Hölder continuity and injectivity of optimal maps*

Alessio Figalli[†], Young-Heon Kim[‡] and Robert J. McCann[§]

Abstract

Consider transportation of one distribution of mass onto another, chosen to optimize the total expected cost, where cost per unit mass transported from x to y is given by a smooth function $c(x, y)$. If the source density $f^+(x)$ is bounded away from zero and infinity in an open region $U' \subset \mathbf{R}^n$, and the target density $f^-(y)$ is bounded away from zero and infinity on its support $\overline{V} \subset \mathbf{R}^n$, which is strongly c-convex with respect to U', and the transportation cost c satisfies the $\mathbf{(A3)_w}$ condition of Trudinger and Wang [51], we deduce local Hölder continuity and injectivity of the optimal map inside U' (so that the associated potential u belongs to $C^{1,\alpha}_{loc}(U')$). Here the exponent $\alpha > 0$ depends only on the dimension and the bounds on the densities, but not on c. Our result provides a crucial step in the low/interior regularity setting: in a sequel [17], we use it to establish regularity of optimal maps with respect to the Riemannian distance squared on arbitrary products of spheres. Three key tools are introduced in the present paper. Namely, we first find a transformation that under $\mathbf{(A3)_w}$ makes c-convex functions level-set convex (as was also obtained independently from us by Liu [41]). We then derive new Alexandrov type estimates for the level-set convex c-convex functions, and a topological lemma showing optimal maps do not mix interior with boundary. This topological lemma, which does not require $\mathbf{(A3)_w}$, is needed by Figalli and Loeper [20] to conclude continuity of optimal maps in two dimensions. In higher dimensions, if the densities $f^\pm$ are Hölder continuous, our result permits continuous differentiability of the map inside U' (in fact, $C^{2,\alpha}_{loc}$ regularity of the associated potential) to be deduced from the work of Liu, Trudinger and Wang [42].

---

*An earlier draft of this paper was circulated under the title *Continuity and injectivity of optimal maps for non-negatively cross-curved costs* at `arxiv.org/abs/0911.3952` The authors are grateful to the Institute for Pure and Applied Mathematics at UCLA, the Institut Fourier at Grenoble, and the Fields Institute in Toronto, for their generous hospitality during various stages of this work. A.F. is supported in part by NSF grant DMS-0969962. Y-H.K. is supported partly by NSF grant DMS-0635607 through the membership at Institute for Advanced Study at Princeton NJ, and also in part by NSERC grant 371642-09. R.J.M. is supported in part by NSERC grants 217006-03 and -08 and NSF grant DMS-0354729. Any opinions, findings and conclusions or recommendations expressed in this material are those of authors and do not reflect the views of either the Natural Sciences and Engineering Research Council of Canada (NSERC) or the United States National Science Foundation (NSF).

[†]Department of Mathematics, University of Texas at Austin, Austin TX USA `figalli@math.utexas.edu`

[‡]Department of Mathematics, University of British Columbia, Vancouver BC Canada `yhkim@math.ubc.ca` and School of Mathematics, Institute for Advanced Study at Princeton NJ USA

[§]Department of Mathematics, University of Toronto, Toronto Ontario Canada M5S 2E4 `mccann@math.toronto.edu`

Contents

1 Introduction

Given probability densities $0 \leq f^{\pm} \in L^1(\mathbf{R}^n)$ with respect to Lebesgue measure $\mathscr{L}^n$ on $\mathbf{R}^n$, and a cost function $c : \mathbf{R}^n \times \mathbf{R}^n \longmapsto [0, +\infty]$, Monge's transportation problem is to find a map $G : \mathbf{R}^n \longmapsto \mathbf{R}^n$ pushing $d\mu^+ = f^+ d\mathscr{L}^n$ forward to $d\mu^- = f^- d\mathscr{L}^n$ which minimizes the expected transportation cost [47]

$$\inf_{G_\# \mu^+ = \mu^-} \int_{\mathbf{R}^n} c(x, G(x)) d\mu^+(x), \tag{1.1}$$

where $G_\# \mu^+ = \mu^-$ means $\mu^-[Y] = \mu^+[G^{-1}(Y)]$ for each Borel $Y \subset \mathbf{R}^n$.

In this context it is interesting to know when a map attaining this infimum exists; sufficient conditions for this were found by Gangbo [25] and by Levin [40], extending work of a number of authors described in [26] [55]. One may also ask when G will be smooth, in which case it must satisfy the prescribed Jacobian equation $|\det DG(x)| = f^+(x)/f^-(G(x))$, which turns out to reduce to a degenerate elliptic partial differential equation of Monge-Ampère type for a scalar potential u satisfying $Du(\tilde{x}) = -D_x c(\tilde{x}, G(\tilde{x}))$. Sufficient conditions for this were discovered by Ma, Trudinger and Wang [46] and Trudinger and Wang [51] [52], after results for the special case $c(x, y) = |x - y|^2/2$ had been worked out by Brenier [4], Delanöe [12], Caffarelli [6] [5] [7] [8] [9], and Urbas [53], and for the cost $c(x, y) = -\log|x - y|$ and measures supported on the unit sphere by Wang [57].

If the ratio $f^+(x)/f^-(y)$ — although bounded away from zero and infinity — is not continuous, the map G will not generally be differentiable, though one may still hope for it to be continuous. This question is not merely of technical interest, since discontinuities in $f^\pm$ arise unavoidably in applications such as partial transport problems [10] [3] [15] [16]. Such results were established for the classical cost $c(x, y) = |x - y|^2/2$ by Caffarelli [5] [7] [8], for its restriction to the product of the boundaries of two strongly convex sets by Gangbo and McCann [27], and for more general costs satisfying the strong regularity hypothesis **(A3)** of Ma, Trudinger and Wang [46] — which excludes the cost $c(x, y) = |x - y|^2/2$ — by Loeper [43]; see also [35] [41] [52]. Under the weaker hypothesis $\mathbf{(A3)_w}$ of Trudinger and Wang [51], which includes the cost $c(x, y) = |x - y|^2/2$ (and whose necessity for regularity was shown by Loeper [43], see also [23]), such a result remained absent from the literature; the aim of this paper is to fill this gap, see Theorem 2.1 below.

A number of interesting cost functions do satisfy hypothesis $\mathbf{(A3)_w}$, and have applications in economics [18] and statistics [49]. Examples include the Euclidean distance squared between two convex graphs over two sufficiently convex sets in $\mathbf{R}^n$ [46], the simple harmonic oscillator action [39], and the Riemannian distance squared on the following spaces: the round sphere [44] and perturbations thereof [13] [21] [22], multiple products of round spheres (and their Riemannian submersion quotients, including products of complex projective spaces) [37], and products of perturbed 2-dimensional spheres [14]. As remarked in [36], for graphs which fail to be strongly convex and all Riemannian product geometries, the stronger condition **(A3)** necessarily fails. In a sequel, we apply the techniques developed here to deduce regularity of optimal maps for the multiple products of round spheres [17]. Moreover, Theorem 2.1 allows one to apply the higher interior regularity results established by Liu, Trudinger and Wang [42], ensuring in particular that the transport map is C^∞-smooth if f^+ and f^- are.

Most of the regularity results quoted above derive from one of two approaches. The continuity method, used by Delanoë, Urbas, Ma, Trudinger and Wang, is a time-honored technique for solving nonlinear equations. Here one perturbs a manifestly soluble problem (such as $|\det DG_0(x)| = f^+(x)/f_0(G_0(x))$ with $f_0 = f^+$, so that $G_0(x) = x$) to the problem of interest ($|\det DG_1(x)| = f^+(x)/f_1(G_1(x))$, $f_1 = f^-$) along a family $\{f_t\}_t$ designed to ensure the set of $t \in [0, 1]$ for which it is soluble is both open and closed. Openness follows from linearization and non-degenerate ellipticity using an implicit function theorem. For the non-degenerate ellipticity and closedness, it is required to establish estimates on the size of derivatives of the solutions (assuming such solutions exist) which depend only on information known a priori about the data (c, f_t). In this way one obtains smoothness of the solution $y = G_1(x)$ from the same argument which shows G_1 to exist.

The alternative approach relies on first knowing existence and uniqueness of a Borel map which solves the problem in great generality, and then deducing continuity or smoothness by close examination of this map after imposing additional conditions on the data $(c, f^\pm)$. Although precursors can be traced back to Alexandrov [2], in the present context this method was largely developed and refined by Caffarelli [5] [7] [8], who used convexity of u crucially to localize the map $G(x) = Du(x)$ and renormalize its behaviour near a point $(\tilde{x}, G(\tilde{x}))$ of interest in the borderline case $c(x, y) = -\langle x, y\rangle$. For non-borderline **(A3)** costs, simpler estimates suffice to deduce continuity of G, as in [27] [11] [43] [52]; in this case Loeper was actually able to deduce an explicit bound $\alpha = (4n-1)^{-1}$ on the Hölder exponent of G when $n > 1$. This bound was recently improved to its sharp value $\alpha = (2n-1)^{-1}$ by Liu [41], using a key observation discovered independently from us (see Section 4 and Theorem 4.3); both Loeper and Liu also obtained explicit exponents $\alpha = \alpha(n,p)$ for $f^+ \in L^p$ with $p > n$ [43] or $p > (n+1)/2$ [41] and $1/f^- \in L^\infty$. For the classical case $c(x,y) = -\langle x, y\rangle$, explicit bounds were found by Forzani and Maldonado [24], depending on $\log \frac{f^+(x)}{f^-(y)} \in L^\infty$ [8] [56].

Our proof introduces at least three significant new tools. Its starting point is that condition $\mathbf{(A3)_w}$ allows one to add a null Lagrangian term to the cost function and exploit diffeomorphism (i.e. gauge) invariance to choose coordinates, which depend on the point of interest, to transform the c-convex functions to level-set convex functions (see Theorem 4.3); this observation was made also by Liu [41], independently from us. Next we establish Alexandrov type estimates (Theorems 6.2 and 6.11) for c-convex functions whose level sets are convex, extending classical estimates for convex functions. These rely on quantitative new aspects of the geometry of convex sets that we derive elsewhere, and which may have independent interest [19]. The resulting Alexandrov type estimates enable us to systematically exploit Caffarelli's approach [5] [7] [8] to prove continuity and injectivity (see Theorems 8.1 and 8.3). Once such results are established, the same estimates permit us to exploit Forzani and Maldonado's [24] approach to extend the engulfing property of Gutierrez and Huang [30] to $\mathbf{(A3)_w}$ c-convex functions (see Theorem 9.3), improving mere continuity of optimal maps to Hölder continuity (Theorem 9.5 and its Corollary 9.6). Along the way, we also have to overcome another serious difficulty, namely the fact that the domain of the cost function (where it is smooth and satisfies appropriate cross-curvature conditions) may not be the whole of $\mathbf{R}^n$. (This situation arises, for example, when optimal transportation occurs between domains in Riemannian manifolds for the distance squared cost or similar type.) This is handled by using Theorem 5.1, where it is first established that optimal transport does not send interior points to boundary points, and vice versa, under the strong c-convexity hypothesis $\mathbf{(B2)_s}$ described in the next section. Theorem 5.1 does not require $\mathbf{(A3)_w}$, however. Let us point out that, in two dimensions, there is an alternate approach to establishing continuity of optimal maps; it was carried out by Figalli and Loeper [20] following Alexandrov's strategy [2], but their result relies on our Theorem 5.1. We may also mention that an alternate approach to Theorem 6.2 was discovered by Vétois [54], inspired by the 2009 draft of the present manuscript; he relaxed the hypothesis of that theorem from **(B4)** below to $\mathbf{(A3)_w}$ independently of us.

2 Main result

Let us begin by formulating the relevant hypothesis on the cost function $c(x,y)$ in a slightly different format than Ma, Trudinger and Wang [46] [51]. We denote their condition $\mathbf{(A3)_w}$ as **(B3)** below, and their stronger condition **(A3)** as $\mathbf{(B3)_s}$. Throughout the paper, D_y will denote the derivative with respect to the variable y, and iterated subscripts as in D^2_{xy} denote iterated derivatives. For each $(\tilde{x}, \tilde{y}) \in \overline{U} \times \overline{V}$ we define the following conditions:

(B0) $U \subset \mathbf{R}^n$ and $V \subset \mathbf{R}^n$ are open and bounded and $c \in C^4\big(\overline{U} \times \overline{V}\big)$;

(B1) (bi-twist) $\left.\begin{array}{l} x \in \overline{U} \longmapsto -D_y c(x, \tilde{y}) \\ y \in \overline{V} \longmapsto -D_x c(\tilde{x}, y) \end{array}\right\}$ are diffeomorphisms onto their ranges;

(B2) (bi-convex) $\left.\begin{array}{l} U_{\tilde{y}} := -D_y c(U, \tilde{y}) \\ V_{\tilde{x}} := -D_x c(\tilde{x}, V) \end{array}\right\}$ are convex subsets of $\mathbf{R}^n$;

(B3) (=$\mathbf{(A3)_w}$) for every curve $t \in [-1,1] \longmapsto \big(D_y c(x(t), y(0)), D_x c(x(0), y(t))\big) \in \mathbf{R}^{2n}$ which is an affinely parameterized line segment,

$$\mathrm{cross}_{(x(0),y(0))}[x'(0), y'(0)] := -\frac{\partial^4}{\partial s^2 \partial t^2}\bigg|_{(s,t)=(0,0)} c(x(s), y(t)) \geq 0 \tag{2.1}$$

provided

$$\frac{\partial^2}{\partial s \partial t}\bigg|_{(s,t)=(0,0)} c(x(s), y(t)) = 0. \tag{2.2}$$

From time to time we may strengthen these hypotheses by writing either:

$\mathbf{(B2)_s}$ if the convex domains $U_{\tilde{y}}$ and $V_{\tilde{x}}$ in **(B2)** are strongly convex;
$\mathbf{(B3)_s}$ (=**(A3)**) if, in the condition **(B3)**, the inequality (2.1) is strict when $x'(0) \neq \mathbf{0} \neq y'(0)$;
(B4) if, in the condition **(B3)**, (2.1) holds even in the absence of the extra assumption (2.2).

Here a convex set $Q \subset \mathbf{R}^n$ is said to be *strongly* convex if there exists a radius $R < +\infty$ (depending only on Q), such that each boundary point $\tilde{x} \in \partial Q$ can be touched from outside by a sphere of radius R enclosing Q; i.e. $Q \subset B_R(\tilde{x} - R\hat{n}_Q(\tilde{x}))$ where $\hat{n}_Q(\tilde{x})$ is an outer unit normal to a hyperplane supporting Q at $\tilde{x}$. When Q is smooth, this means all principal curvatures of its boundary are bounded below by $1/R$.

Hereafter $\overline{U}$ denotes the closure of U, $\mathrm{int}\, U$ denotes its interior, $\mathrm{diam}\, U$ its diameter, and for any measure $\mu^+ \geq 0$ on $\overline{U}$, we use the term *support* and the notation $\mathrm{spt}\, \mu^+ \subset \overline{U}$ to refer to the smallest closed set carrying the full mass of μ^+.

Condition $\mathbf{(A3)_w}$ (=**(B3)**) was used by Trudinger and Wang to show smoothness of optimal maps in the Monge transportation problem (1.1) when the densities are smooth. Necessity of Trudinger and Wang's condition for continuity was shown by Loeper [43], who noted its covariance (as did [36] [50]) and some relations to curvature. Their condition relaxes the hypothesis **(A3)** (=$\mathbf{(B3)_s}$) proposed earlier with Ma [46]. In [36], Kim and McCann showed that the expressions (2.1) and (2.2) correspond to pseudo-Riemannian sectional curvature conditions induced by the cost c on $U \times V$, highlighting their invariance under reparametrization of either U or V by diffeomorphism; see [36, Lemma 4.5]; see also [38] as well as [31], for further investigation of the pseudo-Riemannian aspects of optimal maps. The convexity of $U_{\tilde{y}}$ required in **(B2)** is called *c-convexity of U with respect to $\tilde{y}$* by Ma, Trudinger and Wang (or strong c-convexity if $\mathbf{(B2)_s}$ holds); they call curves $x(s) \in U$, for which $s \in [0,1] \longmapsto -D_y c(x(s), \tilde{y})$ is a line segment, *c-segments with respect to $\tilde{y}$*. Similarly,

V is said to be strongly c^*-convex with respect to $\tilde{x}$ — or with respect to $\overline{U}$ when it holds for all $\tilde{x} \in \overline{U}$ — and the curve $y(t)$ from (2.1) is said to be a c^*-segment with respect to $\tilde{x}$. Such curves correspond to geodesics $(x(t), \tilde{y})$ and $(\tilde{x}, y(t))$ in the geometry of Kim and McCann. Here and throughout, *line segments* are always presumed to be affinely parameterized.

We are now in a position to summarize our main result:

Theorem 2.1 (Interior Hölder continuity and injectivity of optimal maps)**.** *Let $c \in C^4\big(\overline{U} \times \overline{V}\big)$ satisfy* **(B0)**–**(B3)** *and* $\mathbf{(B2)_s}$*. Fix probability densities $f^+ \in L^1\big(U\big)$ and $f^- \in L^1\big(V\big)$ with $(f^+/f^-) \in L^\infty\big(U \times V\big)$ and set $d\mu^\pm := f^\pm d\mathscr{L}^n$. (Note that* $\operatorname{spt}\mu^+$ *may not be c-convex.) If the ratio $(f^-/f^+) \in L^\infty(U' \times V)$ for some open set $U' \subset U$ (U' is not necessarily c-convex), then the minimum (1.1) is attained by a map $G : \overline{U} \longmapsto \overline{V}$ whose restriction to U' is locally Hölder continuous and one-to-one. Moreover, the Hölder exponent depends only on n and $\|\log(f^+/f^-)\|_{L^\infty(U' \times V)}$.*

Proof. As recalled below in Section 3 (or see e.g. [55]) it is well-known by Kantorovich duality that the optimal joint measure $\gamma \in \Gamma(\mu^+, \mu^-)$ from (3.1) vanishes outside the c-subdifferential (3.3) of a potential $u = u^{c^*c}$ satisfying the c-convexity hypothesis (3.2), and that the map $G : \overline{U} \longmapsto \overline{V}$ which we seek is uniquely recovered from this potential using the diffeomorphism **(B1)** to solve (3.5). Thus the Hölder continuity claimed in Theorem 2.1 is equivalent to $u \in C^{1,\alpha}_{loc}(U')$.

Since $\mu^\pm$ do not charge the boundaries of U (or of V), Lemma 3.1(e) shows the c-Monge-Ampère measure defined in (3.6) has density satisfying $|\partial^c u| \leq \|f^+/f^-\|_{L^\infty(U\times V)}$ on $\overline{U}$ and $\|f^-/f^+\|^{-1}_{L^\infty(U'\times V)} \leq |\partial^c u| \leq \|f^+/f^-\|_{L^\infty(U'\times V)}$ on U'. Thus $u \in C^{1,\alpha}_{loc}(U')$ according to Theorem 9.5. Injectivity of G follows from Theorem 8.1, and the fact that the graph of G is contained in the set $\partial^c u \subset \overline{U} \times \overline{V}$ of (3.3). The dependency of the Hölder exponent α only on n and $\|\log(f^+/f^-)\|_{L^\infty(U'\times V)}$ follows by Corollary 9.6. □

Note that in case $f^+ \in C_c(U)$ is continuous and compactly supported, choosing $U' = U'_\varepsilon = \{f^+ > \varepsilon\}$ for all $\varepsilon > 0$, yields local Hölder continuity and injectivity of the optimal map $y = G(x)$ throughout U'_0.

Theorem 2.1 allows to extend the higher interior regularity results established by Liu, Trudinger and Wang in [42], originally given for **(A3)** costs, to the weaker and degenerate case $\mathbf{(A3)_w}$, see [42, Remark 4.1]. Note that these interior regularity results can be applied to manifolds, after getting suitable stay-away-from-the-cut-locus results: this is accomplished for multiple products of round spheres in [17], to yield the first regularity result that we know for optimal maps on Riemannian manifolds which are not flat, yet have some vanishing sectional curvatures.

Let us also point out that different strengthenings of the $\mathbf{(A3)_w}$ condition have been considered in [46] [36] [37] [45] [21] [22]. In particular, one stronger condition is the so-called *non-negative cross-curvature*, which here is denoted by **(B4)**. Although not strictly needed for this paper, under the **(B4)** condition the cost exponential coordinates introduced in Section 3 allow to deduce stronger conclusions with almost no extra effort, and these results play a crucial role in the proof of the regularity of optimal maps on multiple products of spheres [17]. For this reason, we prefer to include here some of the conclusions that one can deduce when $\mathbf{(A3)_w}$ is replaced by $\mathbf{(B4)}$.

Notice that, with respect to the results available for the classical Monge-Ampère equation, we need to strengthen the c-convexity assumption $\mathbf{(B2)}$ to the strong c-convexity $\mathbf{(B2)_s}$. This is due to the fact that the cost function may not be globally defined on the whole space, and for this reason we need $\mathbf{(B2)_s}$ to show that interior points of U are mapped to interior points of V, see Theorem 5.1. In [8] Caffarelli was able to take advantage of the fact that the cost function is smooth on the whole of $\mathbf{R}^n$ to chase potentially singular behaviour to infinity. As explained in Remark 8.2 below, if $-D_y c(U, y) = \mathbf{R}^n$ for all $y \in V$, then we can reproduce Caffarelli's argument in our setting and remove the uniform c-convexity assumption on the target. In particular, this applies for instance to the case of the cost given by the Euclidean squared distance on two entire convex graphs (see [46, Section 6]). We also mention that in a very recent preprint, Guillen and Kitagawa [28] announce they are able to relax the strong c-convexity assumption of our results, at the added cost of assuming compact containment of the supports of both the source and target measures.

3 Background, notation, and preliminaries

Kantorovich discerned [33] [34] that Monge's problem (1.1) could be attacked by studying the linear programming problem

$$\min_{\gamma \in \Gamma(\mu^+, \mu^-)} \int_{\overline{U} \times \overline{V}} c(x, y)\, d\gamma(x, y). \tag{3.1}$$

Here $\Gamma(\mu^+, \mu^-)$ consists of the joint probability measures on $\overline{U} \times \overline{V} \subset \mathbf{R}^n \times \mathbf{R}^n$ having $\mu^\pm$ for marginals. According to the duality theorem from linear programming, the optimizing measures γ vanish outside the zero set of $u(x) + v(y) + c(x, y) \geq 0$ for some pair of functions $(u, v) = (v^c, u^{c^*})$ satisfying

$$v^c(x) := \sup_{y \in \overline{V}} -c(x, y) - v(y), \qquad u^{c^*}(y) := \sup_{x \in \overline{U}} -c(x, y) - u(x); \tag{3.2}$$

these arise as optimizers of the dual program. This zero set is called the c-subdifferential of u, and denoted by

$$\partial^c u = \Big\{(x, y) \in \overline{U} \times \overline{V} \mid u(x) + u^{c^*}(y) + c(x, y) = 0\Big\}; \tag{3.3}$$

we also write $\partial^c u(x) := \{y \mid (x, y) \in \partial^c u\}$, and $\partial^{c^*} u^{c^*}(y) := \{x \mid (x, y) \in \partial^c u\}$, and $\partial^c u(X) := \cup_{x \in X} \partial^c u(x)$ for $X \subset \mathbf{R}^n$. Formula (3.2) defines a generalized Legendre-Fenchel transform called the *c-transform*; any function satisfying $u = u^{c^* c} := (u^{c^*})^c$ is said to be *c-convex*, which reduces to ordinary convexity in the case of the cost $c(x, y) = -\langle x, y\rangle$. In that case $\partial^c u$ reduces to the ordinary subdifferential ∂u of the convex function u, but more generally we define

$$\partial u := \{(x, p) \in \overline{U} \times \mathbf{R}^n \mid u(\tilde{x}) \geq u(x) + \langle p, \tilde{x} - x\rangle + o(|\tilde{x} - x|) \text{ as } \tilde{x} \to x\}, \tag{3.4}$$

$\partial u(x) := \{p \mid (x,p) \in \partial u\}$, and $\partial u(X) := \cup_{x\in X} \partial u(x)$. Assuming $c \in C^2\big(\overline{U} \times \overline{V}\big)$ (which is the case if **(B0)** holds), any c-convex function $u = u^{c^*c}$ will be semi-convex, meaning its Hessian admits a bound from below $D^2 u \geq -\|c\|_{C^2}$ in the distributional sense; equivalently, $u(x) + \|c\|_{C^2}|x|^2/2$ is convex on each ball in U [26]. In particular, u will be twice-differentiable $\mathscr{L}^n$-a.e. on U in the sense of Alexandrov.

