ , and  $S \subset B_R$ , using (2.10), it follows that

$$(5.5) \quad |D_x \mathbf{G}(x, y, v)| = |\nabla \mathbf{u}(x)| \geq |\nabla \mathbf{u}(x) + 2\delta x| - 2\delta|x| \geq \frac{1}{R} - 2\delta R \geq \frac{1}{2R} \quad \forall x \in \partial S$$

provided that  $\delta$  is small enough. Here,  $y := \mathbf{T}_{\mathbf{u}}(x)$  and  $v := \mathbf{H}(x, y, \mathbf{u}(x))$ . Now consider

$$\mathcal{S} := \bigcap_{x \in \partial S} E_x$$

where

$$E_x := \{z \in B_R : \mathbf{G}(z, y, v) \leq 1\}.$$

Clearly,  $S \subset \mathcal{S}$ . Let  $z \notin S$  and  $x \in \partial S$  be a point such that  $\text{dist}(z, \partial S) = |x - z| > 0$ . If  $|z - x| < 1/\delta R$ , then using (5.3) and (5.5), we find that

$$\mathbf{G}(z, y, v) - 1 \geq |\nabla \mathbf{u}(x)| |z - x| - \frac{\delta}{2} |z - x|^2 > 0,$$

and  $z \notin \mathcal{S}$ . On the other hand, if  $|z - x| \geq 1/\delta R$ , then by (5.3) and (5.5), we have that

$$\begin{aligned} \mathbf{G}(z, y, v) - 1 &\geq |z - x| |\nabla \mathbf{u}(x)| + (z - x) \cdot (y - D_x \mathbf{G}(x, y, v)) - 2\delta \\ &\geq \frac{1}{2\delta R^2} - 2R\delta - 2\delta. \end{aligned}$$

And so  $\mathbf{G}(z, y, v) - 1 > 0$  and  $z \notin \mathcal{S}$  provided that  $\delta$  is sufficiently small. In turn,

$$\mathcal{S} = S$$

if  $\delta > 0$  is sufficiently small depending only on  $R$ . It follows that  $S$  is a  $(CR\delta)$ -semiconvex set. Now arguing exactly as in [5, Proposition 5.2], we have that

$$(5.6) \quad \text{osc}_S \mathbf{w} \leq K_{R,n},$$

$$(5.7) \quad 1 - K_{R,n} \delta^{\gamma/n} \leq \mathbf{w} < 1 \quad \text{on } \partial S,$$

and

$$(5.8) \quad D^2 \mathbf{w} \geq \frac{\delta^{\gamma/\tau}}{K_{R,n}} \text{Id} \quad \text{in } \text{co}[S]$$

for some constants  $K_{R,n}$  and  $\tau > 0$  depending only on dimension and  $R$ .

Define

$$\mathbf{w}^+ := (1 + 3\varepsilon + 2\delta^{1/2})\mathbf{w} - 3\varepsilon - 2\delta^{1/2}$$

and

$$\mathbf{w}^- := \left(1 - 3\varepsilon - \frac{\delta^{1/2}}{2}\right)\mathbf{w} + 3\varepsilon + \frac{\delta^{1/2}}{2} + K_{R,n} \delta^{\gamma/n}.$$

We claim that if  $\gamma$  is sufficiently small, then  $\mathbf{w}^- \geq \mathbf{u} \geq \mathbf{w}^+$  in  $S$ . If so, then (5.6) will imply (5.4), as desired.

Choose  $\gamma := \tau/4$ . By (5.7), we have that  $\mathbf{w}^- > \mathbf{u} > \mathbf{w}^+$  on  $\partial S$ . We first show that  $\mathbf{u} \geq \mathbf{w}^+$  in  $S$ . Suppose not. Then, as  $\mathbf{u} > \mathbf{w}^+$  on  $\partial S$ , we see that

$$\emptyset \neq Z := \{\mathbf{u} < \mathbf{w}^+\} \Subset S.$$

Thanks to (5.8) and (5.3), we have that

$$(5.9) \quad D^2 \mathbf{w}^+ > D^2 \mathbf{w} \geq \frac{\delta^{1/4}}{K_{R,n}} \text{Id} > \delta \text{Id} \geq \|D_x^2 \mathbf{G}\|_{C^0(\mathcal{B}(R,S))} \text{Id} \quad \text{in } \text{co}[S]$$