As in [25] [40] [46], hypothesis **(B1)** shows the map $G : \operatorname{dom} Du \longmapsto \overline{V}$ is uniquely defined on the set $\operatorname{dom} Du \subset \overline{U}$ of differentiability for u by

$$D_x c(\tilde{x}, G(\tilde{x})) = -Du(\tilde{x}). \tag{3.5}$$

The graph of G, so-defined, lies in $\partial^c u$. The task at hand is to show (local) Hölder continuity and injectivity of G — the former being equivalent to $u \in C^{1,\alpha}_{loc}(U)$ — by studying the relation $\partial^c u \subset \overline{U} \times \overline{V}$.

To this end, as already done in [46], we define a Borel measure $|\partial^c u|$ on $\mathbf{R}^n$ associated to u by

$$|\partial^c u|(X) := \mathscr{L}^n(\partial^c u(X)) \tag{3.6}$$

for each Borel set $X \subset \mathbf{R}^n$; it will be called the *c-Monge-Ampère measure* of u. (Similarly, we define $|\partial u|$.) We use the notation $|\partial^c u| \geq \lambda$ on U' as a shorthand to indicate $|\partial^c u|(X) \geq \lambda \mathscr{L}^n(X)$ for each $X \subset U'$; similarly, $|\partial^c u| \leq \Lambda$ indicates $|\partial^c u|(X) \leq \Lambda \mathscr{L}^n(X)$. As the next lemma shows, uniform bounds above and below on the marginal densities of a probability measure γ vanishing outside $\partial^c u$ imply similar bounds on $|\partial^c u|$.

Lemma 3.1 (Properties of c-Monge-Ampère measures)**.** *Let c satisfy* **(B0)**-**(B1)**, *while u and u_k denote c-convex functions for each $k \in \mathbf{N}$. Fix $\tilde{x} \in \overline{X}$ and constants $\lambda, \Lambda > 0$.*
(a) Then $\partial^c u(\overline{U}) \subset \overline{V}$ and $|\partial^c u|$ is a Borel measure of total mass $\mathscr{L}^n\big(\overline{V}\big)$ on $\overline{U}$.
(b) If $u_k \to u_\infty$ uniformly, then u_∞ is c-convex and $|\partial^c u_k| \rightharpoonup |\partial^c u_\infty|$ weakly-$*$ in the duality against continuous functions on $\overline{U} \times \overline{V}$.
(c) If $u_k(\tilde{x}) = 0$ for all k, then the functions u_k converge uniformly if and only if the measures $|\partial^c u_k|$ converge weakly-$*$.
(d) If $|\partial^c u| \leq \Lambda$ on $\overline{U}$, then $|\partial^{c^*} u^{c^*}| \geq 1/\Lambda$ on $\overline{V}$.
(e) If a probability measure $\gamma \geq 0$ vanishes outside $\partial^c u \subset \overline{U} \times \overline{V}$, and has marginal densities $f^\pm$, then $f^+ \geq \lambda$ on $U' \subset \overline{U}$ and $f^- \leq \Lambda$ on $\overline{V}$ imply $|\partial^c u| \geq \lambda/\Lambda$ on U', whereas $f^+ \leq \Lambda$ on U' and $f^- \geq \lambda$ on $\overline{V}$ imply $|\partial^c u| \leq \Lambda/\lambda$ on U'.

Proof. (a) The fact $\partial^c u(\overline{U}) \subset \overline{V}$ is an immediate consequence of definition (3.3). Since $c \in C^1(\overline{U} \times \overline{V})$, the c-transform $v = u^{c^*} : \overline{V} \longmapsto \mathbf{R}$ defined by (3.2) can be extended to a Lipschitz function on a neighbourhood of $\overline{V}$, hence Rademacher's theorem asserts $\operatorname{dom} Dv$ is a set of full Lebesgue measure in $\overline{V}$. Use **(B1)** to define the unique solution $F : \operatorname{dom} Dv \longmapsto \overline{U}$ to

$$D_y c(F(\tilde{y}), \tilde{y}) = -Dv(\tilde{y}).$$

As in [25] [40], the vanishing of $u(x) + v(y) + c(x,y) \geq 0$ implies $\partial^{c^*} v(\tilde{y}) = \{F(\tilde{y})\}$, at least for all points $\tilde{y} \in \operatorname{dom} Dv$ where $\overline{V}$ has Lebesgue density greater than one half. For Borel $X \subset \mathbf{R}^n$, this shows $\partial^c u(X)$ differs from the Borel set $F^{-1}(X) \cap \overline{V}$ by a $\mathscr{L}^n$ negligible subset of $\overline{V}$, whence $|\partial^c u| = F_\#\big(\mathscr{L}^n\lfloor_{\overline{V}}\big)$ so claim (a) of the lemma is established.

(b) Let $\|u_k - u_\infty\|_{L^\infty(\overline{U})} \to 0$. It is not hard to deduce c-convexity of u_∞, as in e.g. [18]. Define $v_k = u_k^{c^*}$ and F_k on $\operatorname{dom} Dv_k \subset \overline{V}$ as above, so that $|\partial^c u_k| = F_{k\#}\big(\mathscr{L}^n\lfloor_{\overline{V}}\big)$. Moreover, $v_k \to v_\infty$ in $L^\infty(V)$, where v_∞ is the c^*-dual to u_∞. The uniform semiconvexity of v_k (i.e. convexity of $v_k(y) + \frac{1}{2}\|c\|_{C^2}|y|^2$) ensures pointwise convergence of $Dv_k \to Dv_\infty$ $\mathscr{L}^n$-a.e. on $\overline{V}$. From $D_y c(F_k(\tilde{y}), \tilde{y}) = -Dv_k(\tilde{y})$ we deduce $F_k \to F_\infty$ $\mathscr{L}^n$-a.e. on $\overline{V}$. This is enough to conclude $|\partial^c u_k| \rightharpoonup |\partial^c u_k|$, by testing the convergence against continuous functions and applying Lebesgue's dominated convergence theorem.

(c) To prove the converse, suppose u_k is a sequence of c-convex functions which vanish at $\tilde{x}$ and $|\partial^c u_k| \rightharpoonup \mu_\infty$ weakly-$*$. Since the u_k have Lipschitz constants dominated by $\|c\|_{C^1}$ and $\overline{U}$ is compact, any subsequence of the u_k admits a convergent further subsequence by the Ascoli-Arzelà Theorem. A priori, the limit u_∞ might depend on the subsequences, but (b) guarantees $|\partial^c u_\infty| = \mu_\infty$, after which [43, Proposition 4.1] identifies u_∞ uniquely in terms of $\mu^+ = \mu_\infty$ and $\mu^- = \mathscr{L}^n\lfloor_{\overline{V}}$, up to an additive constant; this arbitrary additive constant is fixed by the condition $u_\infty(\tilde{x}) = 0$. Thus the whole sequence u_k converges uniformly.

(e) Now assume a finite measure $\gamma \geq 0$ vanishes outside $\partial^c u$ and has marginal densities $f^\pm$. Then the second marginal $d\mu^- := f^- d\mathscr{L}^n$ of γ is absolutely continuous with respect to Lebesgue and γ vanishes outside the graph of $F : \overline{V} \longmapsto U$, whence $\gamma = (F \times id)_\# \mu^-$ by e.g. [1, Lemma 2.1]. (Here id denotes the identity map, restricted to the domain $\operatorname{dom} Dv$ of definition of F.) Recalling that $|\partial^c u| = F_\#\big(\mathscr{L}^n\lfloor_{\overline{V}}\big)$ (see the proof of (a) above), for any Borel $X \subset U'$ we have

$$\lambda|\partial^c u|(X) = \lambda\mathscr{L}^n(F^{-1}(X)) \leq \int_{F^{-1}(X)} f^-(y) d\mathscr{L}^n(y) = \int_X f^+(x) d\mathscr{L}^n(x) \leq \Lambda\mathscr{L}^n(X)$$

whenever $\lambda \leq f^-$ and $f^+ \leq \Lambda$. We can also reverse the last four inequalities and interchange λ with Λ to establish claim (e) of the lemma.

(d) The last point remaining follows from (e) by taking $\gamma = (F \times id)_\# \mathscr{L}^n$. Indeed an upper bound λ on $|\partial^c u| = F_\# \mathscr{L}^n$ throughout $\overline{U}$ and lower bound 1 on $\mathscr{L}^n$ translate into a lower bound $1/\lambda$ on $|\partial^{c^*} u^{c^*}|$, since the reflection γ^* defined by $\gamma^*(Y \times X) := \gamma(X \times Y)$ for each $X \times Y \subset U \times V$ vanishes outside $\partial^{c^*} u^{c^*}$ and has second marginal absolutely continuous with respect to Lebesgue by the hypothesis $|\partial^c u| \leq \lambda$. □

Remark 3.2 (Monge-Ampère type equation)**.** Differentiating (3.5) formally with respect to $\tilde{x}$ and recalling $|\det DG(\tilde{x})| = f^+(\tilde{x})/f^-(G(\tilde{x}))$ yields the Monge-Ampère type equation

$$\frac{\det[D^2_{xx} u(\tilde{x}) + D^2_{xx} c(\tilde{x}, G(\tilde{x}))]}{|\det D^2_{xy} c(\tilde{x}, G(\tilde{x}))|} = \frac{f^+(\tilde{x})}{f^-(G(\tilde{x}))} \tag{3.7}$$

on U, where $G(\tilde{x})$ is given as a function of $\tilde{x}$ and $Du(\tilde{x})$ by (3.5). Degenerate ellipticity follows from the fact that $y = G(x)$ produces equality in $u(x) + u^{c^*}(y) + c(x, y) \geq 0$. A condition under which c-convex weak-$*$ solutions are known to exist is given by

$$\int_{\overline{U}} f^+(x) d\mathscr{L}^n(x) = \int_{\overline{V}} f^-(y) d\mathscr{L}^n(y).$$

The boundary condition $\partial^c u(\overline{U}) \subset \overline{V}$ which then guarantees Du to be uniquely determined f^+-a.e. is built into our definition of c-convexity of u. In fact, [43, Proposition 4.1] shows u to be uniquely determined up to additive constant if either $f^+ > 0$ or $f^- > 0$ $\mathscr{L}^n$-a.e. on its connected domain, U or V.

A key result we shall exploit several times is a maximum principle first deduced from Trudinger and Wang's work [51] by Loeper; see [43, Theorem 3.2]. A simple and direct proof, and also an extension can be found in [36, Theorem 4.10], where the principle was also called 'double-mountain above sliding-mountain' (**DASM**). Other proofs and extensions appear in [52] [55] [45] [21]:

Theorem 3.3 (Loeper's maximum principle '**DASM**'). *Assume* **(B0)**–**(B3)** *and fix* $x, \tilde{x} \in \overline{U}$. *If* $t \in [0,1] \longmapsto -D_x c(\tilde{x}, y(t))$ *is a line segment then* $f(t) := -c(x, y(t)) + c(\tilde{x}, y(t)) \leq \max\{f(0), f(1)\}$ *for all* $t \in [0,1]$.

It is through this theorem and the next that hypothesis **(B3)** and the non-negative cross-curvature hypothesis **(B4)** enter crucially. Among the many corollaries Loeper deduced from this result, we shall need two. Proved in [43, Theorem 3.1 and Proposition 4.4] (alternately [36, Theorem 3.1] and [35, A.10]), they include the c-convexity of the so-called *contact set* (meaning the c^*-subdifferential at a point), and a local to global principle.

Corollary 3.4. *Assume* **(B0)**–**(B3)** *and fix* $(\tilde{x}, \tilde{y}) \in \overline{U} \times \overline{V}$. *If* u *is c-convex then* $\partial^c u(\tilde{x})$ *is* c^**-convex with respect to* $\tilde{x} \in U$, *i.e.* $-D_x c(\tilde{x}, \partial^c u(\tilde{x}))$ *forms a convex subset of* $T^*_{\tilde{x}} U$. *Furthermore, any local minimum of the map* $x \in U \longmapsto u(x) + c(x, \tilde{y})$ *is a global minimum.*

As shown in [37, Corollary 2.11], the strengthening **(B4)** of hypothesis **(B3)** improves the conclusion of Loeper's maximum principle. This improvement asserts that the altitude $f(t)$ at each point of the evolving landscape then accelerates as a function of $t \in [0,1]$:

Theorem 3.5 (**Time-convex DASM**). *Assume* **(B0)**–**(B4)** *and fix* $x, \tilde{x} \in \overline{U}$. *If* $t \in [0,1] \longmapsto -D_x c(\tilde{x}, y(t))$ *is a line segment then the function* $t \in [0,1] \longmapsto f(t) := -c(x, y(t)) + c(\tilde{x}, y(t))$ *is convex.*

Remark 3.6. Since all assumptions **(B0)**–**(B4)** on the cost are symmetric in x and y, all the results above still hold when the roles of x and y are exchanged.

c-Monge-Ampère equation

Fix $\lambda, \Lambda > 0$ and an open domain $U^\lambda \subset U$, and let u be a c-convex solution of the c-Monge-Ampère equation

$$\begin{cases} \lambda \mathscr{L}^n \leq |\partial^c u| \leq \frac{1}{\lambda} \mathscr{L}^n & \text{in } U^\lambda \subset U, \\ |\partial^c u| \leq \Lambda \mathscr{L}^n & \text{in } \overline{U}. \end{cases} \tag{3.8}$$

Note that throughout this paper, we require c-convexity **(B2)** not of U^λ but only of U. We sometimes abbreviate (3.8) by writing $|\partial^c u| \in [\lambda, 1/\lambda]$. In the following sections we will prove interior Hölder differentiability of u on U^λ, that is $u \in C^{1,\alpha}_{loc}(U^\lambda)$; see Theorems 8.3 and 9.5.

Convex sets

We close by recalling two nontrivial results for convex sets. These will be essential in Section 6 and later on. The first one is due to Fritz John [32]:

Lemma 3.7 (John's lemma)**.** *For a compact convex set $Q \subset \mathbf{R}^n$ with nonempty interior, there exists an affine transformation $L : \mathbf{R}^n \to \mathbf{R}^n$ such that $\overline{B_1} \subset L^{-1}(Q) \subset \overline{B_n}$.*

The above result can be restated by saying that any compact convex set Q with nonempty interior contains an ellipsoid E, whose dilation nE by factor n with respect to its center contains Q:

$$E \subset Q \subset nE. \tag{3.9}$$

The following is our main result from [19]; it enters crucially in the proof of Theorem 6.11.

Theorem 3.8 (Convex bodies and supporting hyperplanes)**.** *Let $\tilde{Q} \subset \mathbf{R}^n$ be a well-centered convex body, meaning that* (3.9) *holds for some ellipsoid E centered at the origin. Fix $0 \leq s \leq s_0 < 1$. To each $y \in (1-s)\partial\tilde{Q}$ corresponds at least one line ℓ through the origin and hyperplane Π supporting $\tilde{Q}$ such that: Π is orthogonal to ℓ and*

$$\operatorname{dist}(y, \Pi) \leq c(n, s_0) s^{1/2^{n-1}} \operatorname{diam}(\ell \cap \tilde{Q}). \tag{3.10}$$

Here, $c(n, s_0)$ is a constant depending only on n and s_0, namely $c(n, s_0) = n^{3/2}(n-\frac{1}{2})\left(\frac{1+(s_0)^{1/2^n}}{1-(s_0)^{1/2^n}}\right)^{n-1}$.

4 Choosing coordinates which "level-set convexify" c-convex functions

Recall that $c \in C^4(\overline{U} \times \overline{V})$ is a cost function satisfying **(B1)**–**(B3)** on a pair of bounded domains U and V which are strongly c-convex with respect to each other $\mathbf{(B2)_s}$. In the current section, we introduce an important transformation (mixing dependent and independent variables) for the cost $c(x, y)$ and potential $u(x)$, which plays a crucial role in the subsequent analysis. This change of variables and its most relevant properties are encapsulated in the following Definition 4.1 and Theorem 4.3.

Definition 4.1 (Cost-exponential coordinates and apparent properties)**.** *Given $c \in C^4\big(\overline{U} \times \overline{V}\big)$ strongly twisted* **(B0)**–**(B1)***, we refer to the coordinates $(q, p) \in \overline{U}_{\tilde{y}} \times \overline{V}_{\tilde{x}}$ defined by*

$$q = q(x) = -D_y c(x, \tilde{y}), \qquad p = p(y) = -D_x c(\tilde{x}, y), \tag{4.1}$$

as the cost exponential coordinates *from $\tilde{y} \in \overline{V}$ and $\tilde{x} \in \overline{U}$ respectively. We denote the inverse diffeomorphisms by $x : \overline{U}_{\tilde{y}} \subset T^*_{\tilde{y}}V \longmapsto \overline{U}$ and $y : \overline{V}_{\tilde{x}} \subset T^*_{\tilde{x}}U \longmapsto \overline{V}$; they satisfy*

$$q = -D_y c(x(q), \tilde{y}), \qquad p = -D_x c(\tilde{x}, y(p)). \tag{4.2}$$

The cost $\tilde{c}(q, y) = c(x(q), y) - c(x(q), \tilde{y})$ is called the modified cost *at $\tilde{y}$. A subset of $\overline{U}$ or function thereon is said to* appear *from $\tilde{y}$ to have property A, if it has property A when expressed in the coordinates $q \in \overline{U}_{\tilde{y}}$.*

Remark 4.2. Identifying the cotangent vector $0 \oplus q$ with the tangent vector $q^* \oplus 0$ to $U \times V$ using the pseudo-metric of Kim and McCann [36] shows $x(q)$ to be the projection to U of the pseudo-Riemannian exponential map $\exp_{(\tilde{x},\tilde{y})}(q^* \oplus 0)$; similarly $y(p)$ is the projection to V of $\exp_{(\tilde{x},\tilde{y})}(0 \oplus p^*)$. Also, $x(q) =: c^*\text{-}\exp_{\tilde{y}} q$ and $y(p) =: c\text{-}\exp_{\tilde{x}} p$ in the notation of Loeper [43].

In the sequel, whenever we use the expression $\tilde{c}(q,\cdot)$ or $\tilde{u}(q)$, we refer to the modified cost function defined above and level-set convex potential defined below. Since properties **(B0)**–**(B4)** (and $\mathbf{(B2)_s}$) were shown to be tensorial in nature (i.e. coordinate independent) in [36] [43], the modified cost $\tilde{c}$ inherits these properties from the original cost c with one exception: (4.2) defines a C^3 diffeomorphism $q \in \overline{U}_{\tilde{y}} \longmapsto x(q) \in \overline{U}$, so the cost $\tilde{c} \in C^3(\overline{U}_{\tilde{y}} \times \overline{V})$ may not be C^4 smooth. However, its definition reveals that we may still differentiate $\tilde{c}$ four times as long as no more than three of the four derivatives fall on the variable q, and it leads to the same geometrical structure (pseudo-Riemannian curvatures, including (2.1)) as the original cost c since the metric tensor and symplectic form defined in [36] involve only mixed derivatives $D^2_{qy}\tilde{c}$, and therefore remain C^2 functions of the coordinates $(q,y) \in \overline{U}_{\tilde{y}} \times \overline{V}$.

We also use

$$\beta_c^{\pm} \;=\; \beta_c^{\pm}(U \times V) := \;\|(D^2_{xy}c)^{\pm 1}\|_{L^\infty(U\times V)} \tag{4.3}$$

$$\gamma_c^{\pm} \;=\; \gamma_c^{\pm}(U \times V) := \|\det(D^2_{xy}c)^{\pm 1}\|_{L^\infty(U\times V)} \tag{4.4}$$

to denote the bi-Lipschitz constants $\beta_c^{\pm}$ of the coordinate changes (4.1) and the Jacobian bounds $\gamma_c^{\pm}$ for the same transformation. Notice $\gamma_c^+\gamma_c^- \geq 1$ for any cost satisfying **(B1)**, and equality holds whenever the cost function $c(x,y)$ is quadratic. So the parameter $\gamma_c^+\gamma_c^-$ crudely quantifies the departure from the quadratic case. The inequality $\beta_c^+\beta_c^- \geq 1$ is much more rigid, equality implying $D^2_{xy}c(x,y)$ is the identity matrix, and not merely constant.

Our first contribution is the following theorem. If the cost function satisfies **(B3)**, then the level sets of the $\tilde{c}$-convex potential appear convex from $\tilde{y}$, as was discovered independently from us by Liu [41], and exploited by Liu with Trudinger and Wang [42]. Moreover, for a non-negatively cross-curved cost **(B4)**, it shows that any $\tilde{c}$-convex potential appears convex from $\tilde{y} \in \overline{V}$. Note that although the difference between the cost $c(x,y)$ and the modified cost $\tilde{c}(q,y)$ depends on $\tilde{y}$, they differ by a null Lagrangian $c(x,\tilde{y})$ which — being independent of $y \in V$ — does not affect the question of which maps G attain the infimum (1.1). Having a function with convex level sets is a useful starting point, since it opens a possibility to apply the approach and techniques developed by Caffarelli, and refined by Gutieérrez, Forzani, Maldonado and others (see [29] [24]), to address the regularity of c-convex potentials.

Theorem 4.3 (Modified c-convex functions appear level-set convex)**.** *Let $c \in C^4\big(\overline{U} \times \overline{V}\big)$ satisfy* **(B0)**–**(B3)**. *If $u = u^{c^*c}$ is c-convex on $\overline{U}$, then $\tilde{u}(q) = u(x(q)) + c(x(q),\tilde{y})$ has convex level sets, as a function of the cost exponential coordinates $q \in \overline{U}_{\tilde{y}}$ from $\tilde{y} \in \overline{V}$. Moreover,*

$$\tilde{u} + M_c|q|^2 \quad \textit{is convex,} \tag{4.5}$$

where $M_c := (\beta_c^-)^2\|D^2_{xx}c\|_{L^\infty(U\times V)} + (\beta_c^-)^3\|D_x c\|_{L^\infty(U\times V)}\|D^2_{xx}D_y c\|_{L^\infty(U\times V)}$. If, in addition, c is non-negatively cross-curved **(B4)** *then $\tilde{u}$ is convex on $\overline{U}_{\tilde{y}}$. In either case $\tilde{u}$ is minimized* *at q_0 if $\tilde y \in \partial^c u(x(q_0))$. Furthermore, $\tilde u$ is $\tilde c$-convex with respect to the modified cost $\tilde c(q,y) := c(x(q),y) - c(x(q),\tilde y)$ on $\overline{U}_{\tilde y} \times \overline{V}$, and $\partial^{\tilde c}\tilde u(q) = \partial^c u(x(q))$ for all $q \in \overline{U}_{\tilde y}$.*

Proof. The final sentences of the theorem are elementary: c-convexity $u = u^{c^*c}$ asserts

$$u(x) = \sup_{y\in\overline{V}} -c(x,y) - u^{c^*}(y) \quad \text{and} \quad u^{c^*}(y) = \sup_{q\in\overline{U}_{\tilde y}} -c(x(q),y) - u(x(q)) = \tilde u^{\tilde c^*}(y)$$

from (3.2), hence

$$\begin{aligned} \tilde u(q) &= \sup_{y\in\overline{V}} -c(x(q),y) + c(x(q),\tilde y) - u^{c^*}(y) \\ &= \sup_{y\in\overline{V}} -\tilde c(q,y) - \tilde u^{\tilde c^*}(y), \end{aligned}$$

and $\partial^{\tilde c}\tilde u(q) = \partial^c u(x(q))$ since all three suprema above are attained at the same $y \in \overline{V}$. Taking $y = \tilde y$ reduces the inequality $\tilde u(q) + \tilde u^{\tilde c^*}(y) + \tilde c(q,y) \geq 0$ to $\tilde u(q) \geq -\tilde u^{\tilde c^*}(\tilde y)$, with equality precisely if $\tilde y \in \partial^{\tilde c}\tilde u(q)$. It remains to address the convexity claims.