provided that  $\delta$  is sufficiently small depending on  $R$  and  $n$ . Notice that  $\mathbf{w}^+(Z) \subset (-2R^2, 1)$ .<sup>6</sup> Moving any supporting plane to  $\mathbf{w}^+$  in  $Z$  down and then up until it touches  $\mathbf{u}$  from below, we see that

$$\nabla \mathbf{w}^+(Z) \subset \nabla \mathbf{u}(Z).$$

It follows that  $\{(x, \mathbf{w}^+(x), \nabla \mathbf{w}^+(x)) : x \in Z\} \subset \text{dom } \mathbf{G}\text{-exp}$ . Let  $x_0^+ \in Z$ ,  $y_0 := \mathbf{T}_{\mathbf{w}^+}(x_0^+)$ , and  $v_0^+ := \mathbf{H}(x_0^+, y_0, \mathbf{w}^+(x_0^+))$ . Increase and then decrease  $v_0^+$  to  $v_0$  so that  $\mathbf{G}(\cdot, y_0, v_0)$  touches  $\mathbf{u}$  from below at  $x_0$ . Recall that  $\mathbf{T}_{\mathbf{w}^+}(x_0^+) = y_0$  if and only if  $\nabla \mathbf{w}^+(x_0^+) = D_x \mathbf{G}(x_0^+, y_0, v_0^+)$ . (See Section 2.2.) Hence, from (5.9) and as  $\mathbf{G}(x_0^+, y_0, v_0^+) = \mathbf{w}^+(x_0^+)$ , we have that

$$\mathbf{G}(x, y_0, v_0^+) \leq \mathbf{w}^+(x_0^+) + \nabla \mathbf{w}^+(x_0^+) \cdot (x - x_0^+) + \frac{\delta}{2} |x - x_0^+|^2 \leq \mathbf{w}^+(x) \quad \forall x \in \text{co}[S].$$

In turn,  $x_0 \in Z$ . Indeed, if not, then

$$\mathbf{G}(x_0, y_0, v_0) = \mathbf{u}(x_0) \geq \mathbf{w}^+(x_0) \geq \mathbf{G}(x_0, y_0, v_0^+),$$

<sup>6</sup> By (5.7) and the convexity of  $\mathbf{w}^+$ , we see that  $\mathbf{w}^+ < 1$  in  $Z \subset S$ . The inclusion  $\mathbf{w}^+(Z) \subset (-2R^2, 1)$  then follows from considering the lower barrier (for  $\mathbf{w}$ )

$$\frac{\lambda_1^{1/n}}{2\lambda_2^{1/n}} (|x|^2 - (R + \delta^\gamma)^2) + 1$$

and taking  $\delta$  smaller if needed.

from which using (G-Mono), it follows that

$$\mathbf{w}^+(x_0^+) > \mathbf{u}(x_0^+) \geq \mathbf{G}(x_0^+, y_0, v_0) \geq \mathbf{G}(x_0^+, y_0, v_0^+) = \mathbf{w}^+(x_0^+).$$

Impossible; and we deduce that

$$(5.10) \quad \mathbf{T}_{\mathbf{w}^+}(Z) \subset \mathbf{T}_{\mathbf{u}}(Z).$$

Now for any  $x \in Z$ , from (5.8) and taking  $\delta$  even smaller, we compute that

$$D^2\mathbf{w}^+(x) - D_x^2\mathbf{G}(x, \mathbf{T}_{\mathbf{w}^+}(x), \mathbf{H}(x, \mathbf{T}_{\mathbf{w}^+}(x), \mathbf{w}^+(x))) \geq (1 + 3\varepsilon + \delta^{1/2})D^2\mathbf{w}(x).$$

And so by (5.3), we see that

$$\frac{\det(D^2\mathbf{w}^+(x) - D_x^2\mathbf{G}(x, \mathbf{T}_{\mathbf{w}^+}(x), \mathbf{H}(x, \mathbf{T}_{\mathbf{w}^+}(x), \mathbf{w}^+(x))))}{|\det(\mathbf{E}(x, \mathbf{T}_{\mathbf{w}^+}(x), \mathbf{H}(x, \mathbf{T}_{\mathbf{w}^+}(x), \mathbf{w}^+(x))))|} \geq \frac{(1 + 3\varepsilon + \delta^{1/2})^n \lambda_1}{(1 + \delta)^n \lambda_2}.$$

Moreover, from (5.9), for any  $x, z \in Z$  with  $x \neq z$ , setting  $y := \mathbf{T}_{\mathbf{w}^+}(x)$  and  $v := \mathbf{H}(x, y, \mathbf{w}^+(x))$ , we determine that

$$\mathbf{w}^+(z) - \mathbf{G}(z, y, v) = \frac{1}{2} \int_0^1 (D^2\mathbf{w}^+(tz + (1-t)x) - D_x^2\mathbf{G}(tz + (1-t)x, y, v))(z-x) \cdot (z-x) dt > 0.$$