Since the supremum $\tilde u(q)$ of a family of convex functions is again convex, it suffices to establish the convexity of $q \in \overline{U}_{\tilde y} \longmapsto -\tilde c(q,y)$ for each $y \in \overline{V}$ under hypothesis **(B4)**. For a similar reason, it suffices to establish the level-set convexity of the same family of functions under hypothesis **(B3)**.

First assume **(B3)**. Since

$$D_y\tilde c(q,\tilde y) = D_y c(x(q),\tilde y) := -q \tag{4.6}$$

we see that $\tilde c$-segments in $\overline{U}_{\tilde y}$ with respect to $\tilde y$ coincide with ordinary line segments. Let $q(s) = (1-s)q_0 + sq_1$ be any line segment in the convex set $\overline{U}_{\tilde y}$. Define $f(s,y) := -\tilde c(q(s),y) = -c(x(q(s)),y) + c(x(q(s)),\tilde y)$. Loeper's maximum principle (Theorem 3.3 above, see also Remark 3.6) asserts $f(s,y) \leq \max\{f(0,y), f(1,y)\}$, which implies convexity of each set $\{q \in \overline{U}_{\tilde y} \mid -\tilde c(q,y) \leq const\}$. Under hypothesis **(B4)**, Theorem 3.5 goes on to assert convexity of $s \in [0,1] \longmapsto f(s,y)$ as desired.

Finally, (4.5) follows from the fact that $\tilde u$ is a supremum of cost functions and that $|D_{qq}\tilde c| \leq (\beta_c^-)^2|D^2_{xx}c| + (\beta_c^-)^3|D_x c||D^2_{xx}D_y c|$ (observe that $D_q x(q) = -D^2_{xy}c(x(q),\tilde y)^{-1}$). □

Remark 4.4. The above level-set convexity for the modified $\tilde c$-convex functions requires **(B3)** condition, as one can easily derive using Loeper's counterexample [43].

The effect of this change of gauge on the c-Monge-Ampère equation (3.8) is summarized in a corollary:

Corollary 4.5 (Transformed $\tilde c$-Monge-Ampère inequalities)**.** *Using the hypotheses and notation of Theorem 4.3, if $|\partial^c u| \in [\lambda,\Lambda] \subset [0,\infty]$ on $U' \subset \overline{U}$, then $|\partial^{\tilde c}\tilde u| \in [\lambda/\gamma_c^+, \Lambda\gamma_c^-]$ on $U'_{\tilde y} = -D_y c(U',\tilde y)$, where $\gamma_c^\pm = \gamma_c^\pm(U' \times V)$ and $\beta_c^\pm = \beta_c^\pm(U' \times V)$ are defined in* (4.3)–(4.4)*. Furthermore, $\gamma_{\tilde c}^\pm := \gamma_{\tilde c}^\pm(U'_{\tilde y} \times V) \leq \gamma_c^+\gamma_c^-$ and $\beta_{\tilde c}^\pm := \beta_{\tilde c}^\pm(U'_{\tilde y} \times V) \leq \beta_c^+\beta_c^-$.*

Proof. From the Jacobian bounds $|\det D_x q(x)| \in [1/\gamma_c^-, \gamma_c^+]$ on U', we find $\mathscr{L}^n(X)/\gamma_c^- \leq \mathscr{L}^n(q(X)) \leq \gamma_c^+ \mathscr{L}^n(X)$ for each $X \subset U'$. On the other hand, Theorem 4.3 asserts $\partial^{\tilde{c}}\tilde{u}(q(X)) = \partial^c u(X)$, so the claim $|\partial^{\tilde{c}}\tilde{u}| \in [\lambda/\gamma_c^+, \Lambda\gamma_c^-]$ follows from the hypothesis $|\partial^c u| \in [\lambda, \Lambda]$, by definition (3.6) and the fact that $q : \overline{U} \longrightarrow \overline{U}_{\tilde{y}}$ from (4.1) is a diffeomorphism; see **(B1)**. The bounds $\gamma_{\tilde{c}}^{\pm} \leq \gamma_c^+\gamma_c^-$ and $\beta_{\tilde{c}}^{\pm} \leq \beta_c^+\beta_c^-$ follow from $D^2_{qy}\tilde{c}(q,y) = D^2_{xy}c(x(q),y)D_q x(q)$ and $D_q x(q) = -D^2_{xy}c(x(q),\tilde{y})^{-1}$. □

Affine renormalization

We record here an observation that the $\tilde{c}$-Monge-Ampère measure is invariant under an affine renormalization. This is potentially useful in applications, though we *do not* use this fact for the results of the present paper. For an affine transformation $L : \mathbf{R}^n \to \mathbf{R}^n$, define

$$\tilde{u}^*(q) = |\det L|^{-2/n}\tilde{u}(Lq). \tag{4.7}$$

Here $\det L$ denotes the Jacobian determinant of L, i.e. the determinant of the linear part of L.

Lemma 4.6 (Affine invariance of $\tilde{c}$-Monge-Ampère measure)**.** *Assuming* **(B0)**–**(B1)***, given a $\tilde{c}$-convex function $\tilde{u} : U_{\tilde{y}} \longmapsto \mathbf{R}$ and affine bijection $L : \mathbf{R}^n \longmapsto \mathbf{R}^n$, define the renormalized potential $\tilde{u}^*$ by* (4.7) *and renormalized cost*

$$\tilde{c}_*(q,y) = |\det L|^{-2/n}\tilde{c}(Lq, L^*y) \tag{4.8}$$

using the adjoint L^* to the linear part of L. Then, for any Borel set $Q \subset \overline{U}_{\tilde{y}}$,

$$|\partial\tilde{u}^*|(L^{-1}Q) = |\det L|^{-1}|\partial\tilde{u}|(Q), \tag{4.9}$$

$$|\partial^{\tilde{c}_*}\tilde{u}^*|(L^{-1}Q) = |\det L|^{-1}|\partial^{\tilde{c}}\tilde{u}|(Q). \tag{4.10}$$

Proof. From (3.4) we see $\bar{p} \in \partial\tilde{u}(\bar{q})$ if and only if $|\det L|^{-2/n}L^*\bar{p} \in \partial\tilde{u}^*(L^{-1}\bar{q})$, thus (4.9) follows from $\partial\tilde{u}^*(L^{-1}Q) = |\det L|^{-2/n}L^*\big(\partial\tilde{u}(Q)\big)$. Similarly, since (3.2) yields $(\tilde{u}^*)^{\tilde{c}^*_*}(y) = |\det L|^{-2/n}\tilde{u}^{\tilde{c}^*}(L^*y)$, we see $\bar{y} \in \partial^{\tilde{c}}\tilde{u}(\bar{q})$ is equivalent to $|\det L|^{-2/n}L^*\bar{y} \in \partial^{\tilde{c}_*}\tilde{u}^*(L^{-1}\bar{q})$ from (3.3) (and Theorem 4.3), whence $\partial^{\tilde{c}_*}\tilde{u}^*(L^{-1}Q) = |\det L|^{-2/n}L^*\big(\partial^{\tilde{c}}\tilde{u}^*(Q)\big)$ to establish (4.10). □

As a corollary to this lemma, we recover the affine invariance not only of the Monge-Ampère equation satisfied by $\tilde{u}(q)$ — but also of the $\tilde{c}$-Monge-Ampère equation it satisfies — under coordinate changes on V (which induce linear transformations L on $T^*_{\tilde{y}}V$ and L^* on $T_{\tilde{y}}V$): for $q \in U_{\tilde{y}}$,

$$\frac{d|\partial\tilde{u}^*|}{d\mathscr{L}^n}(L^{-1}q) = \frac{d|\partial\tilde{u}|}{d\mathscr{L}^n}(q) \quad \text{and} \quad \frac{d|\partial^{\tilde{c}_*}\tilde{u}^*|}{d\mathscr{L}^n}(L^{-1}q) = \frac{d|\partial^{\tilde{c}}\tilde{u}|}{d\mathscr{L}^n}(q).$$

5 Strongly c-convex interiors and boundaries not mixed by $\partial^c u$

The subsequent sections of this paper are largely devoted to ruling out exposed points in $U_{\tilde{y}}$ of the set where the $\tilde{c}$-convex potential from Theorem 4.3 takes its minimum. This current section rules out exposed points on the boundary of $U_{\tilde{y}}$. We do this by proving an important topological property of the (multi-valued) mapping $\partial^c u \subset \overline{U} \times \overline{V}$. Namely, we show that the subdifferential $\partial^c u$ maps interior points of $\mathrm{spt}\,|\partial^c u| \subset \overline{U}$ only to interior points of V, under hypothesis (3.8), and conversely that $\partial^c u$ maps boundary points of U only to boundary points of V. This theorem may be of independent interest, and was required by Figalli and Loeper to conclude their continuity result concerning maps between *two dimensional* domains which optimize **(B3)** costs [20].

This section does not use the **(B3)** assumption on the cost function $c \in C^4(\overline{U} \times \overline{V})$, but relies crucially on the *strong* c-convexity $\mathbf{(B2)_s}$ of its domains U and V (but importantly, not of $\mathrm{spt}\,|\partial^c u|$). No analog for Theorem 5.1 was needed by Caffarelli to establish $C^{1,\alpha}$ regularity of convex potentials $u(x)$ whose gradients optimize the classical cost $c(x,y) = -\langle x, y\rangle$ [8], since in that case he was able to take advantage of the fact that the cost function is smooth on the whole of $\mathbf{R}^n$ to chase potentially singular behaviour to infinity. (One general approach to showing regularity of solutions for degenerate elliptic partial differential equations is to exploit the threshold-hyperbolic nature of the solution to try to follow either its singularities or its degeneracies to the boundary, where they can hopefully be shown to be in contradiction with boundary conditions; the *degenerate* nature of the ellipticity precludes the possibility of *purely local* regularizing effects.)

Theorem 5.1 (Strongly c-convex interiors and boundaries not mixed by $\partial^c u$)**.** *Let c satisfy* **(B0)**–**(B1)** *and $u = u^{c^*c}$ be a c-convex function (which implies $\partial^c u(\overline{U}) = \overline{V}$), and $\lambda > 0$.*

(a) If $|\partial^c u| \geq \lambda$ on $X \subset \overline{U}$ and V is strongly c^*-convex with respect to X, then interior points of X cannot be mapped by $\partial^c u$ to boundary points of V: i.e. $(X \times \partial V) \cap \partial^c u \subset (\partial X \times \partial V)$.

(b) If $|\partial^c u| \leq \Lambda$ on $\overline{U}$, and U is strongly c-convex with respect to V, then boundary points of U cannot be mapped by $\partial^c u$ into interior points of V: i.e. $\partial U \times V$ is disjoint from $\partial^c u$.

Proof. Note that when X is open the conclusion of (a) implies $\partial^c u$ is disjoint from $X \times \partial V$. We therefore remark that it suffices to prove (a), since (b) follows from (a) exchanging the role x and y and observing that $|\partial^c u| \leq \Lambda$ implies $|\partial^{c^*} u^{c^*}| \geq 1/\Lambda$ as in Lemma 3.1(d).

Let us prove (a). Fix any point $\tilde{x}$ in the interior of X, and $\tilde{y} \in \partial^c u(\tilde{x})$. Assume by contradiction that $\tilde{y} \in \partial V$. The idea of proof is summarized in Figure 1. We first fix appropriate coordinates. At $(\tilde{x}, \tilde{y})$ we use **(B0)**–**(B1)** to define cost-exponential coordinates $(p,q) \longmapsto (x(q), y(p))$ by

$$
\begin{aligned}
p &= -D_x c(\tilde{x}, y(p)) + D_x c(\tilde{x}, \tilde{y}) && \in T^*_{\tilde{x}}(U)\\
q &= D^2_{xy} c(\tilde{x}, \tilde{y})^{-1}(D_y c(x(q), \tilde{y}) - D_y c(\tilde{x}, \tilde{y})) && \in T_{\tilde{x}}(U)
\end{aligned}
$$

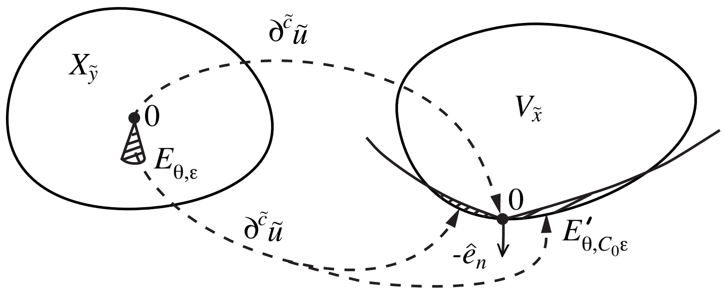

Figure 1: If $\partial^{\tilde c}\tilde u$ sends an interior point onto a boundary point, by $\tilde c$-monotonicity of $\partial^{\tilde c}\tilde u$ the small cone $E_{\theta,\varepsilon}$ (with height ε and opening θ) has to be sent onto $E'_{\theta,C_0\varepsilon}\cap V_{\tilde x}$. Since $\mathscr{L}^n(E_{\theta,\varepsilon})\sim\theta^{n-1}$ for $\varepsilon>0$ small but fixed, while $\mathscr{L}^n(E'_{\theta,C_0\varepsilon}\cap V_{\tilde x})\lesssim\theta^{n+1}$ (by the strong convexity of $\tilde V_{\tilde x}$), we get a contradiction as $\theta\to 0$.

and define a modified cost and potential by subtracting null Lagrangian terms:

$$\begin{aligned}\tilde c(q,p) &:= c(x(q),y(p))-c(x(p),\tilde y)-c(\tilde x,y(p))\\ \tilde u(q) &:= u(x(q))+c(x(q),\tilde y).\end{aligned}$$

Similarly to Corollary 4.5, $|\partial^{\tilde c}\tilde u|\geq\tilde\lambda:=\lambda/(\gamma_c^+\gamma_c^-)$, where $\gamma_c^\pm$ denote the Jacobian bounds (4.4) for the coordinate change. Note $(\tilde x,\tilde y)=(x(\mathbf{0}),y(\mathbf{0}))$ corresponds to $(p,q)=(\mathbf{0},\mathbf{0})$. Since c-segments with respect to $\tilde y$ correspond to line segments in $U_{\tilde y}:=-D_yc(U,\tilde y)$ we see $D_p\tilde c(q,\mathbf{0})$ depends linearly on q, whence $D^3_{qqp}\tilde c(q,\mathbf{0})=0$; similarly c^*-segments with respect to $\tilde x$ become line segments in the p variables, $D_q\tilde c(\mathbf{0},p)$ depends linearly on p, $D^3_{ppq}c(\mathbf{0},p)=0$, and the extra factor $D^2_{xy}c(\tilde x,\tilde y)^{-1}$ in our definition of $x(q)$ makes $-D^2_{pq}\tilde c(\mathbf{0},\mathbf{0})$ the identity matrix (whence $q=-D_p\tilde c(\mathbf{0},q)$ and $p=-D_q\tilde c(p,\mathbf{0})$ for all q in $U_{\tilde y}=x^{-1}(U)$ and p in $V_{\tilde x}:=y^{-1}(V)$). We denote $X_{\tilde y}:=x^{-1}(X)$ and choose orthogonal coordinates on U which make $-\hat e_n$ the outer unit normal to $V_{\tilde x}\subset T^*_{\tilde x}U$ at $\tilde p=\mathbf{0}$. Note that $V_{\tilde x}$ is strongly convex by hypothesis (a).

In these variables, as in Figure 1, consider a small cone of height ε and angle θ around the $-\hat e_n$ axis:

$$E_{\theta,\varepsilon}:=\left\{q\in\mathbf{R}^n \mid \left|-\hat e_n-\frac{q}{|q|}\right|\leq\sin\theta,|q|\leq\varepsilon\right\}$$

Observe that, if θ,ε are small enough, then $E_{\theta,\varepsilon}\subset X_{\tilde y}$, and its measure is of order $\varepsilon^n\theta^{n-1}$. Consider now a slight enlargement

$$E'_{\theta,C_0\varepsilon}:=\left\{p=(P,p_n)\in\mathbf{R}^n \mid p_n\leq\theta|p|+C_0\varepsilon|p|^2\right\},$$

of the polar dual cone, where ε will be chosen sufficiently small depending on the large parameter C_0 forced on us later.

The strong convexity ensures $V_{\tilde x}$ is contained in a ball $B_R(R\hat e_n)$ of some radius $R>1$ contained in the half-space $p_n\geq 0$ with boundary sphere passing through the origin. As long as $C_0\varepsilon < (6R)^{-1}$ we claim $E'_{\theta,C_0\varepsilon}$ intersects this ball — a fortiori $V_{\tilde{x}}$ — in a set whose volume tends to zero like θ^{n+1} as $\theta \to 0$. Indeed, from the inequality

$$p_n \leq \theta\sqrt{|P|^2 + p_n^2} + \frac{1}{6}|P| + \frac{1}{3}p_n$$

satisfied by any $(P, p_n) \in E'_{\theta,C_0\varepsilon} \cap B_R(R\hat{e}_n)$ we deduce $p_n^2 \leq |P|^2(1 + 9\theta^2)/(2 - 9\theta^2)$, i.e. $p_n < |P|$ if θ is small enough. Combined with the further inequalities

$$\frac{|P|^2}{2R} \leq p_n \leq \theta\sqrt{|P|^2 + p_n^2} + C_0\varepsilon|P|^2 + C_0\varepsilon p_n^2$$

(the first inequality follows by the strong convexity of $V_{\tilde{x}}$, and the second from the definition of $E'_{\theta,C_\varepsilon}$), this yields $|P| \leq 6\theta\sqrt{2}$ and $p_n \leq O(\theta^2)$ as $\theta \to 0$. Thus $\mathscr{L}^n(E'_{\theta,C_\varepsilon} \cap V_{\tilde{x}}) \leq C\theta^{n+1}$ for a dimension dependent constant C, provided $C_0\varepsilon < (6R)^{-1}$.

The contradiction now will come from the fact that, thanks to the $\tilde{c}$-cyclical monotonicity of $\partial^{\tilde{c}}\tilde{u}$, if we first choose C_0 big and then we take ε sufficiently small, the image of all $q \in E_{\theta,\varepsilon}$ by $\partial^{\tilde{c}}\tilde{u}$ has to be contained in $E'_{\theta,C_0\varepsilon}$ for θ small enough. Since $\partial^{\tilde{c}}\tilde{u}\Big(\overline{X_{\tilde{y}}}\Big) \subset \overline{V_{\tilde{x}}}$ this will imply

$$\varepsilon^n\theta^{n-1} \sim \tilde{\lambda}\mathscr{L}^n(E_{\theta,\varepsilon}) \leq |\partial^{\tilde{c}}\tilde{u}|(E_{\theta,\varepsilon}) \leq \mathscr{L}^n(V_{\tilde{x}} \cap E'_{\theta,C_0\varepsilon}) \leq C\theta^{n+1},$$

which gives a contradiction as $\theta \to 0$, for $\varepsilon > 0$ small but fixed.

Thus all we need to prove is that, if C_0 is big enough, then $\partial^{\tilde{c}}\tilde{u}(E_{\theta,\varepsilon}) \subset E'_{\theta,C_0\varepsilon}$ for any ε sufficiently small. Let $q \in E_{\theta,\varepsilon}$ and $p \in \partial^{\tilde{c}}\tilde{u}(q)$. Notice that

$$\begin{aligned}
&\int_0^1 ds \int_0^1 dt\, D^2_{qp}\tilde{c}(sq, tp)[q, p] \\
&= \tilde{c}(q, p) + \tilde{c}(\mathbf{0}, \mathbf{0}) - \tilde{c}(q, \mathbf{0}) - c(\mathbf{0}, p) \\
&\leq 0
\end{aligned} \tag{5.1}$$

where the last inequality is a consequence of $\tilde{c}$-monotonicity of $\partial^{\tilde{c}}\tilde{u}$, see for instance [55, Definitions 5.1 and 5.7]. Also note that

$$\begin{aligned}
D^2_{qp}\tilde{c}(sq, tp) =& D^2_{qp}\tilde{c}(\mathbf{0}, tp) + \int_0^s ds' D^3_{qqp}\tilde{c}(s'q, tp)[q] \\
=& D^2_{qp}\tilde{c}(\mathbf{0}, \mathbf{0}) + \int_0^t dt' D^3_{qpp}\tilde{c}(\mathbf{0}, t'p)[p] \\
&+ \int_0^s ds' D^3_{qqp}\tilde{c}(s'q, \mathbf{0})[q] + \int_0^s ds' \int_0^t dt' D^4_{qqpp}\tilde{c}(s'q, t'p)[q, p] \\
=& -I_n + \int_0^s ds' \int_0^t dt' D^4_{qqpp}\tilde{c}(s'q, t'p)[q, p]
\end{aligned} \tag{5.2}$$

since $D^3_{qpp}\tilde{c}(\mathbf{0}, t'p)$ and $D^3_{qqp}\tilde{c}(s'q, \mathbf{0})$ vanish in our chosen coordinates, and $-D^2_{pq}\tilde{c}(\mathbf{0}, \mathbf{0}) = I_n$ is the identity matrix. Then, plugging (5.2) into (5.1) yields

$$\begin{aligned}
-\langle q, p\rangle &\leq -\int_0^1 ds \int_0^1 dt \int_0^s ds' \int_0^t dt' D^4_{qqpp}\tilde{c}(s'q, t'p)[q, q, p, p] \\
&\leq C_0|q|^2|p|^2
\end{aligned}$$

for constant $C_0 = \sup_{(q,p)\in X_{\tilde y}\times V_{\tilde x}} \|D^4_{qqpp}\tilde c\|$. Since the term $-D^4_{qqpp}\tilde c$ is exactly the cross-curvature (2.1), its tensorial nature implies C_0 depends only on $\|c\|_{C^4(U\times V)}$ and the bi-Lipschitz constants $\beta_c^\pm$ from (4.3).

From the above inequality and the definition of $E_{\theta,\varepsilon}$ we deduce

$$p_n = \langle p, \hat e_n + \frac{q}{|q|}\rangle - \langle p, \frac{q}{|q|}\rangle \le \theta|p| + C_0\varepsilon|p|^2$$

so $p \in E'_{\theta,C_0\varepsilon}$ as desired. □

6 Alexandrov type estimates for c-convex functions

In this section we prove the key estimates for c-convex potential functions u which will eventually lead to the Hölder continuity and injectivity of optimal maps. Namely, we extend Alexandrov type estimates commonly used in the analysis of convex solutions to the Monge-Ampère equation with $c(x,y) = -\langle x,y\rangle$, to c-convex solutions of the c-Monge-Ampère for general **(B3)** cost functions. These estimates, Theorems 6.2 and 6.11, are of independent interest: they concern *sections* of u, i.e. the convex sub-level sets of the modified potential $\tilde u$ of Theorem 4.3, and compare the range of values and boundary behaviour of $\tilde u$ on each section with the volume of the section. We describe them briefly before we begin the details.

Fix $p \in \mathbf{R}^n$, $p_0 \in \mathbf{R}$, and a positive symmetric matrix $P > 0$. It is elementary to see that the range of values taken by the parabola $\tilde u(q) = q^t P q + p\cdot q + p_0$ on any non-empty sub-level set $Q = \{\tilde u \le 0\}$ is determined by $\det P$ and the volume of Q:

$$|\min_{q\in Q} \tilde u(q)|^n = \left(\frac{\mathscr{L}^n(Q)}{\mathscr{L}^n(B_1)}\right)^2 \det P. \tag{6.1}$$

Moreover, the parabola tends to zero linearly as the boundary of Q is approached. Should the parabola be replaced by a convex function satisfying $\lambda \le \det D^2\tilde u \le \Lambda$ throughout a fixed fraction of Q, two of the cornerstones of Caffarelli's regularity theory are that the identity (6.1) remains true — up to a factor controlled by λ, Λ and the fraction of Q on which these bounds hold — and moreover that $\tilde u$ tends to zero at a rate no slower than $\mathrm{dist}_{\partial Q}(q)^{1/n}$ as $q \to \partial Q$. The present section is devoted to showing that under **(B3)**, similar estimates hold for level-set convex solutions $\tilde u$ of the c-Monge Ampére equation, on sufficiently small sub-level sets. Although the rate $\mathrm{dist}_{\partial Q}(q)^{1/2^{n-1}}$ we obtain for decay of $\tilde u$ near ∂Q is probably far from optimal, it turns out to be sufficient for our present purpose.