In other words, the function  $\mathbf{G}(\cdot, y, v)$  only touches  $\mathbf{w}^+$  at  $x$ , and the map  $x \mapsto \mathbf{T}_{\mathbf{w}^+}(x)$  is injective in  $Z$ . Therefore, Lemma 5.1 yields

$$|\mathbf{T}_{\mathbf{w}^+}(x)| \geq \frac{(1 + 3\varepsilon + \delta^{1/2})^n \lambda_1}{(1 + \delta)^n \lambda_2} |Z| > (1 + 3\varepsilon) \frac{\lambda_1}{\lambda_2} |Z|$$

if  $\delta$  is small enough depending only on  $R$  and  $n$ . On the other hand, since  $\mathbf{u}$  is  $C^1$  in  $S$ , the push-forward condition and (5.2) imply that

$$|\mathbf{T}_{\mathbf{u}}(Z)| = \int_Z \frac{f(x)}{g(\mathbf{T}_{\mathbf{u}}(x))} dx \leq \frac{1 + \varepsilon \lambda_1}{1 - \varepsilon \lambda_2} |Z|.$$

Combining these last two inequalities, we find that (5.10) is impossible unless  $Z$  is empty. That is,  $\mathbf{w}^+ \leq \mathbf{u}$  in  $S$ .

The argument showing that  $\mathbf{u} \leq \mathbf{w}^-$  in  $S$  is similar to the one just presented, showing that  $\mathbf{u} \geq \mathbf{w}^+$  in  $S$ . So we only provide a sketch. Again, suppose, to the contrary, that  $W := \{\mathbf{u} > \mathbf{w}^-\}$  is non-empty. Now we can find a positive constant  $\mu$  so that  $\mathbf{u}$  touches  $\mathbf{w}^- + \mu$  from below in  $S$ . As both  $\mathbf{u}$  and  $\mathbf{w}^-$  are  $C^1$ , it follows that  $\nabla \mathbf{u} = \nabla \mathbf{w}^-$  on the set  $\{\mathbf{u} = \mathbf{w}^- + \mu\}$ . Therefore, if  $\eta > 0$  is sufficiently small, then the set  $W_\eta := \{\mathbf{u} > \mathbf{w}^- + \mu - \eta\}$  is non-empty and  $\nabla \mathbf{w}^-(W_\eta)$  is contained in a small neighborhood of  $\nabla \mathbf{u}(W_\eta)$ .

Set  $\mathbf{w}_\eta^- := \mathbf{w}^- + \mu - \eta$ . Then, using the same barrier as before, we find that  $\mathbf{w}_\eta^- \in (-R^2, 2)$ . Hence,  $\{(x, \mathbf{w}_\eta^-(x), \nabla \mathbf{w}_\eta^-(x)) : x \in W_\eta\} \subset \text{dom } \mathbf{G}\text{-exp}$ . Let  $\mathbf{G}(\cdot, y_0, v_0^-)$  be the  $\mathbf{G}$ -support for  $\mathbf{u}$  at  $x_0^- \in W_\eta$ . Increase and then decrease  $v_0^-$  to  $v_0$  so that  $\mathbf{G}(\cdot, y_0, v_0)$  touches  $\mathbf{w}_\eta^-$  from below, and let  $x_0$  be the point at which  $\mathbf{G}(\cdot, y_0, v_0)$  touches  $\mathbf{w}_\eta^-$  from below. Notice that  $x_0 \in W_\eta$  and  $v_0 = \mathbf{H}(x_0, y_0, \mathbf{w}_\eta^-(x_0))$ . Therefore,  $D_x \mathbf{G}(x_0, y_0, v_0) = \nabla \mathbf{w}_\eta^-(x_0)$ ; that is,  $y_0 \in \mathbf{T}_{\mathbf{w}_\eta^-}(W_\eta)$  or

$$\mathbf{T}_{\mathbf{u}}(W_\eta) \subset \mathbf{T}_{\mathbf{w}_\eta^-}(W_\eta).$$