We begin with a Lipschitz estimate on the cost function c, which turns out to be useful.

Lemma 6.1 (Modified cost gradient direction is Lipschitz)**.** *Assume* **(B0)**–**(B2)**. *Fix $\tilde y \in \overline{V}$. For $\tilde c \in C^3\big(\overline{U}_{\tilde y} \times \overline{V}\big)$ from Definition 4.1 and each $q, \tilde q \in \overline{U}_{\tilde y}$ and fixed target $y \in \overline{V}$,*

$$|-D_q\tilde c(q,y) + D_q\tilde c(\tilde q,y)| \le \frac{1}{\varepsilon_c}|q-\tilde q|\,|D_q\tilde c(\tilde q,y)|, \tag{6.2}$$

where ε_c *is given by* $\varepsilon_c^{-1} = 2(\beta_c^+)^4(\beta_c^-)^6\|D^3_{xxy}c\|_{L^\infty(U\times V)}$ *in the notation* (4.3)*. If* $y \neq \tilde{y}$ *so neither gradient vanishes, then*

$$\left|-\frac{D_q\tilde{c}(q,y)}{|D_q\tilde{c}(q,y)|} + \frac{D_q\tilde{c}(\tilde{q},y)}{|D_q\tilde{c}(\tilde{q},y)|}\right| \leq \frac{2}{\varepsilon_c}|q-\tilde{q}|. \tag{6.3}$$

Proof. For fixed $\tilde{q} \in \overline{U}_{\tilde{y}}$ introduce the $\tilde{c}$-exponential coordinates $p(y) = -D_q\tilde{c}(\tilde{q},y)$. The bi-Lipschitz constants (4.3) of this coordinate change are estimated by $\beta^\pm_{\tilde{c}} \leq \beta_c^+\beta_c^-$ as in Corollary 4.5. Thus

$$\begin{aligned}\operatorname{dist}(y,\tilde{y}) &\leq \beta^-_{\tilde{c}}|-D_q\tilde{c}(\tilde{q},y) + D_q\tilde{c}(\tilde{q},\tilde{y})|\\ &\leq \beta_c^+\beta_c^-|D_q\tilde{c}(\tilde{q},y)|.\end{aligned}$$

where $\tilde{c}(q,\tilde{y}) \equiv 0$ from Definition 4.1 has been used. Similarly, noting the convexity **(B2)** of $V_{\tilde{q}} := p(V)$,

$$\begin{aligned}|-D_q\tilde{c}(\tilde{q},y) + D_q\tilde{c}(q,y)| &= |-D_q\tilde{c}(\tilde{q},y) + D_q\tilde{c}(q,y) + D_q\tilde{c}(\tilde{q},\tilde{y}) - D_q\tilde{c}(q,\tilde{y})|\\ &\leq \|D^2_{qq}D_p\tilde{c}\|_{L^\infty(U_{\tilde{y}}\times\tilde{V}_{\tilde{q}})}|\tilde{q}-q||p(y)-p(\tilde{y})|\\ &\leq \|D^2_{qq}D_y\tilde{c}\|_{L^\infty(U_{\tilde{y}}\times V)}(\beta_c^-\beta_c^+)^2|\tilde{q}-q|\operatorname{dist}(y,\tilde{y})\end{aligned}$$

Then (6.2) follows since $|D^2_{qq}D_y\tilde{c}| \leq ((\beta_c^-)^2 + \beta_c^+(\beta_c^-)^3))|D^2_{xx}D_yc| \leq 2\beta_c^+(\beta_c^-)^3|D^2_{xx}D_yc|$. (The last inequality follows from $\beta_c^+\beta_c^- \geq 1$.)

Finally, to prove (6.3) we simply use the triangle inequality and (6.2) to get

$$\begin{aligned}\left|-\frac{D_q\tilde{c}(q,y)}{|D_q\tilde{c}(q,y)|} + \frac{D_q\tilde{c}(\tilde{q},y)}{|D_q\tilde{c}(\tilde{q},y)|}\right| &\leq \left|-\frac{D_q\tilde{c}(q,y)}{|D_q\tilde{c}(q,y)|} + \frac{D_q\tilde{c}(q,y)}{|D_q\tilde{c}(\tilde{q},y)|}\right| + \left|-\frac{D_q\tilde{c}(q,y)}{|D_q\tilde{c}(\tilde{q},y)|} + \frac{D_q\tilde{c}(\tilde{q},y)}{|D_q\tilde{c}(\tilde{q},y)|}\right|\\ &= \frac{\big|-|D_q\tilde{c}(q,y)| + |D_q\tilde{c}(\tilde{q},y)|\big|}{|D_q\tilde{c}(\tilde{q},y)|} + \frac{|-D_q\tilde{c}(q,y) + D_q\tilde{c}(\tilde{q},y)|}{|D_q\tilde{c}(\tilde{q},y)|}\\ &\leq \frac{2}{\varepsilon_c}|q-\tilde{q}|.\end{aligned}$$

□

6.1 Alexandrov type lower bounds

In this subsection we prove one of the key estimates of this paper, namely, a bound on the infimum on a c-convex function inside a section in terms of the measure of the section and the mass of the c-Monge-Ampère measure inside that section. The corresponding result for the affine cost function is very easy, but in our case the unavoidable nonlinearity of the cost function requires new ideas.

Theorem 6.2 (Alexandrov type lower bounds)**.** *Assume* **(B0)**–**(B3)**, *define* $\tilde{c} \in C^3\big(\overline{U}_{\tilde{y}}\times\overline{V}\big)$ *as in Definition 4.1, and let* $\gamma^\pm_{\tilde{c}} = \gamma^\pm_{\tilde{c}}(Q\times V)$ *be as in (4.4). Let* $\tilde{u} : \overline{U}_{\tilde{y}} \longmapsto \mathbf{R}$ *be a* $\tilde{c}$*-convex function as in Theorem 4.3, and let* $Q := \{\tilde{u} \leq 0\} \subset \overline{U}_{\tilde{y}}$*. Note that* $Q \subset \mathbf{R}^n$ *is convex.*

Let $\varepsilon_c > 0$ small be given by Lemma 6.1. Finally, let E be the ellipsoid given by John's Lemma (see (3.9)), and assume there exists a small ellipsoid $E_\delta = x_0 + \delta E \subset Q$, with $\delta \le \min\{1, \varepsilon_c/(4\,\mathrm{diam}(E))\}$ and $|\partial^{\tilde c}\tilde u| \ge \lambda > 0$ inside E_δ. We also assume that $E_\delta \subset B \subset 4B \subset U_{\bar y}$, where B, $4B$ are a ball and its dilation. Then there exists a constant $C(n)$, depending only on the dimension, such that

$$\mathscr{L}^n(Q)^2 \le C(n)\frac{\gamma^-_{\tilde c}}{\delta^{2n}\lambda}\,|\inf_Q \tilde u|^n. \tag{6.4}$$

The proof of the theorem above, which is given in the last part of this subsection, relies on the following result, which will also play a key role in Section 9 (see (9.4)) to show the engulfing property and obtain Hölder continuity of optimal maps.

Lemma 6.3 (Dual norm estimates)**.** *With the same notation and assumptions as in Theorem 6.2, let $\mathscr{K} \subset Q$ be an open convex set such that* $\mathrm{diam}(\mathscr{K}) \le \varepsilon_c/4$. *We also assume that there exists a ball B such that $\mathscr{K} \subset B \subset 4B \subset U_{\bar y}$, where $4B$ denotes the dilation of B by a factor 4 with respect to its center. (This assumption is to locate all relevant points inside the domain where the assumptions on $\tilde c$ hold.) Up to a translation, assume that the ellipsoid associated to $\mathscr{K}$ by John's Lemma (see 3.9) is centered at the origin $\mathbf{0}$, and for any $\rho \in (0,1)$ let $\rho\mathscr{K}$ denote the dilation of $\mathscr{K}$ with respect to the origin. Moreover, let $\|\cdot\|^*_{\mathscr{K}}$ denote the "dual norm" associated to $\mathscr{K}$, that is*

$$\|v\|^*_{\mathscr{K}} := \sup_{w\in\mathscr{K}} w\cdot v. \tag{6.5}$$

Then, for any $\rho < 1$ setting $C_*(n,\rho) := \frac{8n}{(1-\rho)^2}$ implies

$$\|-D_q\tilde c(q,y)\|^*_{\mathscr{K}} \le C_*(n,\rho)|\inf_{\mathscr{K}} \tilde u| \qquad \forall\, q \in \rho\mathscr{K},\ y \in \partial^{\tilde c}\tilde u(\rho\mathscr{K}). \tag{6.6}$$

Before proving the above lemma, let us explain the geometric intuition behind the result: for $v \in \mathbf{R}^n \setminus \{\mathbf{0}\}$, let H_v be the supporting hyperplane to $\mathscr{K}$ orthogonal to v and contained inside the half-space $\{q \mid q\cdot v > 0\}$. Then

$$\mathrm{dist}(\mathbf{0}, H_v) = \sup_{w\in\mathscr{K}} w\cdot\frac{v}{|v|} = \frac{\|v\|^*_{\mathscr{K}}}{|v|},$$

see Figure 2, and Lemma 6.3 states that, for all $v = -D_q\tilde c(q,y)$, with $q \in \rho\mathscr{K}$ and $y \in \partial^{\tilde c}\tilde u(\rho\mathscr{K})$,

$$\mathrm{dist}(\mathbf{0}, H_v)|v| \le C_*(n,\rho)|\inf_{\mathscr{K}} \tilde u|.$$

Hence, roughly speaking, (6.6) is just telling us that the size of gradient at a point inside $\rho\mathscr{K}$ times the width of $\mathscr{K}$ in the direction orthogonal to the gradient is controlled, up to a factor $C_*(n,\rho)$, by the infimum of $\tilde u$ inside $\mathscr{K}$ (all this provided the diameter of $\mathscr{K}$ is sufficiently small). This would be a standard estimate if we were working with convex functions and c was the standard quadratic cost in $\mathbf{R}^n$, but in our situation the proof is both subtle and involved.

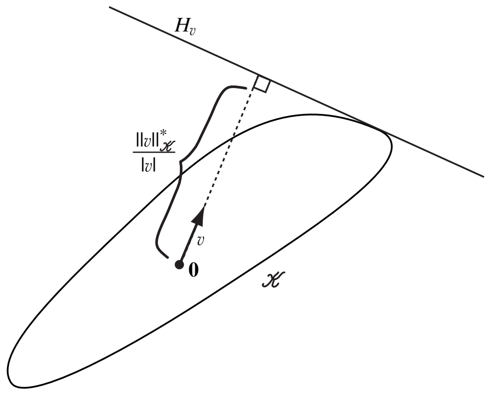

Figure 2: The quantity $\frac{\|v\|^*_{\mathscr{K}}}{|v|}$ represents the distance between the origin and the supporting hyperplane orthogonal to v.

To prove (6.6) let us start observing that, since $|w| \le \operatorname{diam}(\mathscr{K})$ for $w \in \mathscr{K}$, the following useful inequality holds:

$$\|v\|^*_{\mathscr{K}} \le \operatorname{diam}(\mathscr{K})\,|v| \qquad \forall\, v \in \mathbf{R}^n. \tag{6.7}$$

We will need two preliminary results.

Lemma 6.4. *With the same notation and assumptions as in Lemma 6.3, let $q \in \rho\mathscr{K}$, $y \in V$, and let m_y be a function of the form $m_y := -\tilde{c}(\cdot, y) + C_y$ for some constant $C_y \in \mathbf{R}$. Set $v := -D_q m_y(q)$, assume that $\mathscr{K} \subset \{m_y < 0\}$, and let q_0 denote the intersection of the half-line $\ell_v^q := q + \mathbf{R}^+ v = \{q + tv \,|\, t > 0\}$ with $\{m_y = 0\}$ (assuming it exists in $U_{\tilde{y}}$). Define*

$$\hat{q}_0 := \begin{cases} q_0 & \text{if } |q_0 - q| \le \operatorname{diam}\mathscr{K}, \\ q + \frac{\operatorname{diam}(\mathscr{K})}{|v|} v & \text{if } |q_0 - q| \ge \operatorname{diam}(\mathscr{K}) \text{ or } q_0 \text{ does not exist in } U_{\tilde{y}}, \end{cases}$$

(Notice that, by the assumption $\mathscr{K} \subset B \subset 4B \subset U_{\tilde{y}}$, we have $\hat{q}_0 \in U_{\tilde{y}}$.) Then

$$2|\hat{q}_0 - q| \ge (1 - \rho)\frac{\|v\|^*_{\mathscr{K}}}{|v|}. \tag{6.8}$$

Proof. In the case $\hat{q}_0 = q + \operatorname{diam}(\mathscr{K})v$, the inequality (6.8) follows from (6.7). Thus, we can assume that q_0 exists in $U_{\tilde{y}}$ and that $\hat{q}_0 = q_0$. Let us recall that $\{m_y < 0\}$ is convex (see Theorem 4.3). Now, let $\tilde{H}_v$ denote the hyperplane tangent to $\{m_y = 0\}$ at q_0. Let us observe that $\tilde{H}_v$ is orthogonal to the vector $v_0 := -D_q m_y(q_0)$, see Figure 3. Since $\mathscr{K} \subset \{m_y < 0\}$, we have $\tilde{H}_v \cap \mathscr{K} = \emptyset$, which implies

$$\operatorname{dist}(\mathbf{0}, \tilde{H}_v) \ge \sup_{w \in \mathscr{K}} w \cdot \frac{v_0}{|v_0|} = \frac{\|v_0\|^*_{\mathscr{K}}}{|v_0|}.$$

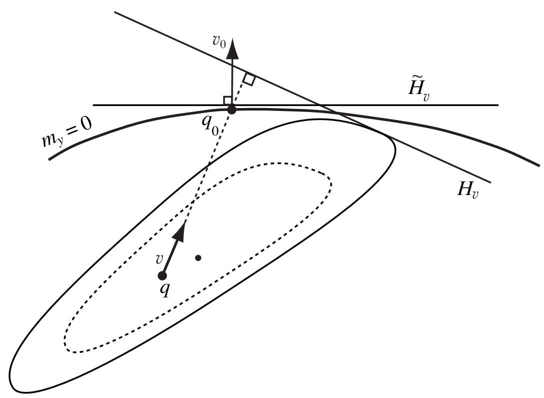

Figure 3: It is easily seen that $|q_0 - q|$ controls $\mathrm{dist}(\mathbf{0}, H_{\tilde{v}})$, and using that the gradient of m_y does not vary much we deduce that the latter is close to $\mathrm{dist}(\mathbf{0}, H_v)$ in terms of $|q_0 - q|$.

Moreover, since $q \in \rho\mathscr{K}$ and $\tilde{H}_v \cap \mathscr{K} = \emptyset$, we also have

$$\mathrm{dist}(q, \tilde{H}_v) \geq (1-\rho)\,\mathrm{dist}(\mathbf{0}, \tilde{H}_v).$$

Hence, observing that $q_0 \in \tilde{H}_v$ we obtain

$$|q_0 - q| \geq \mathrm{dist}(q, \tilde{H}_v) \geq (1-\rho)\frac{\|v_0\|^*_{\mathscr{K}}}{|v_0|}.$$

Now, to conclude the proof, we observe that (6.3) applied with $v_0 = -D_q\tilde{c}(q_0, y)$ and $v = -D_q\tilde{c}(q, y)$, together with (6.7) and the assumption $\mathrm{diam}(\mathscr{K}) \leq \varepsilon_c/4$, implies

$$\left\|\frac{v}{|v|} - \frac{v_0}{|v_0|}\right\|^*_{\mathscr{K}} \leq \mathrm{diam}(\mathscr{K})\left|\frac{v}{|v|} - \frac{v_0}{|v_0|}\right| \leq \frac{2\,\mathrm{diam}(\mathscr{K})}{\varepsilon_c}|q_0 - q| \leq \frac{|q_0 - q|}{2}.$$

Combining all together and using the triangle inequality for $\|\cdot\|^*_{\mathscr{K}}$ (which is a consequence of the convexity of $\mathscr{K}$) we finally obtain

$$\begin{aligned}
\left\|\frac{v}{|v|}\right\|^*_{\mathscr{K}} &\leq \left\|\frac{v_0}{|v_0|}\right\|^*_{\mathscr{K}} + \left\|\frac{v}{|v|} - \frac{v_0}{|v_0|}\right\|^*_{\mathscr{K}} \\
&\leq \frac{|q_0 - q|}{1-\rho} + \frac{|q_0 - q|}{2} \\
&\leq \frac{2|q_0 - q|}{1-\rho},
\end{aligned}$$

as desired. □

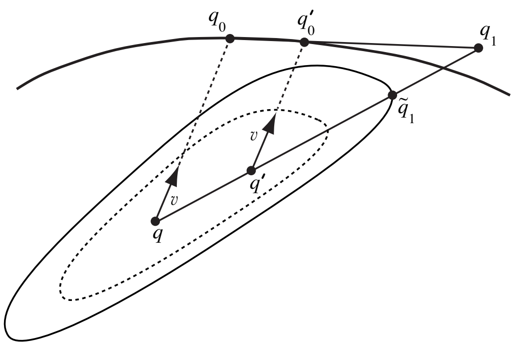

Figure 4: In this figure we assume $\hat{q}_0 = q_0$. Using similarity, we only need to bound from below the ratio $\frac{|q'-q_1|}{|q-q_1|}$, and the latter is greater or equal than $\frac{|q'-\tilde{q}_1|}{|q-\tilde{q}_1|}$. An easy geometric argument allows to control this last quantity. .

Lemma 6.5. *With the same notation and assumptions as in Lemmata 6.3 and 6.4, fix $q' \in \rho\mathscr{K}$, and let $\ell_v^{q'}$ denote the half line $\ell_v^{q'} := q' + \mathbf{R}^+ v = \{q' + tv \mid t > 0\}$. Denote by q'_0 the intersection of $\ell_v^{q'}$ with $\{m_y = 0\}$ (assuming it exists in $U_{\tilde{y}}$). Then,*

$$|q'_0 - q'| \geq \frac{1-\rho}{2n} |\hat{q}_0 - q|.$$

Proof. Let $\ell_{qq'}$ and $\ell_{\hat{q}_0 q'_0}$ denote the lines passing through q, q' and $\hat{q}_0, q'_0$ respectively, and denote by $q_1 := \ell_{qq'} \cap \ell_{\hat{q}_0 q'_0}$ their intersection point (see Figure 4); since the first four points lie in the same plane, we can (after slightly perturbing q or q' if necessary) assume this intersection exists and is unique. Note that in the following we will only use convexity of $\mathscr{K}$, and in particular we do not require $q_1 \in U_{\tilde{y}}$.

Let us now distinguish two cases, depending whether $|q-q_1| \leq |q'-q_1|$ or $|q-q_1| \geq |q'-q_1|$.

• $|q - q_1| \leq |q' - q_1|$: In this case we simply observe that, since $\hat{q}_0 - q$ and $q'_0 - q'$ are parallel, by similarity

$$1 \leq \frac{|q' - q_1|}{|q - q_1|} = \frac{|q'_0 - q'|}{|\hat{q}_0 - q|},$$

and so the result is proved since $1 \geq \frac{1-\rho}{2n}$.

• $|q - q_1| \geq |q' - q_1|$: In this second case, we first claim that $q_1 \notin \mathscr{K}$. Indeed, we observe that since $\hat{q}_0 - q$ and $q'_0 - q'$ are parallel, the point q_1 cannot belong to the segment joining $\hat{q}_0$ and q'_0. Hence, since $\hat{q}_0 \in \{m_y \leq 0\}$, $q'_0 \in \{m_y = 0\}$, and $\mathscr{K} \subset \{m_y < 0\}$ by assumption, by the convexity of the set $\{m_y \leq 0\}$ we necessarily have q_1 outside $\{m_y < 0\}$, thus outside of $\mathscr{K}$, which proves the claim.

Thus, we can find the point $\tilde{q}_1$ obtained by intersecting $\partial\mathscr{K}$ with the segment going from

q' to q_1, and by the elementary inequality

$$\frac{a+c}{b+c} \geq \frac{a}{b} \qquad \forall\, 0 < a \leq b,\, c \geq 0,$$

we get

$$\frac{|q' - q_1|}{|q - q_1|} \geq \frac{|q' - \tilde{q}_1|}{|q - \tilde{q}_1|}.$$

Now, to estimate the right hand side from below, let L be the affine transformation provided by John's Lemma such that $B_1 \subset L^{-1}(\mathscr{K}) \subset B_n$. Since the points $q, q', \tilde{q}_1$ are aligned, we have

$$\frac{|q' - \tilde{q}_1|}{|q - \tilde{q}_1|} = \frac{|L^{-1}q' - L^{-1}\tilde{q}_1|}{|L^{-1}q - L^{-1}\tilde{q}_1|}.$$

Now, since $L^{-1}q, L^{-1}\tilde{q}_1 \in B_n$, we immediately get

$$|L^{-1}q - L^{-1}\tilde{q}_1| \leq 2n.$$

On the other hand, since $L^{-1}q' \in \rho\big(L^{-1}(\mathscr{K})\big)$ while $L^{-1}\tilde{q}_1 \in \partial\big(L^{-1}(\mathscr{K})\big)$, it follows from $B_1 \subset L^{-1}(\mathscr{K})$ that

$$|L^{-1}q' - L^{-1}\tilde{q}_1| \;\geq\; 1 - \rho.$$

To see this last inequality, for each point $w \in \partial\big(L^{-1}(\mathscr{K})\big)$ consider the convex hull $\mathcal{C}_w \subset L^{-1}(\mathscr{K})$ of $\{w\} \cup B_1$. Then, centered at the point $\rho w \in \partial\Big(\rho\big(L^{-1}(\mathscr{K})\big)\Big)$ there exists a ball of radius $1-\rho$ contained in $\mathcal{C}_w$, thus in $L^{-1}(\mathscr{K})$. This shows that the distance from any point in $\rho\big(L^{-1}(\mathscr{K})\big)$ to $\partial\big(L^{-1}(\mathscr{K})\big)$ is at least $1-\rho$, proving the inequality.

From these inequalities and using again similarity, we get

$$\frac{|q_0' - q'|}{|\hat{q}_0 - q|} = \frac{|q' - q_1|}{|q - q_1|} \geq \frac{1-\rho}{2n},$$

concluding the proof. ☐

Proof of Lemma 6.3. Fix $v = -D_q\tilde{c}(q,y)$, with $q \in \rho\mathscr{K}$, $q' \in \rho\mathscr{K}$ and $y \in \partial^{\tilde{c}}\tilde{u}(q')$, and define the function $m_y(\cdot) := -\tilde{c}(\cdot, y) + \tilde{c}(q', y) + \tilde{u}(q')$. Notice that since $y \in \partial^{\tilde{c}}\tilde{u}(q')$, the inclusion $Q = \{\tilde{u} \leq 0\} \subset \{m_y \leq 0\}$ holds.