Observe that from (5.8),

$$(1 - 3\varepsilon - \delta^{1/2})D^2\mathbf{w}(x) \leq D^2\mathbf{w}_\eta^-(x) - D_x^2\mathbf{G}(x, \mathbf{T}_{\mathbf{w}_\eta^-}(x), \mathbf{H}(x, \mathbf{T}_{\mathbf{w}_\eta^-}(x), \mathbf{w}_\eta^-(x)))$$

if  $\delta^{3/4} \geq 2K_{R,n}\delta$ . Also, taking  $\delta$  even smaller (so that  $\delta^{3/4} \geq 4K_{R,n}\delta$  or, equivalently,  $\delta^{1/2} \geq 4K_{R,n}\delta^{3/4}$ ), we find that

$$D^2\mathbf{w}_\eta^-(x) - D_x^2\mathbf{G}(x, \mathbf{T}_{\mathbf{w}_\eta^-}(x), \mathbf{H}(x, \mathbf{T}_{\mathbf{w}_\eta^-}(x), \mathbf{w}_\eta^-(x))) \leq \left(1 - 3\varepsilon - \frac{\delta^{1/2}}{4}\right) D^2\mathbf{w}.$$

Without loss of generality, we assume that  $\delta^{1/2} \leq 1/4$ ; whence,  $1 - 3\varepsilon - \delta^{1/2} \geq 0$ . Hence, by Lemma 5.1,

$$|\mathbf{T}_{\mathbf{w}_\eta^-}(W_\eta)| \leq \frac{(1 - 3\varepsilon - \delta^{1/2}/4)^n \lambda_2}{(1 - \delta)^n \lambda_2} |W_\eta| < (1 - 3\varepsilon) \frac{\lambda_2}{\lambda_2} |W_\eta|.$$

Moreover,

$$|\mathbf{T}_{\mathbf{u}}(W_\eta)| \geq \frac{1 - \varepsilon \lambda_1}{1 + \varepsilon \lambda_2} |W_\eta|.$$

Like before, combining these last two inequalities, we arrive at a contradiction unless  $|W_\eta| = 0$ , so long as  $\delta$  is sufficiently small depending on  $R$  and dimension.  $\square$

With Proposition 5.3 in hand, our next proposition is a higher regularity version of Proposition 4.2.

**Proposition 5.4.** *In addition to the hypotheses of Proposition 4.2, suppose  $f \in C^{0,\alpha}(\mathcal{C})$ ,  $g \in C^{0,\alpha}(\mathcal{K})$ , and  $B_6 \times (-27, 27) \times B_6 \subset \text{dom } \mathbf{G}\text{-exp}$ . There exist positive constants  $\delta' \leq \delta$  and  $\eta' \leq \eta$  such that the following holds: if*

$$(5.11) \quad \begin{aligned} & \|f - 1\|_{L^\infty(\mathcal{C})} + \|g - 1\|_{L^\infty(\mathcal{K})} \leq \delta', \\ & \|\mathbf{G} - x \cdot y + v\|_{C^2(\mathcal{B}_3)} + \|\mathbf{H} - x \cdot y + v\|_{C^2(\mathcal{B}_3)} \leq \delta', \end{aligned}$$

and

$$\left\| \mathbf{u} - \frac{1}{2}|x|^2 \right\|_{C^0(\mathcal{B}_3)} \leq \eta',$$

then  $\mathbf{u} \in C^{2,\alpha'}(0)$  for some  $\alpha' < \alpha$ .

The proof of Proposition 5.4 follows arguing exactly as in the proof of [5, Theorem 5.3]. That said, let us make some remarks. An inspection of the proof of [5, Theorem 5.3] reveals that, apart from a comparison-type principle like Proposition 5.3, we will need that the sum of the norms in (5.11) decays under parabolically quadratic rescalings. (The remainder of the proof uses classical estimates for the Monge–Ampère equation.) The assumption  $\mathbf{G}, \mathbf{H} \in C^{2,\alpha}(\mathcal{B}_3)$  plus (4.5) and (4.6) guarantee this decay. Finally, since the domain of the  $\mathbf{G}\text{-exp}$  map includes the product of three open sets at the beginning, the third of which compactly contains  $\nabla \mathbf{u}(B_{1/3})$ , the set inclusions in the hypotheses of Proposition 5.3 will be satisfied at each stage of the iteration by construction; we are zooming in with the correct rescaling. So applying Proposition 5.4 at every point in  $B_{1/7}$  and then classical Schauder estimates, we obtain our final theorem.