Recall the point q_0' from Lemma 6.5 (when it exists in $U_{\tilde{y}}$), and the point $\hat{q}_0 \in U_{\tilde{y}}$ from Lemma 6.4. Define,

$$\hat{q}_0' := \begin{cases} q_0' & \text{if } |q_0' - q'| \leq \operatorname{diam}\mathscr{K}, \\ q' + \frac{\operatorname{diam}(\mathscr{K})}{|v|}v & \text{if } |q_0' - q'| \geq \operatorname{diam}(\mathscr{K}) \text{ or } q_0' \text{ does not exist in } U_{\tilde{y}}, \end{cases}$$

Notice that by the assumption $\mathscr{K} \subset B \subset 4B \subset U_{\tilde{y}}$, the points $\hat{q}_0$, $\hat{q}_0'$ belong to $U_{\tilde{y}}$. Let $[q', \hat{q}_0'] \subset U_{\tilde{y}}$ denote the segment going from q' to $\hat{q}_0'$. Then, since $\hat{q}_0' \in \{m_y \leq 0\}$ and $\tilde{u}$ is negative inside $\mathscr{K} \subset Q$, we get

$$\begin{aligned} |\inf_{\mathscr{K}} \tilde{u}| &\geq -\tilde{u}(q') \geq -\tilde{u}(q') + \tilde{u}(\hat{q}_0') \\ &\geq -\tilde{c}(\hat{q}_0', y) + \tilde{c}(q', y) = \int_{[q', \hat{q}_0']} \langle -D_q\tilde{c}(w,y), \frac{v}{|v|}\rangle \, dw \end{aligned} \tag{6.9}$$

Since $q, q' \in \mathscr{K}$ and $|\hat{q}_0' - q'| \le \operatorname{diam}(\mathscr{K})$, we have

$$|q - w| \le 2\operatorname{diam}(\mathscr{K}) \le \varepsilon_c/2 \qquad \forall\, w \in [q', \hat{q}_0'].$$

Hence, recalling that $-D_q\tilde{c}(q,y) = v$ and applying Lemma 6.1, we obtain

$$\langle -D_q\tilde{c}(w,y), \frac{v}{|v|}\rangle \ge \frac{|v|}{2} \qquad \forall\, w\in [q', \hat{q}_0'].$$

Applying this and Lemmas 6.4 and 6.5 to (6.9), we finally conclude

$$\begin{aligned}
|\inf_{\mathscr{K}} \tilde{u}| &\ge \min\left\{\operatorname{diam}(\mathscr{K}), |q_0' - q'|\right\} \frac{|v|}{2} && \text{(from (6.9))}\\
&\ge \min\left\{\operatorname{diam}(\mathscr{K}), \frac{1-\rho}{2n}|\hat{q}_0 - q|\right\} \frac{|v|}{2} && \text{(from Lemma 6.5)}\\
&\ge \min\left\{\operatorname{diam}(\mathscr{K}), \frac{(1-\rho)^2}{4n}\frac{\|v\|^*_{\mathscr{K}}}{|v|}\right\} \frac{|v|}{2} && \text{(from Lemma 6.4)}\\
&= \min\left\{\frac{1}{2}\operatorname{diam}(\mathscr{K})|v|, \frac{(1-\rho)^2}{8n}\|v\|^*_{\mathscr{K}}\right\}\\
&\ge \min\left\{\frac{1}{2}\|v\|^*_{\mathscr{K}}, \frac{(1-\rho)^2}{8n}\|v\|^*_{\mathscr{K}}\right\} && \text{(from (6.7))}\\
&= \frac{(1-\rho)^2}{8n}\|v\|^*_{\mathscr{K}}
\end{aligned}$$

as desired, completing the proof. □

Proof of Theorem 6.2. Set $h := |\inf_Q \tilde{u}|$. Since $\mathscr{L}^n(Q) \le n^n \mathscr{L}^n(E)$, it is enough to show

$$\lambda \mathscr{L}^n(E)^2 \le C(n)\gamma^-_{\tilde{c}} \frac{h^n}{\delta^{2n}}. \tag{6.10}$$

So, the rest of the proof is devoted to (6.10). Let $E_\delta = x_0 + \delta E$ be defined as in the statement of the proposition. With no loss of generality, up to a change of coordinates we can assume that E_δ is of the form $\{q \mid \sum_i a_i^2 {q_i}^2 < 1\}$. Define

$$\mathcal{C}_0 = \left\{q_0 \in T_{\mathbf{0}}Q \mid q_0 = -D_q\tilde{c}(\mathbf{0}, y), \quad y \in \partial^{\tilde{c}}\tilde{u}\big(\tfrac{1}{2}E_\delta\big)\right\}.$$

Notice that

$$\mathscr{L}^n\big(\partial^{\tilde{c}}\tilde{u}\big(\tfrac{1}{2}E_\delta\big)\big) \le \gamma^-_{\tilde{c}}\, \mathscr{L}^n(\mathcal{C}_0) \tag{6.11}$$

(see (4.4)). Now, let us define the norm

$$\|q\|^*_{E_\delta} := \sqrt{\sum_{i=1}^n \frac{q_i^2}{a_i^2}} = \sup_{v\in E_\delta} v \cdot q.$$

Since $\mathscr{K} = E_\delta$ satisfies the assumptions for Lemma 6.3, choosing $\rho = 1/2$ we get

$$\|q_0\|^*_{E_\delta} \leq 32nh \text{ for all } q_0 \in \mathcal{C}_0. \tag{6.12}$$

Let $\Phi_h : \mathbf{R}^n \to \mathbf{R}^n$ denote the linear map $\Phi_h(x) := 32nh\,(a_1x_1, \ldots, a_nx_n)$. Then from (6.12) it follows

$$\mathcal{C}_0 \subset \Phi_h(B_1). \tag{6.13}$$

Since

$$\mathscr{L}^n(\Phi_h(B_1))) \leq C(n)\frac{h^n}{\mathscr{L}^n(E_\delta)}, \tag{6.14}$$

combining (6.11), (6.13), and (6.14), we get

$$\mathscr{L}^n\big(\partial^{\tilde{c}}\tilde{u}\big(\tfrac{1}{2}E_\delta\big)\big) \leq C(n)\gamma^-_{\tilde{c}}\frac{h^n}{\mathscr{L}^n(E_\delta)}.$$

As $\mathscr{L}^n\big(\partial^{\tilde{c}}\tilde{u}\big(\tfrac{1}{2}E_\delta\big)\big) \geq \lambda 2^{-n}\mathscr{L}^n(E_\delta)$ and $\mathscr{L}^n(E_\delta) = \delta^n\mathscr{L}^n(E)$, we obtain the desired conclusion. □

6.2 Bounds for $\tilde{c}$-cones over convex sets

We now progress toward the Alexandrov type upper bounds in Theorem 6.11. In this subsection we construct and study the $\tilde{c}$-cone associated to the section of a $\tilde{c}$-convex function. This $\tilde{c}$-cone — whose entire $\tilde{c}$-Monge-Ampère mass concentrates at a single prescribed point — plays an essential role in our proof of Lemma 6.10.

Definition 6.6 ($\tilde{c}$-cone)**.** *Assume* **(B0)**–**(B3)***, and let* $\tilde{u} : \overline{U}_{\tilde{y}} \longmapsto \mathbf{R}$ *be the* $\tilde{c}$*-convex function with convex level sets from Theorem 4.3. Let* Q *denote the section* $\{\tilde{u} \leq 0\}$*, fix* $\tilde{q} \in \operatorname{int} Q$*, and assume* $\tilde{u} = 0$ *on* ∂Q *and* $\overline{Q} \subset U_{\tilde{y}}$*. The* $\tilde{c}$-cone $h^{\tilde{c}} : U_{\tilde{y}} \longmapsto \mathbf{R}$ generated by $\tilde{q}$ and Q with height $-\tilde{u}(\tilde{q}) > 0$ *is given by*

$$h^{\tilde{c}}(q) := \sup_{y\in\overline{V}}\{-\tilde{c}(q,y) + \tilde{c}(\tilde{q},y) + \tilde{u}(\tilde{q}) \mid -\tilde{c}(q,y) + \tilde{c}(\tilde{q},y) + \tilde{u}(\tilde{q}) \leq 0 \text{ on } \partial Q\}. \tag{6.15}$$

Notice the $\tilde{c}$-cone $h^{\tilde{c}}$ depends only on the convex set $Q \subset U_{\tilde{y}}$, $\tilde{q} \in \operatorname{int} Q$, and the value $\tilde{u}(\tilde{q})$, but is otherwise independent of $\tilde{u}$. Recalling that $\tilde{c}(q, \tilde{y}) \equiv 0$ on $U_{\tilde{y}}$, we record several key properties of the $\tilde{c}$-cone:

Lemma 6.7 (Basic properties of $\tilde{c}$-cones)**.** *Adopting the notation and hypotheses of Definition 6.6, let* $h^{\tilde{c}} : U_{\tilde{q}} \longmapsto \mathbf{R}$ *be the* $\tilde{c}$*-cone generated by* $\tilde{q}$ *and* Q *with height* $-\tilde{u}(\tilde{q}) > 0$*. Then*

(a) $h^{\tilde{c}}$ *has convex level sets; furthermore, it is a convex function if* **(B4)** *holds;*

(b) $h^{\tilde{c}}(q) \geq h^{\tilde{c}}(\tilde{q}) = \tilde{u}(\tilde{q})$ *for all* $q \in Q$*;*

(c) $h^{\tilde c} = 0$ *on* ∂Q;

(d) $\partial^{\tilde c} h^{\tilde c}(\tilde q) \subset \partial^{\tilde c}\tilde u(Q)$.

Proof. Property (a) is a consequence of the level-set convexity of $q \longmapsto -\tilde c(q,y)$ proved in Theorem 4.3, or its convexity assuming **(B4)**. Moreover, since $-\tilde c(q,\tilde y)+\tilde c(\tilde q,\tilde y)+\tilde u(\tilde q) = \tilde u(\tilde q)$ for all $q \in U_{\tilde y}$, (b) follows.

For each pair $q_0 \in \partial Q$ and $y_0 \in \partial^{\tilde c}\tilde u(q_0)$, consider the supporting mountain $m_0(q) = -\tilde c(q,y_0) + \tilde c(q_0,y_0)$, i.e. $m_0(q_0) = 0 = \tilde u(q_0)$ and $m_0 \le \tilde u$. Consider the $\tilde c$-segment $y(t)$ connecting $y(0) = y_0$ and $y(1) = \tilde y$ in V with respect to q_0. Since $-\tilde c(q,\tilde y) \equiv 0$, by continuity there exists some $t \in [0,1[$ for which $m_t(q) := -\tilde c(q,y(t)) + \tilde c(q_0,y(t))$ satisfies $m_t(\tilde q) = \tilde u(\tilde q)$. From Loeper's maximum principle (Theorem 3.3 above), we have

$$m_t \le \max[m_0, -\tilde c(\cdot,\tilde y)+\tilde c(q_0,\tilde y)] = \max[m_0, 0],$$

and therefore, from $m_0 \le \tilde u$,

$$m_t \le 0 \text{ on } Q.$$

By the construction, m_t is of the form

$$-\tilde c(\cdot,y(t)) + \tilde c(\tilde q,y(t)) + \tilde u(\tilde q),$$

and vanishes at q_0. This proves (c). Finally (d) follows from (c) and the fact that $h^{\tilde c}(\tilde q) = \tilde u(\tilde q)$. Indeed, it suffices to move down the supporting mountain of $h^{\tilde c}$ at $\tilde q$ until the last moment at which it touches the graph of $\tilde u$ inside Q. The conclusion then follows from Loeper's local to global principle, Corollary 3.4 above. □

The following estimate shows that the Monge-Ampère measure, and the relative location of the vertex within the section which generates it, control the height of any well-localized $\tilde c$-cone. Together with Lemma 6.7(d), this proposition plays a key role in the proof of our Alexandrov type estimate (Lemma 6.10).

Proposition 6.8 (Lower bound on the Monge-Ampère measure of a small $\tilde c$-cone)**.** *Assume* **(B0)**–**(B3)**, *and define $\tilde c \in C^3\big(\overline{U}_{\tilde y}\times\overline{V}\big)$ as in Definition 4.1. Let $Q \subset U_{\tilde y}$ be a compact convex set, and $h^{\tilde c}$ the $\tilde c$-cone generated by $\tilde q \in \operatorname{int} Q$ of height $-h^{\tilde c}(\tilde q) > 0$ over Q. Let Π^+, Π^- be two parallel hyperplanes contained in $T^*_{\tilde y}V \setminus Q$ and touching ∂Q from two opposite sides. We also assume that there exists a ball B such that $Q \subset B \subset 4B \subset U_{\tilde y}$. Then there exist $\varepsilon_c > 0$ small, depending only on the cost (and given by Lemma 6.1), and a constant $C(n) > 0$ depending only on the dimension, such that the following holds:*

If $\operatorname{diam}(Q) \le \varepsilon'_c := \varepsilon_c/C(n)$, *then*

$$|h^{\tilde c}(\tilde q)|^n \le C(n)\frac{\min\{\operatorname{dist}(\tilde q,\Pi^+),\operatorname{dist}(\tilde q,\Pi^-)\}}{\ell_{\Pi^+}}|\partial h^{\tilde c}|(\{\tilde q\})\mathscr{L}^n(Q), \tag{6.16}$$

where ℓ_{Π^+} denotes the maximal length among all the segments obtained by intersecting Q with a line orthogonal to Π^+.

Proof. We fix $\tilde{q} \in \operatorname{int} Q$. Let Π^i, $i = 1, \cdots n$, (with Π^1 equal either Π^+ or Π^-) be hyperplanes contained in $T^*_{\tilde{y}}V \setminus Q \simeq \mathbf{R}^n \setminus Q$, touching ∂Q, and such that $\{\Pi^+, \Pi^2, \ldots, \Pi^n\}$ are all mutually orthogonal (so that $\{\Pi^-, \Pi^2, \ldots, \Pi^n\}$ are also mutually orthogonal). Moreover we choose $\{\Pi^2, \ldots, \Pi^n\}$ in such a way that, if $\pi^1(Q)$ denotes the projection of Q on Π^1 and $\mathscr{H}^{n-1}(\pi^1(Q))$ denotes its $(n-1)$-dimensional Hausdorff measure, then

$$C(n)\mathscr{H}^{n-1}(\pi^1(Q)) \geq \prod_{i=2}^{n} \operatorname{dist}(\tilde{q}, \Pi^i), \tag{6.17}$$

for some universal constant $C(n)$. Indeed, as $\pi^1(Q)$ is convex, by Lemma 3.7 we can find an ellipsoid E such that $E \subset \pi^1(Q) \subset (n-1)E$, and for instance we can choose $\{\Pi^2, \ldots, \Pi^n\}$ among the hyperplanes orthogonal to the axes of the ellipsoid (for each axis we have two possible hyperplanes, and we can choose either of them).

Each hyperplane Π^i touches Q from outside, say at $q^i \in T^*_{\tilde{y}}V$. Let $p_i \in T_{\tilde{y}}V$ be the outward (from Q) unit vector at q^i orthogonal to Π^i. Then $s_i p_i \in \partial h^{\tilde{c}}(q^i)$ for some $s_i > 0$, and by Corollary 3.4 there exists $y_i \in \partial^{\tilde{c}} h^{\tilde{c}}(q^i)$ such that

$$-D_q\tilde{c}(q^i, y_i) = s_i p_i.$$

Define $y_i(t)$ as

$$-D_q\tilde{c}(q^i, y_i(t)) = t\, s_i p_i,$$

i.e. $y_i(t)$ is the $\tilde{c}$-segment from $\tilde{y}$ to y_i with respect to q^i. As in the proof of Lemma 6.7(c), there exists $0 < t_i \leq 1$ such that the function

$$m_{y_i(t_i)}(\cdot) := -\tilde{c}(\cdot, y_i(t_i)) + \tilde{c}(\tilde{q}, y_i(t_i)) + h^{\tilde{c}}(\tilde{q})$$

satisfies

$$m_{y_i(t_i)} \leq 0 \qquad \text{on } Q \text{ with equality at } q^i, \tag{6.18}$$

see Figure 5. By the definition of $h^{\tilde{c}}$, (6.18) implies $y_i(t_i) \in \partial^{\tilde{c}} h^{\tilde{c}}(\tilde{q}) \cap \partial^{\tilde{c}} h^{\tilde{c}}(q^i)$,

$$-D_q\tilde{c}(\tilde{q}, y_i(t_i)) \in \partial h^{\tilde{c}}(\tilde{q}) \quad \text{and} \quad t_i s_i p_i = -D_q\tilde{c}(q^i, y_i(t_i)) \in \partial h^{\tilde{c}}(q^i).$$

Note that the sub-level set $S_{y_i(t_i)} := \{z \in U_{\tilde{y}} \mid m_{y_i(t_i)} \leq 0\}$ is convex. Draw from $\tilde{q}$ a half-line orthogonal to Π^i. Let $\tilde{q}^i$ be point where this line meets with the boundary $\partial S_{y_i(t_i)}$. By the assumption $Q \subset B \subset 4B \subset U_{\tilde{y}}$, we see $\tilde{q}^i \in U_{\tilde{y}}$. By convexity of $S_{y_i(t_i)}$, we have

$$|\tilde{q} - \tilde{q}^i| \leq \operatorname{dist}(\tilde{q}, \Pi^i). \tag{6.19}$$

Let $\operatorname{diam} Q \leq \delta_n \varepsilon_c$ for some small constant $0 < \delta_n < 1/3$ to be fixed. By (6.2) and the trivial inequality $\operatorname{dist}(\tilde{q}, \Pi^i) \leq \operatorname{diam} Q$, we have

$$\begin{aligned} \big| - D_q\tilde{c}(\tilde{q}, y_i(t_i)) + D_q\tilde{c}(\ell\tilde{q} + (1-\ell)\tilde{q}^i, y_i(t_i))\big| &\leq \frac{1}{\varepsilon_c}|\tilde{q} - q^i|\,\big| - D_q\tilde{c}(\tilde{q}, y_i(t_i)\big| \\ &\leq \delta_n\big| - D_q\tilde{c}(\tilde{q}, y_i(t_i))\big| \end{aligned} \tag{6.20}$$

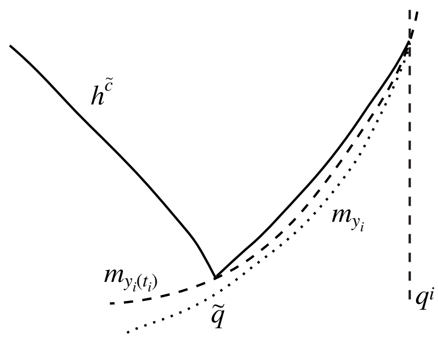

Figure 5: The dotted line represents the graph of $m_{y_i} := -\tilde{c}(\cdot, y_i) + \tilde{c}(\tilde{q}, y_i) + h^{\tilde{c}}(\tilde{q})$, while the dashed one represents the graph of $m_{y_i(t_i)} := -\tilde{c}(\cdot, y_i(t_i)) + \tilde{c}(\tilde{q}, y_i(t_i)) + h^{\tilde{c}}(\tilde{q})$. The idea is that, whenever we have m_{y_i} a supporting function for $h^{\tilde{c}}$ at a point $q^i \in \partial Q$, we can let y vary continuously along the $\tilde{c}$-segment from $\tilde{y}$ to y_i with respect to q^i, to obtain a supporting function $m_{y_i(t_i)}$ which touches $h^{\tilde{c}}$ also at $\tilde{q}$ as well as q^i.

for all $0 \leq \ell \leq 1$. Therefore, we see

$$\begin{aligned}
|h^{\tilde{c}}(\tilde{q})| &= |\tilde{c}(\tilde{q}, y_i(t_i)) - \tilde{c}(\tilde{q}^i, y_i(t_i))| \\
&= \Big| \int_0^1 -D_q\tilde{c}\big(\ell\tilde{q} + (1-\ell)\tilde{q}^i, y_i(t_i)\big) \cdot (\tilde{q} - \tilde{q}^i) d\ell \Big| \\
&\leq (1+\delta_n)\big| - D_q\tilde{c}(\tilde{q}, y_i(t_i))\big|\big|\tilde{q} - \tilde{q}^i\big| \qquad \text{(by (6.20))} \\
&\leq (1+\delta_n)\big| - D_q\tilde{c}(\tilde{q}, y_i(t_i))\big| \operatorname{dist}(\tilde{q}, \Pi^i) \qquad \text{by (6.19)).}
\end{aligned}$$

Thanks to this estimate it follows

$$| - D_q\tilde{c}(\tilde{q}, y_i(t_i))| \geq \frac{|h^{\tilde{c}}(\tilde{q})|}{2\operatorname{dist}(\tilde{q}, \Pi^i)}. \tag{6.21}$$

Moreover, similarly to (6.20), we have

$$\big| - D_q\tilde{c}(q^i, y_i(t_i)) + D_q\tilde{c}(\tilde{q}, y_i(t_i))\big| \leq \delta_n\big| - D_q\tilde{c}(\tilde{q}, y_i(t_i))\big|. \tag{6.22}$$

Since the vectors $\{-D_q\tilde{c}(q^i, y_i(t_i))\}_{i=1}^n$ are mutually orthogonal, (6.22) implies that for δ_n small enough the convex hull of $\{-D_q\tilde{c}(\tilde{q}, y_i(t_i))\}_{i=1}^n \subset \partial h^{\tilde{c}}(\tilde{q})$ has measure of order

$$\prod_{i=1}^n \big| - D_q\tilde{c}(\tilde{q}, y_i(t_i))\big|.$$

Thus, by the lower bound (6.21) and the convexity of $\partial h^{\tilde{c}}(\tilde{q})$, we obtain

$$\mathscr{L}^n(\partial h^{\tilde{c}}(\tilde{q})) \geq C(n)\frac{|h^{\tilde{c}}(\tilde{q})|^n}{\prod_{i=1}^n \operatorname{dist}(\tilde{q}, \Pi^i)}.$$

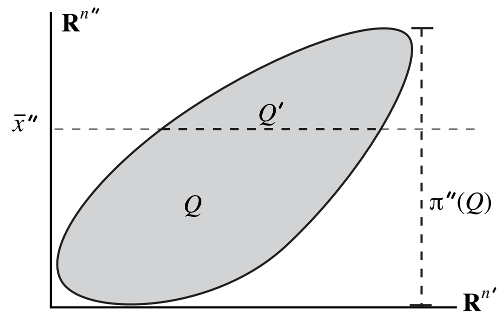

Figure 6: The volume of any convex set always controls the product (measure of one slice) · (measure of the projection orthogonal to the slice).

Since Π^1 was either Π^+ or Π^-, we have proved that

$$|h^{\tilde{c}}(\tilde{q})|^n \leq C(n)|\partial h^{\tilde{c}}|(\{\tilde{q}\}) \min\{\mathrm{dist}(\tilde{q}, \Pi^+), \mathrm{dist}(\tilde{q}, \Pi^-)\} \prod_{i=2}^{n} \mathrm{dist}(\tilde{q}, \Pi^i).$$

To conclude the proof, we apply Lemma 6.9 below with Q' given by the segment obtained intersecting Q with a line orthogonal to Π^+. Combining that lemma with (6.17), we obtain

$$C(n)\mathscr{L}^n(Q) \geq \ell_{\Pi^+} \prod_{i=2}^{n} \mathrm{dist}(\tilde{q}, \Pi^i),$$

and last two inequalities prove the proposition (taking $C(n) \geq 1/\delta_n$ larger if necessary). □

Lemma 6.9 (Estimating a convex volume using one slice and an orthogonal projection)**.** *Let Q be a convex set in $\mathbf{R}^n = \mathbf{R}^{n'} \times \mathbf{R}^{n''}$. Let π', π'' denote the projections to the components $\mathbf{R}^{n'}$, $\mathbf{R}^{n''}$, respectively. Let Q' be a slice orthogonal to the second component, that is*

$$Q' = (\pi'')^{-1}(\bar{x}'') \cap Q \qquad \textit{for some } \bar{x}'' \in \pi''(Q).$$

Then there exists a constant $C(n)$, depending only on $n = n' + n''$, such that

$$C(n)\mathscr{L}^n(Q) \geq \mathscr{H}^{n'}(Q')\mathscr{H}^{n''}(\pi''(Q)),$$

where $\mathscr{H}^d$ denotes the d-dimensional Hausdorff measure.