**Theorem 5.5.** *In addition to the hypotheses of Theorem 4.1, suppose  $\mathbf{G}, \mathbf{H} \in C^{k,\alpha}(\mathcal{B}_4)$ ,  $f \in C^{k,\alpha}(\mathcal{C})$ ,  $g \in C^{k,\alpha}(\mathcal{K})$ , for some  $k \geq 0$  and  $\alpha \in (0, 1)$ , and  $B_8 \times (-64, 64) \times B_8 \subset \text{dom } \mathbf{G}\text{-exp}$ . There exist positive constants  $\delta_1 \leq \delta_0$  and  $\eta_1 \leq \eta_0$  such that the following holds: if*

$$\begin{aligned} & \|f - \mathbf{1}_{\mathcal{C}}\|_{L^\infty(\mathcal{B}_4)} + \|g - \mathbf{1}_{\mathcal{K}}\|_{L^\infty(\mathcal{B}_4)} \leq \delta_1, \\ & \|\mathbf{G} - x \cdot y + v\|_{C^2(\mathcal{B}_4)} + \|\mathbf{H} - x \cdot y + v\|_{C^2(\mathcal{B}_4)} \leq \delta_1, \end{aligned}$$

and

$$\left\| \mathbf{u} - \frac{1}{2}|x|^2 \right\|_{C^0(\mathcal{B}_4)} \leq \eta_1,$$

then  $\mathbf{u} \in C^{k+2,\alpha}(B_{1/8})$ .

**Acknowledgments.** I would like to thank Jun Kitagawa for a careful reading of preliminary versions of this paper and discussions and Nestor Guillen for a helpful conversation. Also, I am grateful to both Connor Mooney and Alessio Figalli for their support, encouragement, and valuable conversations. Finally, I wish to thank the referee for his or her detailed reading and comments.

## REFERENCES

- [1] G. A. Beer, *The Hausdorff metric and convergence in measure*, Michigan Math. J. **21** (1974), 63-64.
- [2] L. A. Caffarelli, *The regularity of mappings with a convex potential*, J. Amer. Math. Soc. **5**(1) (1992), 99-104.
- [3] S. Chen and A. Figalli, *Boundary  $\varepsilon$ -regularity in optimal transportation*, Adv. Math. **273** (2015), 540-567.
- [4] S. Chen and A. Figalli, *Partial  $W^{2,p}$  regularity for optimal transport maps*, J. Funct. Anal. **272** (2017), no. 11, 4588-4605.
- [5] G. De Philippis and A. Figalli, *Partial regularity for optimal transport maps*, Publ. Math. Inst. Hautes Études Sci. (2015), no. 121, 81-112.
- [6] L. C. Evans and R. F. Gariepy, “Measure Theory and Fine Properties of Functions”, CRC Press Inc. Boca Raton, 1992.
- [7] A. Fathi and A. Figalli, *Optimal transportation on non-compact manifolds*, Israel J. Math. **175** (2010), no. 1, 1-59.
- [8] A. Figalli, *Regularity properties of optimal maps between nonconvex domains in the plane*, Comm. Partial Differential Equations **35** (2010), no. 3, 465-479.
- [9] A. Figalli and Y.-H. Kim, *Partial regularity of Brenier solutions of the Monge-Ampère equation*, Discrete Contin. Dyn. Syst. **28** (2010), no. 2, 559-565.
- [10] M. Goldman and F. Otto, *A variational proof of partial regularity for optimal transportation maps*, arXiv:1704.05339.
- [11] N. Guillen and J. Kitagawa, *Pointwise estimates and regularity in geometric optics and other generated Jacobian equations*, Comm. Pure Appl. Math. **70** (2017), no. 6, 1146-1220.

- [12] Y. Jhaveri, *On the (in)stability of the identity map in optimal transportation*, arXiv:1710.03708.
- [13] A. Karakhanyan and X.-J. Wang, *On the reflector shape design*, *J. Differential Geom.* **84** (2010), no. 3, 561-610.
- [14] X.-N. Ma, N. S. Trudinger, and X.-J. Wang, *Regularity of potential functions of the optimal transportation problem*, *Arch. Ration. Mech. Anal.* **177** (2005), no. 2, 151-183.
- [15] N. S. Trudinger, *On the local theory of prescribed Jacobian equations*, *Discrete Contin. Dyn. Syst.* **34** (2014), no. 4, 1663-1681.
- [16] C. Villani, "Optimal Transport, Old and New", *Grundlehren des mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*, Vol. 338, Springer-Verlag Berlin Heidelberg, 2009.

ETH ZÜRICH, DEPARTMENT OF MATHEMATICS, RÄMISTRASSE 101, ZÜRICH 8092, SWITZERLAND  
Email address: yash.jhaveri@math.ethz.ch