Proof. Let $L : \mathbf{R}^{n''} \to \mathbf{R}^{n''}$ be an affine map with determinant 1 given by Lemma 3.7 such that $B_r \subset L(\pi''(Q)) \subset B_{n''r}$ for some $r > 0$. Then, if we extend L to the whole $\mathbf{R}^n$ as $\tilde{L}(x', x'') = (x', Lx'')$, we have $\mathscr{L}^n(L(Q)) = \mathscr{L}^n(Q)$, $\mathscr{H}^{n'}(\tilde{L}(Q')) = \mathscr{H}^{n'}(Q')$, and

$$\mathscr{H}^{n''}(\pi''(\tilde{L}(Q))) = \mathscr{H}^{n''}(L(\pi''(Q))) = \mathscr{H}^{n''}(\pi''(Q)).$$

Hence, we can assume from the beginning that $B_r \subset \pi''(Q) \subset B_{n''r}$. Let us now consider the point $\bar{x}''$, and we fix an orthonormal basis $\{\hat{e}_1, \ldots, \hat{e}_{n''}\}$ in $\mathbf{R}^{n''}$ such that $\bar{x}'' = c\hat{e}_1$ for some $c \leq 0$. Since $\{r\hat{e}_1, \ldots, r\hat{e}_{n''}\} \subset \pi''(Q)$, there exist points $\{x_1, \ldots, x_{n''}\} \subset Q$ such that $\pi''(x_i) = r\hat{e}_i$. Let C' denote the convex hull of Q' with x_1, and let V' denote the $(n'+1)$-dimensional strip obtained taking the convex hull of $\mathbf{R}^{n'} \times \{\bar{x}''\}$ with x_1. Observe that $C' \subset V'$, so

$$\mathscr{H}^{n'+1}(C') = \frac{1}{n'+1} \operatorname{dist}(x_1, \mathbf{R}^{n'} \times \{\bar{x}''\}) \mathscr{H}^{n'}(Q') \geq \frac{r}{n'+1} \mathscr{H}^{n'}(Q'). \tag{6.23}$$

We now remark that, since $\pi''(x_i) = r\hat{e}_i$ and $\hat{e}_i \perp V'$ for $i = 2, \ldots, n''$, we have $\operatorname{dist}(x_i, V') = r$ for all $i = 2, \ldots, n''$. Moreover, if $y_i \in V'$ denotes the closest point to x_i, then the segments joining x_i to y_i parallels $\hat{e}_i$, hence these segments are all mutually orthogonal, and they are all orthogonal to V' too. From this fact it is easy to see that, if we define the convex hull

$$C := \operatorname{co}(x_2, \ldots, x_{n''}, C'),$$

then, since $|x_i - y_i| = r$ for $i = 2, \ldots, n''$, by (6.23) and the inclusion $\pi''(Q) \subset B_{n''r} \subset \mathbf{R}^{n''}$ we get

$$\mathscr{L}^n(C) = \frac{(n'+1)!}{n!} \mathscr{H}^{n'+1}(C') r^{n''-1} \geq \frac{n'!}{n!} \mathscr{H}^{n'}(Q') r^{n''} \geq C(n) \mathscr{H}^{n'}(Q') \mathscr{H}^{n''}(\pi''(Q)).$$

This concludes the proof, as $C \subset Q$. □

6.3 Alexandrov type upper bounds

The next Alexandrov type lemma holds for localized sections Q of $\tilde{c}$-convex functions.

Lemma 6.10 (Alexandrov type estimate and lower barrier)**.** *Assume* **(B0)**–**(B3)***, and let $\tilde{u} : \overline{U}_{\tilde{y}} \longmapsto \mathbf{R}$ be a $\tilde{c}$-convex function from Theorem 4.3. Let Q denote the section $\{\tilde{u} \leq 0\} \subset \overline{U}_{\tilde{y}}$, assume $\tilde{u} = 0$ on ∂Q, and fix $\tilde{q} \in \operatorname{int} Q$. Let Π^+, Π^- be two parallel hyperplanes contained in $\mathbf{R}^n \setminus Q$ and touching ∂Q from two opposite sides. We also assume that there exists a ball B such that $Q \subset B \subset 4B \subset U_{\tilde{y}}$. Then there exist $\varepsilon_c' > 0$ (given by Proposition 6.8) such that, if $\operatorname{diam}(Q) \leq \varepsilon_c'$ then*

$$|\tilde{u}(\tilde{q})|^n \leq C(n) \gamma^+_{\tilde{c}}(Q \times V) \frac{\min\{\operatorname{dist}(\tilde{q}, \Pi^+), \operatorname{dist}(\tilde{q}, \Pi^-)\}}{\ell_{\Pi^+}} |\partial^{\tilde{c}} \tilde{u}|(Q) \mathscr{L}^n(Q),$$

where ℓ_{Π^+} denotes the maximal length among all the segments obtained by intersecting Q with a line orthogonal to Π^+, and $\gamma^+_{\tilde{c}}(Q \times V)$ is defined as in (4.4).

Proof. Fix $\tilde{q} \in Q$. Observe that $\tilde{u} = 0$ on ∂Q and consider the $\tilde{c}$-cone $h^{\tilde{c}}$ generated by $\tilde{q}$ and Q of height $-h^{\tilde{c}}(\tilde{q}) = -\tilde{u}(\tilde{q})$ as in (6.15). From Lemma 6.7(d) we have

$$|\partial^{\tilde{c}} h^{\tilde{c}}|(\{\tilde{q}\}) \leq |\partial^{\tilde{c}} \tilde{u}|(Q),$$

and from Loeper's local to global principle, Corollary 3.4 above,

$$\partial h^{\tilde{c}}(\tilde{q}) = -D_q \tilde{c}(\tilde{q}, \partial^{\tilde{c}} h^{\tilde{c}}(\tilde{q})).$$

Therefore,

$$|\partial h^{\tilde{c}}|(\{\tilde{q}\}) \leq \| \det D^2_{qy} \tilde{c}\|_{C^0(\{\tilde{q}\}\times V)} |\partial^c h^c|(\{\tilde{q}\}).$$

The lower bound on $|\partial h^{\tilde{c}}|(\{\tilde{q}\})$ comes from (6.16). This finishes the proof. □

Combining this with Theorem 3.8, we get the following important estimates:

Theorem 6.11 (Alexandrov type upper bound). *Assume* **(B0)**–**(B3)***, and let* $\tilde{u} : \overline{U}_{\tilde{y}} \longmapsto \mathbf{R}$ *be a* $\tilde{c}$*-convex function from Theorem 4.3. There exist* $\varepsilon'_c > 0$ *small, depending only on the dimension and the cost function, and constant* $C(n)$*, depending only on the dimension, such that the following holds:*

Letting Q *denote the section* $\{\tilde{u} \leq 0\} \subset U_{\tilde{y}}$*, assume* $|\partial^{\tilde{c}} \tilde{u}| \leq 1/\lambda$ *in* Q *and* $\tilde{u} = 0$ *on* ∂Q*. We also assume that there exists a ball* B *such that* $Q \subset B \subset 4B \subset U_{\tilde{y}}$*, and that* $\mathrm{diam}(Q) \leq \varepsilon'_c$*. For* $\frac{1}{2n} < t \leq 1$*, let* $q_t \in Q$ *be a point such that* $q_t \in t\partial Q$*, where* $t\partial Q$ *denotes the dilation with the factor* t *with respect to the center of* E*, the ellipsoid given by John's Lemma (see* (3.9)*). Then,*

$$|\tilde{u}(q_t)|^n \leq C(n) \frac{\gamma^+_{\tilde{c}}}{\lambda} (1-t)^{1/2^{n-1}} \mathscr{L}^n(Q)^2, \tag{6.24}$$

where $\gamma^+_{\tilde{c}} = \gamma^+_{\tilde{c}}(Q \times V)$ *is defined as in* (4.4)*, which satisfies* $\gamma^+_{\tilde{c}} \leq \gamma^+_c \gamma^-_c$ *from Corollary 4.5. Moreover,*

$$\frac{|\inf_Q \tilde{u}|^n}{\mathscr{L}^n(Q)^2} \leq \ C(n) \frac{\gamma^+_{\tilde{c}}}{\lambda}. \tag{6.25}$$

Remark 6.12. Aficionados of the Monge-Ampère theory may be less surprised by these estimates once it is recognized that the localization in coordinates ensures the cost is approximately affine, at least in one of its two variables. However, no matter how well we approximate, the non-affine nature of the cost function remains relevant and persistent. Controlling the departure from affine is vital to our analysis and requires the new ideas developed above.

Proof of Theorem 6.11. For $0 < s_0 \leq t \leq 1$, let $q_t \in Q$ be a point such that $q_t \in t\partial Q$. By Theorem 3.8 applied with $s_0 = 1/(2n)$ we can find $\Pi^+ \neq \Pi^-$ parallel hyperplanes contained in $T^*_{\tilde{y}} V \setminus Q$, supporting Q from two opposite sides, and such that

$$\frac{\min\{\mathrm{dist}(\tilde{q}, \Pi^+), \mathrm{dist}(\tilde{q}, \Pi^-)\}}{\ell_{\Pi^+}} \leq C(n)(1-t)^{1/2^{n-1}}.$$

Then (6.24) follows from Lemma 6.10 and the assumption $|\partial^{\tilde{c}} \tilde{u}| \leq 1/\lambda$.

To prove (6.25), observe that, since $\frac{1}{2n} Q \subset \frac{1}{2} E$,

$$\frac{|\inf_{Q \setminus (\frac{1}{2}E)} \tilde{u}|^n}{\mathscr{L}^n(Q)^2} \leq \ C(n) \frac{\gamma^+_{\tilde{c}}}{\lambda} (1 - \frac{1}{2n})^{1/2^{n-1}}.$$

On the other hand, taking Π^+ and Π^- orthogonal to one of the longest axes of E and choosing $q \in \frac{1}{2}E$ in Lemma 6.10 yields

$$|\tilde{u}(q)|^n \leq C(n)\frac{\gamma_{\tilde{c}}^+}{\lambda} n\mathscr{L}^n(Q)^2, \qquad \forall q \in \tfrac{1}{2}E.$$

Combining these two estimates we obtain (6.25) to complete the proof. □

7 The contact set is either a single point or crosses the domain

The previous estimates (Theorems 6.2 and 6.11) may have some independent interest, but they also provide key ingredients which we use to deduce injectivity and Hölder continuity of optimal maps. In this and the subsequent section, we prove the strict c-convexity of the c-convex optimal transport potentials $u : \overline{U} \longmapsto \mathbf{R}$, meaning $\partial^c u(x)$ should be disjoint from $\partial^c u(\tilde{x})$ whenever $x, \tilde{x} \in U^\lambda$ are distinct. This shows the injectivity of optimal maps, and is accomplished in Theorem 8.1. In the present section we show that, if the contact set does not consist of a single point, then it extends to the boundary of U. Our method relies on the **(B3)** assumption on the cost c.

Recall that a point x of a convex set $S \subset \mathbf{R}^n$ is *exposed* if there is a hyperplane supporting S exclusively at x. Although the *contact set* $S := \partial^{c^*} u^{c^*}(\tilde{y})$ may not be convex, it appears convex from $\tilde{y}$ by Corollary 3.4, meaning its image $q(S) \subset U_{\tilde{y}}$ in the coordinates (4.1) is convex. The following theorem shows this convex set is either a singleton, or contains a segment which stretches across the domain. We prove it by showing the solution geometry near certain exposed points of $q(S)$ inside $U_{\tilde{y}}$ would be inconsistent with the bounds (Theorem 6.2 and 6.11) established in the previous section.

Theorem 7.1 (The contact set is either a single point or crosses the domain)**.** *Assume* **(B0)**–**(B3)***, and let u be a c-convex solution of* (3.8) *with $U^\lambda \subset U$ open. Fix $\tilde{x} \in U^\lambda$ and $\tilde{y} \in \partial^c u(\tilde{x})$, and define the contact set $S := \{x \in \overline{U} \mid u(x) = u(\tilde{x}) - c(x, \tilde{y}) + c(\tilde{x}, \tilde{y})\}$. Assume that $S \neq \{\tilde{x}\}$, i.e. it is not a singleton. Then S intersects ∂U.*

Proof. To derive a contradiction, we assume that $S \neq \{\tilde{x}\}$ and $S \subset\subset U$ (i.e. $S \cap \partial U = \emptyset$).

As in Definition 4.1, we transform $(x, u) \longmapsto (q, \tilde{u})$ with respect to $\tilde{y}$, i.e. we consider the transformation $q \in \overline{U}_{\tilde{y}} \longmapsto x(q) \in \overline{U}$, defined on $\overline{U}_{\tilde{y}} := -D_y c(\overline{U}, \tilde{y}) \subset T^*_{\tilde{y}}V$ by the relation

$$-D_y c(x(q), \tilde{y}) = q,$$

and the modified cost function $\tilde{c}(q, y) := c(x(q), y) - c(x(q), \tilde{y})$ on $\overline{U}_{\tilde{y}} \times \overline{V}$, for which the $\tilde{c}$-convex potential function $q \in \overline{U}_{\tilde{y}} \longmapsto \tilde{u}(q) := u(x(q)) + c(x(q), \tilde{y})$ is level-set convex. We observe that $\tilde{c}(q, \tilde{y}) \equiv 0$ for all q, and moreover the set $S = \partial^{c^*} u^{c^*}(\tilde{y})$ appears convex from $\tilde{y}$, meaning $S_{\tilde{y}} := -D_y c(S, \tilde{y})$ is convex, by Corollary 3.4.

Our proof is reminiscent of Caffarelli's for the cost $\tilde{c}(q, y) = -\langle q, y\rangle$ [8, Lemma 3]. Observe $\tilde{q} := -D_y c(\tilde{x}, \tilde{y})$ lies in the interior of the set $U^\lambda_{\tilde{y}} := -D_y c(U^\lambda, \tilde{y})$ where $|\partial^{\tilde{c}} \tilde{u}| \in [\lambda/\gamma_c^+, \gamma_c^-/\lambda]$, according to Corollary 4.5. Choose the point $q^0 \in S_{\tilde{y}} \subset\subset U_{\tilde{y}}$ furthest from $\tilde{q}$; it is an exposed

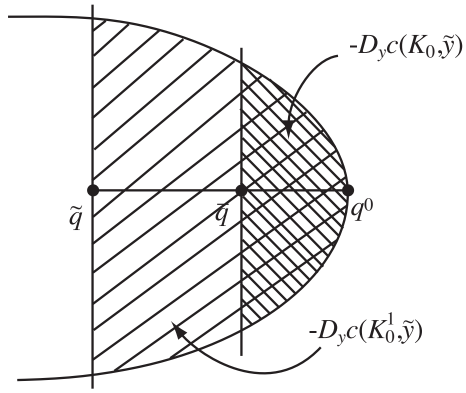

Figure 7: If the contact set $S_{\tilde{y}}$ has an exposed point q_0, we can cut two portions of $S_{\tilde{y}}$ with two hyperplanes orthogonal to $\tilde{q} - q^0$. The diameter of $-D_yc(K_0, \tilde{y})$ needs to be sufficiently small to apply the Alexandrov estimates from Theorem 6.11, while $-D_yc(K_0^1, \tilde{y})$ has to intersect $U^\lambda_{\tilde{y}}$ in some nontrivial set to apply Theorem 6.2 (this is not automatic, as q^0 may not be an interior point of $\operatorname{spt}|\partial^{\tilde{c}} u|$).

point of $S_{\tilde{y}}$. We will see below that the presence of such exposed point gives a contradiction, proving the theorem.

Before we proceed further, note that because of our assumption $S \subset\subset U$, the sets in the following argument, which will be sufficiently close to S, are also contained in U; the same holds for the corresponding sets in different coordinates. This is to make sure that we can perform the analysis using the assumptions on c.

For a suitable choice of Cartesian coordinates on V we may, without loss of generality, take $q^0 - \tilde{q}$ parallel to the positive y^1 axis. Denote by $\hat{e}_i$ the associated orthogonal basis for $T_{\tilde{y}}V$, and set $b^0 := \langle q^0, \hat{e}_1 \rangle$ and $\tilde{b} := \langle \tilde{q}, \hat{e}_1 \rangle$, so the halfspace $\{q \in T^*_{\tilde{y}}V \simeq \mathbf{R}^n \mid q_1 := \langle q, \hat{e}_1 \rangle \geq b^0\}$ intersects $S_{\tilde{y}}$ only at q^0. Use the fact that q^0 is an exposed point of $S_{\tilde{y}}$ to cut a corner K_0 off the contact set S by choosing $\bar{s} > 0$ small enough that $\bar{b} = (1 - \bar{s})b^0 + \bar{s}\tilde{b}$ satisfies:

(i) $-D_yc(K_0, \tilde{y}) := S_{\tilde{y}} \cap \{q \in \overline{U}_{\tilde{y}} \mid q_1 \geq \bar{b}\}$ is a compact convex set in the interior of $U_{\tilde{y}}$;

(ii) $\operatorname{diam}(-D_yc(K_0, \tilde{y})) \leq \varepsilon'_c/2$, where ε'_c is from Theorem 6.11.

(iii) $-D_yc(K_0, \tilde{y}) \subset B \subset 5B \subset U_{\tilde{y}}$ for some ball B as in the assumptions of Theorem 6.11.

Defining $q^s := (1 - s)q^0 + s\tilde{q}$, $x^s := x(q^s)$ the corresponding c-segment with respect to $\tilde{y}$, and $\bar{q} = q^{\bar{s}}$, note that $S_{\tilde{y}} \cap \{q_1 = \bar{b}\}$ contains $\bar{q}$, and K_0 contains $\bar{x} := x^{\bar{s}}$ and x^0. Since the corner K_0 may not intersect the support of $|\partial^c u|$ (especially, when q^0 is not an interior point of $\operatorname{spt}|\partial^c u|$), we shall need to cut a larger corner K_0^1 as well, defined by $-D_yc(K_0^1, \tilde{y}) := S_{\tilde{y}} \cap \{q \in \overline{U}_{\tilde{y}} \mid q_1 \geq \tilde{b}\}$, which intersects U^λ at $\tilde{x}$. By tilting the supporting function slightly, we shall now define sections $K_\varepsilon \subset K_\varepsilon^1$ of u whose interiors include the extreme point x^0

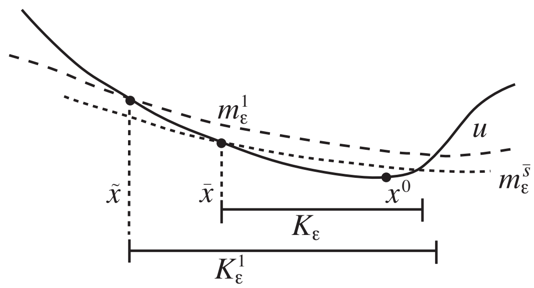

Figure 8: We cut the graph of u with the two functions $m_\varepsilon^{\bar{s}}$ and m_ε^1 to obtain two sets $K_\varepsilon \approx K_0$ and $K_\varepsilon^1 \approx K_0^1$ inside which we can apply our Alexandrov estimates to get a contradiction (Theorem 6.11 to K_ε, and Theorem 6.2 to K_ε^1). The idea is that the value of $u - m_\varepsilon^{\bar{s}}$ at x_0 is comparable to its minimum inside K_ε, but this is forbidden by our Alexandrov estimates since x_0 is too close to the boundary of K_0^ε. However, to make the argument work we need also to take advantage of the section K_ε^1, in order to "capture" some positive mass of the c-Monge-Ampère measure.

and whose boundaries pass through $\bar{x}$ and $\tilde{x}$ respectively, but which converge to K_0 and K_0^1 respectively as $\varepsilon \to 0$.

Indeed, set $y_\varepsilon := \tilde{y} + \varepsilon \hat{e}_1$ and observe

$$\begin{aligned} m_\varepsilon^s(x) &:= -c(x, y_\varepsilon) + c(x, \tilde{y}) + c(x^s, y_\varepsilon) - c(x^s, \tilde{y}) \\ &= \varepsilon\langle -D_y c(x, \tilde{y}) + D_y c(x^s, \tilde{y}), \hat{e}_1 \rangle + o(\varepsilon) \\ &= \varepsilon(\langle -D_y c(x, \tilde{y}), \hat{e}_1 \rangle - (1-s)b^0 - s\tilde{b}) + o(\varepsilon). \end{aligned} \tag{7.1}$$

Taking $s \in \{\bar{s}, 1\}$ in this formula and $\varepsilon > 0$ shows the sections defined by

$$\begin{aligned} K_\varepsilon &:= \{x \mid u(x) \le u(\bar{x}) - c(x, y_\varepsilon) + c(\bar{x}, y_\varepsilon)\}, \\ K_\varepsilon^1 &:= \{x \mid u(x) \le u(\tilde{x}) - c(x, y_\varepsilon) + c(\tilde{x}, y_\varepsilon)\}, \end{aligned}$$

both include a neighbourhood of x_0 but converge to K_0 and K_0^1 respectively as $\varepsilon \to 0$.

We remark that there exist a priori no coordinates in which all sets K_ε are simultaneously convex. However for each fixed $\varepsilon > 0$, we can change coordinates so that both K_ε and K_ε^1 become convex: use y_ε to make the transformations

$$\begin{aligned} q &:= -D_y c(x_\varepsilon(q), y_\varepsilon), \\ \tilde{c}_\varepsilon(q, y) &:= c(x_\varepsilon(q), y) - c(x_\varepsilon(q), y_\varepsilon), \end{aligned}$$

so that the functions

$$\begin{aligned} \tilde{u}_\varepsilon(q) &:= u(x_\varepsilon(q)) + c(x_\varepsilon(q), y_\varepsilon) - u(\bar{x}) - c(\bar{x}, y_\varepsilon), \\ \tilde{u}_\varepsilon^1(q) &:= u(x_\varepsilon(q)) + c(x_\varepsilon(q), y_\varepsilon) - u(\tilde{x}) - c(\tilde{x}, y_\varepsilon). \end{aligned}$$

are level-set convex on $U_{y_\varepsilon} := D_y c(U, y_\varepsilon)$. Observe that, in these coordinates, K_ε and K_ε^1 become convex:

$$\tilde{K}_\varepsilon := -D_y c(K_\varepsilon, y_\varepsilon) = \{q \in \overline{U}_{y_\varepsilon} \mid \tilde{u}_\varepsilon(q) \leq 0\},$$
$$\tilde{K}_\varepsilon^1 := -D_y c(K_\varepsilon^1, y_\varepsilon) = \{q \in \overline{U}_{y_\varepsilon} \mid \tilde{u}_\varepsilon^1(q) \leq 0\},$$

and either $\tilde{K}_\varepsilon \subset \tilde{K}_\varepsilon^1$ or $\tilde{K}_\varepsilon^1 \subset \tilde{K}_\varepsilon$ since $\tilde{u}_\varepsilon(q) - \tilde{u}_\varepsilon^1(q) = const$. For $\varepsilon > 0$ small, the inclusion must be the first of the two since the limits satisfy $\tilde{K}_0 \subset \tilde{K}_0^1$ and $\tilde{q} \in \tilde{K}_0^1 \setminus \tilde{K}_0$.

In the new coordinates, our original point $\tilde{x} \in U^\lambda$, the exposed point x^0, and the c-convex combination $\bar{x}$ with respect to $\tilde{y}$, correspond to

$$\tilde{q}_\varepsilon := -D_y c(\tilde{x}, y_\varepsilon), \quad q_\varepsilon^0 := -D_y c(x^0, y_\varepsilon), \quad \bar{q}_\varepsilon := -D_y c(\bar{x}, y_\varepsilon).$$

Thanks to (ii) and (iii), for ε sufficiently small we have $\operatorname{diam}(\tilde{K}_\varepsilon) \leq \varepsilon_c'$ and $\tilde{K}_\varepsilon \subset B \subset 4B \subset U_{y_\varepsilon}$ for some ball B (note that this ball can be different from the one in (iii)), so that all the estimates of Theorem 6.11 apply.

Let us observe that, since $\lim_{\varepsilon\to 0} q_\varepsilon^0 - \bar{q}_\varepsilon = q^0 - \bar{q}$ and $q^0 \in \partial Q$, (6.24) combines with $K_\varepsilon \subset K_\varepsilon^1$ and $|\partial^{\tilde{c}_\varepsilon} \tilde{u}_\varepsilon|(K_\varepsilon) \leq \Lambda \gamma_c^- \mathscr{L}^n(K_\varepsilon)$ from (3.8) and Corollary 4.5, to yield

$$\frac{|\tilde{u}_\varepsilon(q_\varepsilon^0)|^n}{\Lambda \gamma_c^- \mathscr{L}^n(\tilde{K}_\varepsilon^1)^2} \to 0 \qquad \text{as } \varepsilon \to 0. \tag{7.2}$$

On the other hand, $\bar{x} \in S$ implies $\tilde{u}_\varepsilon(q_\varepsilon^0) = m_\varepsilon^{\bar{s}}(x^0)$, and $\tilde{x} \in S$ implies $\tilde{u}_\varepsilon^1(q_\varepsilon^0) = m_\varepsilon^1(x^0)$ similarly. Thus (7.1) yields

$$\frac{\tilde{u}_\varepsilon(q_\varepsilon^0)}{\tilde{u}_\varepsilon^1(q_\varepsilon^0)} = \frac{\varepsilon(b^0 - \bar{b}) + o(\varepsilon)}{\varepsilon(b^0 - \tilde{b}) + o(\varepsilon)} \to \bar{s} \quad \text{as} \quad \varepsilon \to 0. \tag{7.3}$$

Our contradiction with (7.2)–(7.3) will be established by bounding the ratio $|\tilde{u}^1(q_\varepsilon^0)|^n / \mathscr{L}^n(K_\varepsilon^1)^2$ away from zero.

Recall that

$$b^0 = \langle -D_y c(x^0, \tilde{y}), \hat{e}_1 \rangle \;=\; \max\{q_1 \mid q \in -D_y c(K_0, \tilde{y})\}) \;>\; \tilde{b}$$

and $u(x) - u(\tilde{x}) \geq -c(x, \tilde{y}) + c(\tilde{x}, \tilde{y})$ with equality at x^0. From the convergence of K_ε^1 to K_0^1 and the asymptotic behaviour (7.1) of $m_\varepsilon^1(x)$ we get

$$\begin{aligned}
\frac{\tilde{u}_\varepsilon^1(q_\varepsilon^0)}{\inf_{\tilde{K}_\varepsilon^1} \tilde{u}_\varepsilon^1} &= \frac{-u(x^0) - c(x^0, y_\varepsilon) + u(\tilde{x}) + c(\tilde{x}, y_\varepsilon)}{\sup_{q \in \tilde{K}_\varepsilon^1}[-u(x(q)) - c(x(q), y_\varepsilon) + u(\tilde{x}) + c(\tilde{x}, y_\varepsilon)]} \\
&\geq \frac{-c(x^0, y_\varepsilon) + c(\tilde{x}, y_\varepsilon) + c(x^0, \tilde{y}) - c(\tilde{x}, \tilde{y})}{\sup_{x \in K_\varepsilon^1}[-c(x, y_\varepsilon) + c(\tilde{x}, y_\varepsilon) + c(x, \tilde{y}) - c(\tilde{x}, \tilde{y})]} \\
&\geq \frac{\varepsilon(\langle -D_y c(x^0, \tilde{y}), e_1 \rangle - \tilde{b}) + o(\varepsilon)}{\varepsilon(\max\{q_1 \mid q \in -D_y c(K_\varepsilon^1, \tilde{y})\} - \tilde{b}) + o(\varepsilon)} \\
&\geq \frac{1}{2}
\end{aligned} \tag{7.4}$$

for ε sufficiently small (because, by our construction, $\max\{q_1 \mid q \in -D_y c(K^1_\varepsilon, \tilde{y})\}$ is exactly $\langle -D_y c(x^0,\tilde{y}), e_1\rangle$). This shows $\tilde{u}^1(q^0_\varepsilon)$ is comparable to the minimum value of $\tilde{u}^1_\varepsilon$. To conclude the proof we would like to apply Theorem 6.2, but we need to show that $|\partial^c u| \geq \lambda$ on a stable fraction of K^1_ε as $\varepsilon \to 0$. We shall prove this as in [8, Lemma 3].

Since K^1_ε converges to K^1_0 for sufficiently small ε, observe that K^1_ε (thus, $\tilde{K}^1_\varepsilon$) is bounded uniformly in ε. Therefore the affine transformation $(L^1_\varepsilon)^{-1}$ that sends $\tilde{K}^1_\varepsilon$ to $B_1 \subset \tilde{K}^{1,*}_\varepsilon \subset \overline{B}_n$ as in Lemma 3.7 is an expansion, i.e. $|(L^1_\varepsilon)^{-1}q - (L^1_\varepsilon)^{-1}q'| \geq C_0|q-q'|$, with a constant $C_0 > 0$ independent of ε. Since $\tilde{x}$ is an interior point of U^λ, $B_{2\beta_c^-\delta/C_0}(\tilde{x}) \subset U^\lambda$ for sufficiently small $\delta > 0$ (here β_c^- is from (4.3)), hence $B_{2\delta/C_0}(\tilde{q}_\varepsilon) \subset U^\lambda_{y_\varepsilon}$, where $U^\lambda_{y_\varepsilon} := -D_y c(U^\lambda, y_\varepsilon)$. Let $\tilde{q}^*_\varepsilon := (L^1_\varepsilon)^{-1}(\tilde{q}_\varepsilon)$. Then, by the expansion property of $(L^1_\varepsilon)^{-1}$, we have

$$U^{\lambda,*}_{y_\varepsilon} := (L^1_\varepsilon)^{-1}(U^\lambda_{y_\varepsilon}) \supset B_{2\delta}(\tilde{q}^*_\varepsilon).$$

Since $\tilde{K}^{1,*}_\varepsilon$ is convex, it contains the convex hull $\mathcal{C}$ of $B_1 \cup \{\tilde{q}^*_\varepsilon\}$. Consider $\mathcal{C} \cap B_{2\delta}(\tilde{q}^*_\varepsilon)$. Since $\operatorname{dist}(\mathbf{0}, \tilde{q}^*_\varepsilon) \leq n$, there exists a ball B^* of radius δ/n (centered somewhere in $\mathcal{C} \cap B_{2\delta}(\tilde{q}^*_\varepsilon)$) such that

$$B^* \subset \mathcal{C} \cap B_{2\delta}(\tilde{q}^*_\varepsilon) \subset B_{2\delta}(\tilde{q}^*_\varepsilon) \cap \tilde{K}^{1,*}_\varepsilon.$$

Therefore, the ellipsoid

$$E_\delta := L^1_\varepsilon(B^*)$$

is contained in $U^\lambda_{y_\varepsilon}$. Notice that E_δ is nothing but a dilation and translation of the ellipsoid $E = L^1_\varepsilon(B_1)$ associated to $\tilde{K}^1_\varepsilon$ by John's Lemma in (3.9). By dilating further if necessary (but, with a factor independent of ε), one may assume that $E_\delta \subset B \subset 4B \subset U_{y_\varepsilon}$ for some ball B (as before, this ball can be different from that for K_0 or $\tilde{K}_\varepsilon$). Thus, we can apply Theorem 6.2 (with $Q = \tilde{K}^1_\varepsilon$) and obtain

$$\frac{|\inf_{\tilde{K}^1_\varepsilon} \tilde{u}^1_\varepsilon|^n}{\mathscr{L}^n(\tilde{K}^1_\varepsilon)^2} \gtrsim \delta^{2n}$$

where the inequality $\gtrsim$ is independent of ε. As $\varepsilon \to 0$ this contradicts (7.2)–(7.4) to complete the proof. □

Remark 7.2. As can be easily seen from the proof, when $U^\lambda = U$ one can actually show that if S is not a singleton, then $S_{\tilde{y}}$ has no exposed points in the interior of $U_{\tilde{y}}$. Indeed, if by contradiction there exists q^0 an exposed point of $S_{\tilde{y}}$ belonging to the interior of $U_{\tilde{y}}$, we can choose a point $\tilde{q} \in S_{\tilde{y}}$ sufficiently close to q^0 in the interior of $U_{\tilde{y}} = U^\lambda_{\tilde{y}}$ such that the segment $q^0 - \tilde{q}$ is orthogonal to a hyperplane supporting $S_{\tilde{y}}$ at q^0. Then it can be immediately checked that the above proof (which could even be simplified in this particular case, since one may choose $K^1_0 = K_0$ and avoid the last part of the proof) shows that such a point q^0 cannot exist.

8 Continuity and injectivity of optimal maps

The first theorem below combines results of Sections 5 and 7 to deduce strict c-convexity of the c-potential for an optimal map, if its target is strongly c-convex. This strict c-convexity — which is equivalent to injectivity of the map — will then be combined with an adaptation of Caffarelli's argument [5, Corollary 1] to obtain interior continuity of the map — or equivalently C^1-regularity of its c-potential function — for **(B3)** costs.

Theorem 8.1 (Injectivity of optimal maps to a strongly c-convex target)**.** *Let c satisfy* **(B0)**–**(B3)** *and* $\mathbf{(B2)_s}$. *If u is a c-convex solution of* (3.8) *on $U^\lambda \subset U$ open, then u is strictly c-convex on U^λ, meaning $\partial^c u(x)$ and $\partial^c u(\tilde{x})$ are disjoint whenever $x, \tilde{x} \in U^\lambda$ are distinct.*

Proof. Suppose by contradiction that $\tilde{y} \in \partial^c u(x) \cap \partial^c u(\tilde{x})$ for two distinct points $x, \tilde{x} \in U^\lambda$, and set $S = \partial^{c^*} u^{c^*}(\tilde{y})$. According to Theorem 7.1, the set S intersects the boundary of U at a point $\bar{x} \in \partial U \cap \partial^{c^*} u^{c^*}(\tilde{y})$. Since (3.8) asserts $\lambda \le |\partial^c u|$ on U^λ and $|\partial^c u| \le \Lambda$ on $\overline{U}$, Theorem 5.1(a) yields $\tilde{y} \in V$ (since $x, \tilde{x} \in U^\lambda$), and hence $\bar{x} \in U$ by Theorem 5.1(b). This contradicts $\bar{x} \in \partial U$ and proves the theorem. □

Remark 8.2. As already mentioned at the end of Section 2, if $-D_y c(U, y) = \mathbf{R}^n$ for any $y \in V$ we can remove the strong c-convexity assumption $\mathbf{(B2)_s}$. Indeed, assume that u is not strictly c-convex, and fix $\tilde{y} \in \partial^c u(U)$ such that the dimension of the convex set $S_{\tilde{y}} := -D_y c(S, \tilde{y}) \subset \mathbf{R}^n$ is maximal, say equal to $k \in \{1, \dots, n\}$. Since this set has no exposed points (see Theorem 7.1), it has to contain a k-dimensional plane $P_{\tilde{y}}$.

Let $x_y(q)$ denote the inverse of $x \mapsto q(x) := -D_y c(x, y)$. Since the function $u_{\tilde{y}}(q) := u(x_{\tilde{y}}(q)) + c(x_{\tilde{y}}(q), \tilde{y})$ is level-set convex and attains its minimum on $S_{\tilde{y}}$, all its level sets are convex and contain $P_{\tilde{y}}$, hence they have to be of the form $\{u_{\tilde{y}} \le M\} = P_{\tilde{y}} \times K_M$, where K_M is a convex set in $P_{\tilde{y}}^\perp \simeq \mathbf{R}^{n-k}$. This implies that $u_{\tilde{y}}$ depends only on the $(n-k)$ variables in $P_{\tilde{y}}^\perp$. In particular, if we set $c_{\tilde{y}}(q, y) := c(x_{\tilde{y}}(q), y) - c(x_{\tilde{y}}(q), \tilde{y})$, then, for any $y \in \partial^{c_{\tilde{y}}} u_{\tilde{y}}(\mathbf{R}^n) = \partial^c u(\mathbf{R}^n)$, the contact set is always k-dimensional (since it contains a translate of $P_{\tilde{y}}$, and cannot be larger by maximality of k). By repeating the above considerations with a different y, we deduce that $u_y(q) := u(x_y(q)) + c(x_y(q), y)$ is constant along a k-dimensional plane P_y for any $y \in \partial^c u(U)$.

Now, given $\bar{q} \in -D_y c(U, \tilde{y}) = \mathbf{R}^n$ and $\bar{y} \in \partial^{c_{\tilde{y}}} u_{\tilde{y}}(\bar{q})$, consider the supporting function $m_{\bar{q},\bar{y}}(q) := -c_{\tilde{y}}(q, \bar{y}) + c_{\tilde{y}}(\bar{q}, \bar{y}) + u_{\tilde{y}}(\bar{q})$. Set $M := u_{\tilde{y}}(\bar{q})$. Since $m_{\bar{q},\bar{y}} \le u_{\tilde{y}}$ we have

$$\{m_{\bar{q},\bar{y}} \le M\} \supset \{\tilde{u} \le M\} = P_{\tilde{y}} \times K_M.$$

Hence, being $\{m_{\bar{q},\bar{y}} \le M\}$ convex, it must split as $P_{\tilde{y}} \times K_M^{\bar{q},\bar{y}}$, where $K_M^{\bar{q},\bar{y}} \subset P_{\tilde{y}}^\perp \simeq \mathbf{R}^{n-k}$.

We now observe that the set $\{m_{\bar{q},\bar{y}} \le M\}$ corresponds, in the original variables, to a set of the form $\{-c(x, \bar{y}) + c(x, \tilde{y}) \le L\}$. Hence, by a symmetric argument (exchanging the roles of $\tilde{y}$ and $\bar{y}$), the set

$$\{-c(x, \tilde{y}) + c(x, \bar{y}) \le -L\} = \mathbf{R}^n \setminus \{-c(x, \bar{y}) + c(x, \tilde{y}) \le L\}$$

is of the form $P_{\bar{y}} \times K_M^{\tilde{q},\tilde{y}}$ when we look at it in the coordinates $-D_y c(x, \bar{y})$. This implies that

$$-D_y c\big(x_{\tilde{y}}(q + P_{\tilde{y}}), \bar{y}\big) \text{ is a } k\text{-plane parallel to } P_{\bar{y}} \qquad \forall\, q \in \mathbf{R}^n,\ \forall\, \tilde{y}, \bar{y} \in \partial^c u(U). \tag{8.1}$$

In other words, if we consider the foliation induces by the k-plane $P_{\tilde{y}}$ in the system of coordinates $-D_y c(\cdot, \tilde{y})$ and we look at it from a different system of coordinates $-D_y c(\cdot, \bar{y})$, we simply obtain the foliation induced by $P_{\bar{y}}$

By differentiating the relation (8.1) at the point $q = \mathbf{0}$, we deduce the following: let $x_0 := x_{\tilde{y}}(\mathbf{0}) \in U$, and $e := -D^2_{xy} c(x_0, \tilde{y})^{-1} \cdot v$ for some unit vector v parallel to $P_{\tilde{y}}$. Then

$$- D^2_{xy} c(x_0, y) \cdot e \parallel P_y \qquad \forall\, y \in \partial^c u(U). \tag{8.2}$$

Recalling that u_y is constant along directions parallel to P_y, using (8.2) we obtain

$$0 = D_q u_y(q)|_{q = x_y(x_0)} \cdot \big(D^2_{xy} c(x_0, y) \cdot e\big) = \langle p_0 + D_x c(x_0, y), e \rangle \qquad \forall\, y \in \partial^c u(U),$$

where p_0 is any element in $\partial u(x_0)$. By **(B1)** this implies that $\partial^c u(U)$ is contained a $(n-1)$-dimensional manifold. Thus the c-Monge-Ampère measure of u is singular, contradicting the c-Monge-Ampère equation (3.8).

By adapting Caffarelli's argument [5, Corollary 1], we now show continuity of the optimal map. Although in the next section we will actually prove a stronger result (i.e., optimal maps to strongly c-convex targets are locally Hölder continuous), we prefer to prove this result for two reasons: first, the proof is much simpler than the one of Hölder continuity. Second, although not strictly necessary, knowing in advance that solutions of (3.8) are C^1 will avoid some technical issues in the proof of the $C^{1,\alpha}$ regularity.

Theorem 8.3 (Continuity of optimal maps to strongly c-convex targets)**.** *Let c satisfy* **(B0)**–**(B3)** *and* $\mathbf{(B2)_s}$. *If u is a c-convex solution of* (3.8) *on $U^\lambda \subset U$ open, then u is continuously differentiable inside U^λ.*

Proof. Recalling that c-convexity implies semiconvexity (see Section 3 and also (4.5)), all we need to show is that the c-subdifferential $\partial^c u(\tilde{x})$ of u at every point $\tilde{x} \in U_\lambda$ is a singleton. Notice that $\partial^c u(\tilde{x}) \subset\subset V$ by Theorem 5.1(a).

Assume by contradiction that is not. As $\partial^c u(\tilde{x})$ is compact, one can find a point y_0 in the set $\partial^c u(\tilde{x})$ such that $-D_x c(\tilde{x}, y_0) \in \partial u(\tilde{x})$ is an exposed point of the compact convex set $\partial u(\tilde{x})$. Similarly to Definition 4.1, we transform $(x, u) \longmapsto (q, \tilde{u})$ with respect to y_0, i.e. we consider the transformation $q \in \overline{U}_{y_0} \longmapsto x(q) \in \overline{U}$, defined on $\overline{U}_{y_0} = -D_y c(\overline{U}, y_0) + D_y c(\tilde{x}, y_0) \subset T^*_{y_0} V$ by the relation

$$-D_y c(x(q), y_0) + D_y c(\tilde{x}, y_0) = q,$$

and the modified cost function $\tilde{c}(q, y) := c(x(q), y) - c(x(q), y_0)$ on $\overline{U}_{y_0} \times \overline{V}$, for which the $\tilde{c}$-convex potential function $q \in \overline{U}_{y_0} \longmapsto \tilde{u}(q) := u(x(q)) - u(\tilde{x}) + c(x(q), y_0) - c(\tilde{x}, y_0)$ is level-set convex. We observe that $\tilde{c}(q, y_0) \equiv 0$ for all q, the point $\tilde{x}$ is sent to $\mathbf{0}$, $\tilde{u} \geq \tilde{u}(\mathbf{0}) = 0$, and $\tilde{u}$ is strictly $\tilde{c}$-convex thanks to Theorem 8.1. Moreover, since $-D_x c(\tilde{x}, y_0) \in \partial u(\tilde{x})$ was an exposed point of $\partial u(\tilde{x})$, $\mathbf{0} = -D_q \tilde{c}(\mathbf{0}, y_0)$ is an exposed point of $\partial \tilde{u}(\mathbf{0})$. Hence, we can find a vector $v \in \partial \tilde{u}(\mathbf{0}) \setminus \{\mathbf{0}\}$ such that the hyperplane orthogonal to v is a supporting hyperplane for $\partial \tilde{u}(\mathbf{0})$ at $\mathbf{0}$, see Figure 9. By the semiconvexity of $\tilde{u}$ (see (4.5)), this implies that

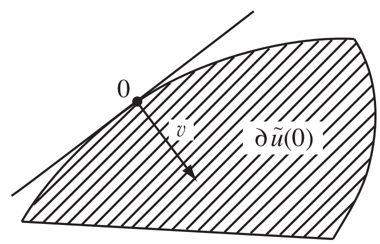

Figure 9: $v \in \partial\tilde{u}(\mathbf{0})$ and the hyperplane orthogonal to v is supporting $\partial\tilde{u}(\mathbf{0})$ at 0.

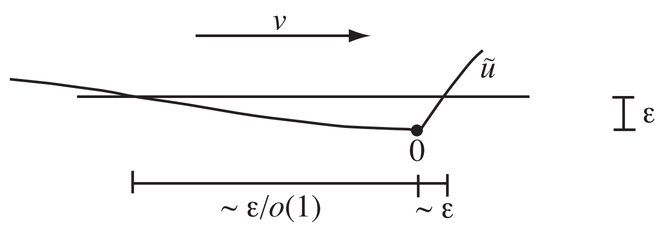

Figure 10: Since the hyperplane orthogonal to v is supporting $\partial\tilde{u}(\mathbf{0})$ at 0, we have $\tilde{u}(-tv) = o(t)$ for $t \geq 0$. Moreover, by the semiconvexity, $\tilde{u}$ grows at least linearly in the direction of v.

$$\tilde{u}(-tv) = o(t) \quad \text{for } t \geq 0, \qquad \tilde{u}(q) \geq \langle v, q\rangle - M_c|q|^2 \quad \text{for all } q \in U_{y_0}. \tag{8.3}$$

Let us now consider the (convex) section $K_\varepsilon := \{\tilde{u} \leq \varepsilon\}$. Recalling that $\mathbf{0}$ the unique minimum point with $\tilde{u}(\mathbf{0}) = 0$, we have that K_ε shrinks to $\mathbf{0}$ as $\varepsilon \to 0$, and $\tilde{u} > 0$ on $U_{y_0} \setminus \{\mathbf{0}\}$. Thus by (8.3) it is easily seen that for ε sufficiently small the following hold:

$$K_\varepsilon \subset \{q \mid \langle q, v\rangle \leq 2\varepsilon\}, \qquad -\alpha(\varepsilon)v \in K_\varepsilon,$$

where $\alpha(\varepsilon) > 0$ is a positive constant depending on ε and such that $\alpha(\varepsilon)/\varepsilon \to +\infty$ as $\varepsilon \to 0$. Since $\mathbf{0}$ is the minimum point of $\tilde{u}$, this immediately implies that one between our Alexandrov estimates (6.4) or (6.24) must be violated by $\tilde{u}$ inside K_ε for ε sufficiently small, which is the desired contradiction, see Figure 10. □

9 Engulfing property and Hölder continuity of optimal maps

Our goal is to prove the $C^{1,\alpha}_{loc}$ regularity of u inside U^λ. The proof will rely on the strict c-convexity of u inside U^λ (Theorem 8.1) and the Alexandrov type estimates from Theorems 6.2 and 6.11. Indeed, these results will enable us to show the engulfing property under the **(B3)** condition (Theorem 9.3), thus extending the result of Gutierrez and Huang [30, Theorem 2.2] given for the classical Monge-Ampère equation. Then, the engulfing property allows us to apply the method of Forzani and Maldonado [24] to obtain local $C^{1,\alpha}$ estimates (Theorem 9.5). Finally, by a covering argument, we show that the Hölder exponent α depends only on λ and n, and not on particular cost function (Corollary 9.6).

Remark 9.1. We point out that in this section the assumptions **(B2)** and $\mathbf{(B2)_s}$ are used only to ensure the strict c-convexity (Theorems 8.1). Since the following arguments are performed locally, i.e. after restricting to small neighborhoods, having already obtained the strict c-convexity of u, the c-convexity **(B2)** and $\mathbf{(B2)_s}$ of the ambient domains are not further required. This is a useful remark for the covering argument in Corollary 9.6.

Given points $\tilde{x} \in U^\lambda$ and $\tilde{y} \in \partial^c u(\tilde{x})$, and $\tau > 0$, we denote by $S(\tilde{x}, \tilde{y}, \tau)$ the section

$$S(\tilde{x}, \tilde{y}, \tau) := \{x \in U \mid u(x) \le u(\tilde{x}) - c(x, \tilde{y}) + c(\tilde{x}, \tilde{y}) + \tau\}. \tag{9.1}$$

Notice that by the strict c-convexity of u (Theorem 8.1), $S(\tilde{x}, \tilde{y}, \tau) \to \{\tilde{x}\}$ as $\tau \to 0$.

In the following, we assume that all points $x, \tilde{x}$ that we choose inside U^λ are close to each other, and "relatively far" from the boundary of U^λ, i.e.

$$\mathrm{dist}(x, \tilde{x}) \ll \min\big(\mathrm{dist}(x, \partial U^\lambda), \mathrm{dist}(\tilde{x}, \partial U^\lambda)\big).$$

This assumption ensures all the relevant sets, i.e. sections around x or $\tilde{x}$, stay strictly inside U^λ.

As we already did many times in the previous section, given a point $\tilde{y} \in \partial^c u(\tilde{x})$ with $\tilde{x} \in U^\lambda$, we consider the transformation $(x, u) \longmapsto (q, \tilde{u})$ with respect to $\tilde{y}$ (see Definition 4.1), and define the sections

$$Q_\tau = \{q \in T_{\tilde{y}}{}^*V \mid \tilde{u}(q) \le \tau\}, \qquad \tau \ge 0.$$

Note that Q_τ corresponds to $S(\tilde{x}, \tilde{y}, \tau)$ under the coordinate change. By the **(B3)** condition, each Q_τ is a convex set (Theorem 4.3). We also keep using the notation ρQ_τ to denote the dilation of Q_τ by a factor $\rho > 0$ with respect to the center of the ellipsoid given by John's Lemma (see (3.9)).

In our analysis, we will only need to consider the case $\tau \ll 1$. This is useful since, thanks to the strict c-convexity of u (Theorem 8.1), we can consider in the sequel only sections contained inside U^λ and sufficiently small. In particular, for every section $Q = S(x, y, \tau)$ or $Q = S(\tilde{x}, \tilde{y}, K\tau)$, after the transformation in Definition 4.1:

- we can apply Theorem 6.2 with $E_\delta = E$ (so (6.4) holds with $\delta = 1$);
- we can apply Lemma 6.3 with $\mathscr{K} = Q$;
- Theorem 6.11 holds.

The following result generalizes [30, Theorem 2.1(ii)] for $c(x, y) = -x \cdot y$ to **(B3)** costs:

Lemma 9.2 (Comparison of sections with different heights)**.** *Assume* **(B0)**-**(B3)** *and let u be a strictly c-convex solution to* (3.8) *on U^λ. Take $\tilde{x} \in U^\lambda$, $\tilde{y} \in \partial^c u(\tilde{x})$ and τ sufficiently small so that*

$$S_\tau = S(\tilde{x}, \tilde{y}, \tau) \subset U^\lambda$$

and set $Q_\tau := -D_y c(S_\tau, \tilde{y})$. Then, there exist $0 < \rho_0 < 1$, depending only on the dimension n and $\gamma_c^+ \gamma_c^- / \lambda$ (in particular independent of $\tilde{x}, \tilde{y}$ and t), such that

$$Q_{\tau/2} \subseteq \rho_0 Q_\tau.$$

Proof. This is a simple consequence of Theorems 6.2 and 6.11: indeed, considering $\tilde{u} - \tau$, for $\rho > 1/2n$ (6.24) gives

$$|\tilde{u}(q) - \tau|^n \leq C(n) \frac{\gamma_{\tilde{c}}^+}{\lambda} (1-\rho)^{2^{-n+1}} \mathscr{L}^n(Q_\tau)^2 \qquad \forall\, q \in Q_\tau \setminus \rho Q_\tau,$$

while by (6.4) (with $\delta = 1$)

$$\mathscr{L}^n(Q_\tau)^2 \leq C(n) \frac{\gamma_{\tilde{c}}^-}{\lambda} \tau^n.$$

Hence, we get

$$\begin{aligned} |\tilde{u}(q) - \tau| &\leq C(n) \frac{\gamma_{\tilde{c}}^+ \gamma_{\tilde{c}}^-}{\lambda^2} (1-\rho)^{2^{-n+1}} \tau^n \\ &\leq C(n) \left[\frac{\gamma_c^+ \gamma_c^-}{\lambda} \right]^2 (1-\rho)^{2^{-n+1}} \tau^n \end{aligned}$$

where the last inequality follows from $\gamma_{\tilde{c}}^+ \gamma_{\tilde{c}}^- / \lambda^2 \leq \left[\gamma_c^+ \gamma_c^- / \lambda \right]^2$ as in Corollary 4.5. Therefore, for $1 - \rho_0$ sufficiently small (depending only on the dimension n and $\gamma_c^+ \gamma_c^- / \lambda$) we get

$$|\tilde{u}(q) - \tau| \leq \tau/2 \qquad \forall\, q \in Q_\tau \setminus \rho_0 Q_\tau,$$

so $Q_{\tau/2} \subseteq \rho_0 Q_\tau$ as desired. □

Thanks to Lemma 9.2 and Lemma 6.3, we can prove the engulfing property under **(B3)**, extending the classical Monge-Ampère case of [30, Theorem 2.2]:

Theorem 9.3 (Engulfing)**.** *Assume* **(B0)**-**(B3)** *and let u be a strictly c-convex solution to* (3.8) *on U^λ. Let $x, \tilde{x} \in U^\lambda$ be close: i.e.*

$$\mathrm{dist}(x, \tilde{x}) \ll \min\big(\mathrm{dist}(x, \partial U^\lambda), \mathrm{dist}(\tilde{x}, \partial U^\lambda) \big).$$

Then, there exists a constant $K > 1$, depending only on the dimension n and $\gamma_c^+ \gamma_c^- / \lambda$, such that, for all $y \in \partial^c u(x)$, $\tilde{y} \in \partial^c u(\tilde{x})$, and $\tau > 0$ small (so that the relevant sets are in U^λ),

$$x \in S(\tilde{x}, \tilde{y}, \tau) \quad \Rightarrow \quad \tilde{x} \in S(x, y, K\tau). \tag{9.2}$$

Proof. First, fix $(\tilde{x}, \tilde{y}) \in \partial^c u$. We consider the transformation $(x, u) \longmapsto (q, \tilde{u})$ with respect to $\tilde{y}$ (see Definition 4.1). Let $\tilde{q} = -D_x c(\tilde{x}, \tilde{y}) \in T^*_{\tilde{y}} V$ denote the point corresponding to $\tilde{x}$ in these new coordinates. To show (9.2) we will find $K > 0$ such that

$$q \in Q_\tau \quad \Longrightarrow \quad \tilde{u}(\tilde{q}) \leq \tilde{u}(q) + \tilde{c}(q, y) - \tilde{c}(\tilde{q}, y) + K\tau, \qquad \forall\, y \in \partial^{\tilde{c}} \tilde{u}(q). \tag{9.3}$$

Fix $\tau > 0$ small, $q \in Q_\tau$ and $y \in \partial^{\tilde{c}} \tilde{u}(q)$. Assume by translation that the John ellipsoid of $Q_{2\tau}$ is centered at the origin. By Lemma 9.2 we have $Q_\tau \subset \rho_0 Q_{2\tau}$ for some $\rho_0 < 1$ (depending only on the dimension n and $\gamma_c^+ \gamma_c^- / \lambda$). Hence, by Lemma 6.3 applied with $\mathscr{K} = Q_{2\tau}$ we obtain

$$\|D_q \tilde{c}(\bar{q}, y)\|^*_{Q_{2\tau}} \leq C\Big(n, \frac{\gamma_c^+ \gamma_c^-}{\lambda}\Big)\, \tau \qquad \forall\, \bar{q} \in Q_\tau, \tag{9.4}$$

where $\| \cdot \|^*_{Q_{2\tau}}$ is the dual norm associated to $Q_{2\tau}$ (see (6.5)).

Recall from Theorem 4.3 that $\tilde{q}$ minimizes $\tilde{u}$, so $\tilde{u}(\tilde{q}) \leq \tilde{u}(q)$. To obtain (9.3) from this, we need only estimate the difference between the two costs. Since both $q, \tilde{q} \in Q_\tau \subset Q_{2\tau}$, by the definition of $\| \cdot \|^*_{Q_{2\tau}}$ and (9.4), (recalling $sq + (1-s)\tilde{q} \in Q_\tau$ due to convexity of Q_τ) we find

$$\begin{aligned}
&\big|\tilde{c}(q, y) - \tilde{c}(\tilde{q}, y)\big| \\
&= \left| \int_0^1 D_q \tilde{c}\big(sq + (1-s)\tilde{q}, y\big)\, ds \cdot (q - \tilde{q}) \right| \\
&\leq \left| \int_0^1 D_q \tilde{c}\big(sq + (1-s)\tilde{q}, y\big) \cdot q\, ds \right| + \left| \int_0^1 D_q \tilde{c}\big(sq + (1-s)\tilde{q}, y\big) \cdot \tilde{q}\, ds \right| \\
&\leq 2 \int_0^1 \|D_q \tilde{c}\big(sq + (1-s)\tilde{q}, y\big)\|^*_{Q_{2\tau}}\, ds \\
&\leq C\Big(n, \frac{\gamma_c^+ \gamma_c^-}{\lambda}\Big)\, \tau \\
&= K\tau,
\end{aligned}$$

thus establishing (9.3). □

Having established the engulfing property, the $C^{1,\alpha}$ estimates of the potential functions follows by applying a modified version of Forzani and Maldonado's method [24]. Here is a key consequence of the engulfing property.

Lemma 9.4 (Gain in c-monotonicity due to engulfing)**.** *Assume the engulfing property* (9.2) *holds. Let $c, u, x, \tilde{x}$ and K be as in Theorem 9.3, and let $y \in \partial^c u(x)$ and $\tilde{y} \in \partial^c u(\tilde{x})$. Then,*

$$\frac{1+K}{K}[u(x) - u(\tilde{x}) - c(\tilde{x}, \tilde{y}) + c(x, \tilde{y})] \leq c(\tilde{x}, y) - c(x, y) - c(\tilde{x}, \tilde{y}) + c(x, \tilde{y}).$$

Proof. Given $x, \tilde{x}$, notice that $u(\tilde{x}) - u(x) + c(\tilde{x}, y) - c(x, y) \geq 0$. Fix $\varepsilon > 0$ small, to ensure

$$\tau := u(\tilde{x}) - u(x) + c(\tilde{x}, y) - c(x, y) + \varepsilon > 0.$$

Then $\tilde{x} \in S(x, y, \tau)$, which by the engulfing property implies $x \in S(\tilde{x}, \tilde{y}, K\tau)$, that is

$$u(x) \leq u(\tilde{x}) + c(\tilde{x}, \tilde{y}) - c(x, \tilde{y}) + K[u(\tilde{x}) - u(x) + c(\tilde{x}, y) - c(x, y) + \varepsilon].$$

Letting $\varepsilon \to 0$ and rearranging terms we get

$$(K+1)u(x) \leq (K+1)u(\tilde{x}) + c(\tilde{x}, \tilde{y}) - c(x, \tilde{y}) + K[c(\tilde{x}, y) - c(x, y)],$$

or equivalently

$$u(x) - u(\tilde{x}) \leq \frac{1}{1+K}[c(\tilde{x}, \tilde{y}) - c(x, \tilde{y})] + \frac{K}{1+K}[c(\tilde{x}, y) - c(x, y)].$$

This gives

$$u(x) - u(\tilde{x}) - c(\tilde{x}, \tilde{y}) + c(x, \tilde{y}) \leq \frac{K}{1+K}[c(\tilde{x}, y) - c(x, y) - c(\tilde{x}, \tilde{y}) + c(x, \tilde{y})],$$

as desired □

In the above lemma, it is crucial to have a factor $(1+K)/K > 1$. Indeed, the above result implies the desired Hölder continuity of u, with a Hölder exponent independent of the particular choice of c (see Corollary 9.6):

Theorem 9.5 (Hölder continuity of optimal maps to strongly c-convex targets)**.** *Let c satisfy* **(B0)**–**(B3)** *and* $\mathbf{(B2)_s}$. *If u is a c-convex solution of* (3.8) *on $U^\lambda \subset U$ open, then $u \in C^{1,1/K}_{loc}(U^\lambda)$, with K as in Theorem 9.3 which depends only on the dimension n and $\gamma_c^+\gamma_c^-/\lambda$.*

Proof. As we already pointed out in the previous section, although not strictly needed, we will use the additional information that $u \in C^1(U^\lambda)$ (Theorem 8.3) to avoid some technical issues in the following proof. However, it is interesting to point out the argument below works with minor modifications even if u is not C^1, replacing the gradient by subdifferentials (recall that u is semiconvex), and using that semiconvex function are Lipschitz and so differentiable a.e. We leave the details to the interested reader.

The proof uses the idea of Forzani and Maldonado [24]. The c-convexity of u is strict on U^λ, according to Theorem 8.1. Given a point $x_s \in U^\lambda$, we denote by y_s the unique element in $\partial^c u(x_s)$; the uniqueness of y_s follows from the C^1 regularity of u, since y_s is uniquely identified by the relation $Du(x_s) = -D_x c(x_s, y_s)$.

Let $x_0 \in U^\lambda$. We will show that for $x_1 \in U^\lambda$ sufficiently close to x_0,

$$|u(x_0) - u(x_1) - Du(x_1) \cdot (x_0 - x_1)| \lesssim |x_0 - x_1|^{1+1/K},$$

from which the local $C^{1,1/K}$ regularity of u follows by standard arguments.

Fix a direction v with $|v|$ small, set $x_s = x_0 + sv$, and consider the function

$$\phi(s) := u(x_s) - u(x_0) + c(x_s, y_0) - c(x_0, y_0) \geq 0,$$

for $s \in [0,1]$. The idea is to use Lemma 9.4 to derive a differential inequality, which controls the growth of ϕ. First, observe that

$$\phi'(s) = Du(x_s) \cdot v + D_x c(x_s, y_0) \cdot v.$$

Since $Du(x_s) = -D_x c(x_s, y_s)$, we get

$$\begin{aligned}\phi'(s)s &= [D_x c(x_s, y_0) - D_x c(x_s, y_s)] \cdot (sv) \\ &\geq c(x_0, y_s) - c(x_s, y_s) - c(x_0, y_0) + c(x_s, y_0) - \|D^2_{xx} c\|_{L^\infty(U\times V)} s^2 |v|^2.\end{aligned}$$

So, by Lemma 9.4 we get

$$\frac{1+K}{K}\phi(s) \leq \phi'(s)s + \|D^2_{xx} c\|_{L^\infty(U\times V)} s^2 |v|^2,$$

that is

$$\frac{d}{dt}\left(\frac{\phi(s)}{s^{1+1/K}}\right) \geq -\frac{\|D^2_{xx} c\|_{L^\infty(U\times V)} |v|^2}{s^{1/K}}.$$

Hence $\phi(s)/s^{1+1/K} \leq \phi(1) + \|D^2_{xx} c\|_{L^\infty(U\times V)} |v|^2 \int_s^1 \tau^{-1/K}\, d\tau \leq \phi(1) + C_1$ (since $1 - \frac{1}{K} > 0$). So,

$$\begin{aligned}\phi(s) &= u(x_s) - u(x_0) - c(x_0, y_0) + c(x_s, y_0) \\ &\leq s^{1+1/K}[u(x_1) - u(x_0) - c(x_1) + c(x_0 + v, y_0) + C_1] \\ &\leq 2C_1 s^{1+1/K} \text{ (choosing } |v| \text{ small enough and using the continuity of } u \text{ and } c).\end{aligned}$$

By the arbitrariness of x_0, v and s we easily deduce that, for all $x_0, x_1 \in U^\lambda$ sufficiently close,

$$u(x_1) - u(x_0) - c(x_0, y_0) + c(x_1, y_0) \leq 2C_1 |x_0 - x_1|^{1+1/K}.$$

Since c is smooth, $Du(x_1) = -D_x c(x_1, y_1)$, and $u(x_0) - u(x_1) - c(x_1, y_0) + c(x_0, y_0) \geq 0$, the last inequality implies

$$\begin{aligned}&|u(x_0) - u(x_1) - Du(x_1) \cdot (x_0 - x_1)| \\ &\leq |u(x_0) - u(x_1) - c(x_1, y_0) + c(x_0, y_0)| + \|D^2_{xx} c\|_{L^\infty(U\times V)} |x_0 - x_1|^2 \\ &\leq C_2(|x_0 - x_1|^{1+1/K} + |x_0 - x_1|^2) \\ &\leq 2C_2 |x_0 - x_1|^{1+1/K}\end{aligned}$$

for all $x_0, x_1 \in U^\lambda$ sufficiently close. This proves the desired estimate, and concludes the proof of the $C^{1,1/K}_{loc}$ regularity of u inside U^λ. □

In fact, the Hölder exponent in the previous theorem does not depend on the particular cost function:

Corollary 9.6 (Universal Hölder exponent)**.** *With the same notation and assumptions as in Theorem 9.5, $u \in C^{1,\alpha}_{loc}(U^\lambda)$, where the Hölder exponent $\alpha > 0$ depends only on n and $\lambda > 0$.*

Proof. Recalling $\gamma_c^\pm = \gamma_c^\pm(U^\lambda \times V) := \|(\det D^2_{xy}c)^{\pm 1}\|_{L^\infty(U^\lambda \times V)}$ from (4.4), we see $\gamma_c^+ \to 1/\gamma_c^-$ as the set $U^\lambda \times V$ shrinks to a point. Since u is C^1 by Theorem 9.5, $\partial^c u$ gives a single-valued continuous map. Compactness of $\overline{U}^\lambda$ combined with **(B0)-(B1)** allows the set $\partial^c u \cap (\overline{U}^\lambda \times \overline{V})$ to be covered with finitely many neighborhoods $U_k \times V_k$, such that $\gamma_c^+(U_k \times V_k)\gamma_c^-(U_k \times V_k) \leq 2$ for all k. Hence, thanks to Remark 9.1, we can apply Theorem 9.5 on each such neighborhood $(U^\lambda \cap U_k)$, which in turn yields a Hölder exponent $0 < \alpha < 1$ depending only on n and $\lambda > 0$. □

References

[1] N. Ahmad, H.K. Kim and R.J. McCann. Optimal transportation, topology and uniqueness. *Bull. Math. Sci.* **1** (2011), no. 1, 13–32.

[2] A. Alexandroff. Smoothness of the convex surface of bounded Gaussian curvature. *C.R. (Doklady) Acad. Sci. URSS (N.S.)* **36** (1942) 195–199.

[3] J.W. Barrett and L. Prigozhin. Partial L^1 Monge-Kantorovich problem: Variational formulation and numerical approximation. *Interfaces and Free Boundaries* **11** (2009) 201–238.

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

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

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

[7] L.A. Caffarelli. Some regularity properties of solutions of Monge-Ampère equation. *Comm. Pure Appl. Math.* **64** (1991) 965–969.

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

[9] L.A. Caffarelli. Boundary regularity of maps with convex potentials — II. *Ann. of Math. (2)* **144** (1996) 453–496.

[10] L.A. Caffarelli and R.J. McCann. Free boundaries in optimal transport and Monge-Ampère obstacle problems. *Ann. of Math. (2)* **171** (2010) 673–730.

[11] L.A. Caffarelli, C.E. Gutiérrez and Q. Huang. On the regularity of reflector antennas. *Ann. of Math. (2)* **167** (2008), no. 1, 299–323.

[12] P. Delanoë. Classical solvability in dimension two of the second boundary-value problem associated with the Monge-Ampère operator. *Ann. Inst. H. Poincarè Anal. Non Linèaire* **8** (1991) 443–457.

[13] P. Delanoë and Y. Ge. Regularity of optimal transport on compact, locally nearly spherical, manifolds. *J. Reine Angew. Math.* **646** (2010) 65-115.

[14] P. Delanoë and Y. Ge. Locally nearly spherical surfaces are almost-positively c-curved. *Methods Appl. Anal.* **18** (2011), no. 3, 269–302.

[15] A. Figalli. The optimal partial transport problem. *Arch. Ration. Mech. Anal.* **195** (2010), no. 2, 533–560.

[16] A. Figalli. A note on the regularity of the free boundaries in the optimal partial transport problem. *Rend. Circ. Mat. Palermo.* **58** (2009), no. 2, 283–286.

[17] A. Figalli, Y.-H. Kim and R.J. McCann. Regularity of optimal transport maps on multiple products of spheres. To appear in *J. Eur. Math. Soc. (JEMS)*.

[18] A. Figalli, Y.-H. Kim and R.J. McCann. When is multidimensional screening a convex program? *J. Econom. Theory* **146** (2011) 454–478.

[19] A. Figalli, Y.-H. Kim and R.J. McCann. On supporting hyperplanes to convex bodies. *Preprint,* 2011.

[20] A. Figalli and G. Loeper. C^1 regularity of solutions of the Monge-Ampère equation for optimal transport in dimension two. *Calc. Var. Partial Differential Equations* **35** (2009) 537–550.

[21] A. Figalli and L. Rifford. Continuity of optimal transport maps on small deformations of $\mathbb{S}^2$. *Comm. Pure Appl. Math.* **62** (2009), no. 12, 1670–1706.

[22] A. Figalli, L. Rifford and C. Villani. Nearly round spheres look convex. *Amer. J. Math.* **134** (2012), no. 1, 109–139.

[23] A. Figalli, L. Rifford and C. Villani. Necessary and sufficient conditions for continuity of optimal transport maps on Riemannian manifolds. *Tohoku Math. J. (2)* **63** (2011), no. 4, 855–876.

[24] L. Forzani and D. Maldonado. Properties of the solutions to the Monge-Ampère equation. *Nonlinear Anal.* **57** (2004) 815–829.

[25] W. Gangbo. *Habilitation thesis.* Université de Metz, 1995.

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

[27] W. Gangbo and R.J. McCann. Shape recognition via Wasserstein distance. *Quart. Appl. Math.* **58** (2000) 705–737.

[28] N. Guillen and J. Kitagawa. On the local geometry of maps with c-convex potentials. *Preprint at arXiv:math/1212.4865v1.*

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

[30] C.E. Gutiérrez and Q. Huang. Geometric properties of the sections of solutions to the Monge-Ampère equation. *Trans. Amer. Math. Soc.* **352** (2000), no. 9, 4381–4396.

[31] F.R. Harvey and H.B. Lawson. Split Special-Lagrangian geometry. Preprint 2010.

[32] F. John. Extremum problems with inequalities as subsidiary conditions. In *Studies and Essays Presented to R. Courant on his 60th Birthday, January 8, 1948*, pages 187–204. Interscience, New York, 1948.

[33] L. Kantorovich. On the translocation of masses. *C.R. (Doklady) Acad. Sci. URSS (N.S.)* **37** (1942) 199–201.

[34] L. Kantorovich. On a problem of Monge (In Russian). *Uspekhi Math. Nauk.* **3** (1948) 225–226.

[35] Y.-H. Kim and R.J. McCann. Appendices to original version of Continuity, curvature, and the general covariance of optimal transportation. *Preprint at arXiv:math/0712.3077v1.*

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

[37] Y.-H. Kim and R.J. McCann. Towards the smoothness of optimal maps on Riemannian submersions and Riemannian products (of round spheres in particular). *J. Reine Angew. Math.* **664** (2012) 1–27.

[38] Y.-H. Kim, R.J. McCann and M. Warren Pseudo-Riemannian geometry calibrates optimal transportation. *Math. Res. Lett.* **17** (2010) 1183–1197.

[39] P.W.Y. Lee and R.J. McCann. The Ma-Trudinger-Wang curvature for natural mechanical actions. *Calc. Var. Partial Differential Equations.* **41** (2011) 285–299.

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

[41] J. Liu. Hölder regularity of optimal mappings in optimal transportation. *Calc Var. Partial Differential Equations* **34** (2009) 435–451.

[42] J. Liu, N.S. Trudinger and X.-J. Wang. Interior $C^{2,\alpha}$ regularity for potential functions in optimal transportation. *Comm. Partial Differential Equations* **35** (2010), no.1, 165–184.

[43] G. Loeper. On the regularity of maps solutions of optimal transportation problems. *Acta. Math.* **202** (2009), no. 2, 241–283.

[44] G. Loeper. Regularity of optimal maps on the sphere: The quadratic cost and the reflector antenna. *Arch. Ration. Mech. Anal.* **199** (2011), no.1, 269–289.

[45] G. Loeper and C. Villani. Regularity of optimal transport in curved geometry: the non-focal case. *Duke Math. J.* **151** (2010), no. 3, 431–485.

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

[47] G. Monge. Mémoire sur la théorie des déblais et de remblais. *Histoire de l'Académie Royale des Sciences de Paris, avec les Mémoires de Mathématique et de Physique pour la même année*, pages 666–704, 1781.

[48] R.T. Rockafellar. *Convex analysis.* Reprint of the 1970 original. *Princeton Landmarks in Mathematics.* Princeton Paperbacks. Princeton University Press, Princeton, NJ, 1997.

[49] T. Sei. A Jacobian inequality for gradient maps on the sphere and its application to directional statistics. *Preprint at arXiv:0906.0874.* To appear in *Communications in Statistics - Theory and Methods.*

[50] N.S. Trudinger. Recent developments in elliptic partial differential equations of Monge-Ampère type, pp 291–301 in *International Congress of Mathematicians* **3**, Zürich: Eur. Math. Soc., 2006.

[51] N.S. Trudinger and X.-J. Wang. On the second boundary value problem for Monge-Ampère type equations and optimal transportation. *Ann. Sc. Norm. Super. Pisa Cl. Sci. (5)* **8** (2009), no. 1, 143–174.

[52] N.S. Trudinger and X.-J. Wang. On strict convexity and C^1-regularity of potential functions in optimal transportation. *Arch. Ration. Mech. Anal.* **192** (2009), no. 3, 403–418.

[53] J. Urbas. On the second boundary value problem for equations of Monge-Ampère type. *J. Reine Angew. Math.* **487** (1997) 115–124.

[54] J. Vétois. Continuity and injectivity of optimal maps. *Preprint,* 2011.

[55] C. Villani. *Optimal Transport. Old and New*, volume 338 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences].* Springer, New York, 2009.

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

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