On the regularity of maps solutions of optimal transportation problems

Grégoire Loeper*

May 31, 2006

Abstract

We give a necessary and sufficient condition on the cost function so that the map solution of Monge's optimal transportation problem is continuous when the measures are smooth. This condition was first introduced by Ma, Trudinger and Wang [17, 22] for a priori estimates of the corresponding Monge-Ampère equation. It is expressed by a so-called *cost-sectional curvature* being non-negative. We show that when the cost function is the squared distance of a Riemannian manifold, the cost-sectional curvature yields the sectional curvature. As a consequence, if the manifold does not have non-negative sectional curvature everywhere, the optimal transport map *can not be continuous* for arbitrary smooth positive measures. The non-negativity of the cost-sectional curvature is shown to be equivalent to the connectedness of the contact set between any cost-convex function (the proper generalization of a convex function) and any of its supporting functions. When the cost-sectional curvature is uniformly positive, we obtain that optimal maps are continuous or Hölder continuous under quite weak assumptions on the measures, compared to what is needed in the Euclidean case. This case includes the reflector antenna problem and the squared Riemannian distance on the sphere.

---

[1] Institut Camille Jordan, Université Claude Bernard Lyon 1

Contents

1 Introduction 3

2 Preliminaries 6
- 2.1 Notation 6
- 2.2 Kantorovitch duality and c-convex potentials 6
- 2.3 The Riemannian case 9
- 2.4 Notion of c-convexity for sets 9
- 2.5 The Monge-Ampère equation 10
- 2.6 Generalized solutions 10
- 2.7 Cost-sectional curvature and conditions **Aw**, **As** 11
- 2.8 Previous regularity results for optimal maps 11

3 Results 13
- 3.1 Condition **Aw**, connectedness of the contact set and regularity issues 13
- 3.2 Improved partial regularity under **As** 15
- 3.3 Conditions **Aw**, **As** for the quadratic cost of a Riemannian manifold 16
- 3.4 Examples of costs that satisfy **As** or **Aw** 18

4 Proof of Theorems 3.1, 3.2 18
- 4.1 Proof of Theorem 3.1 19
- 4.2 Proof of Theorem 3.2 20

5 Proof of Theorem 3.5 21
- 5.1 Sketch of the proof 21
- 5.2 Geometric interpretation of condition **As** 22
- 5.3 Construction of supporting functions 25
- 5.4 Continuity estimates 27

6 Proof of Theorem 3.11 31
- 6.1 Monge-Kantorovitch problem on a Riemannian manifold 31
- 6.2 Reduction of the problem to an Euclidean problem 32
- 6.3 Verification of assumption **As**: Proof of Theorem 3.10 33
- 6.4 Conclusion of the proof 36

7 Proof of Theorem 3.7 37

8 Proof of Theorem 3.8 37
- 8.1 Counterexample to regularity for a manifold with negative curvature 38

9 Appendix 39

1 Introduction

Given A, B two topological spaces, a cost function $c : A \times B \to \mathbb{R}$, and μ_0, μ_1 two probability measures respectively on A and B, Monge's problem of optimal transportation consists in finding among all measurable maps $T : A \to B$ that push forward μ_0 onto μ_1 (hereafter $T_\# \mu_0 = \mu_1$) in the sense that

$$\forall E \subset B \text{ Borel }, \mu_1(E) = \mu_0(T^{-1}(E)), \tag{1}$$

a (the ?) map that realizes

$$\operatorname{Argmin}\Big\{ \int_A c(x, T(x))\, d\mu_0(x), T_\# \mu_0 = \mu_1 \Big\}. \tag{2}$$

Optimal transportation has undergone a rapid and important development since the pioneering work of Brenier, who discovered that when $A = B = \mathbb{R}^n$ and the cost is the distance squared, optimal maps for the problem (2) are gradients of convex functions [1]. Following this result and its subsequent extensions, the theory of optimal transportation has flourished, with generalizations to other cost functions [10, 15], more general spaces such as Riemannian manifolds [19], applications in many other areas of mathematics such as geometric analysis, functional inequalities, fluid mechanics, dynamical systems, and other more concrete applications such as irrigation, cosmology.

When A, B are domains of the Euclidean space $\mathbb{R}^n$, or of a Riemannian manifold, a common feature to all optimal transportation problems is that optimal maps derive from a (cost-convex) potential, which, assuming some smoothness, is in turn solution to a fully non-linear elliptic PDE: the Monge-Ampère equation. In all cases, the Monge-Ampère equation arising from an optimal transportation problem reads in local coordinates

$$\det(D^2\phi - \mathcal{A}(x, \nabla\phi)) = f(x, \nabla\phi), \tag{3}$$

where $(x, p) \to \mathcal{A}(x, p)$ is a symmetric matrix valued function, that depends on the cost function $c(x, y)$ through the formula

$$\mathcal{A}(x, p) = -D^2_{xx} c(x, y) \text{ for } y \text{ such that } -\nabla_x c(x, y) = p. \tag{4}$$

That there is indeed a unique y such that $-\nabla_x c(x, y) = p$ will be guaranteed by condition **A1** given hereafter. The optimal map will then be

$$x \to y : -\nabla_x c(x, y) = \nabla\phi(x).$$

In the case $\mathcal{A} = 0$, equation (3) was well known and studied before optimal transportation since it appears in Minkowsky's problem: find a convex hypersurface with prescribed Gauss curvature. In the case of optimal transportation, the boundary condition consists in prescribing that the image of the optimal map lies in a certain domain. It is known as the second boundary value problem.

Until recently, except in the particular case of the so-called reflector antenna, treated by Wang [28], the regularity of optimal maps was only known in the case where the cost function is the (Euclidean) squared distance $c(x,y)=|x-y|^2$, which is the cost considered by Brenier in [1], for which the matrix $\mathcal{A}$ in (3) is the identity (which is trivially equivalent to the case $\mathcal{A}=0$). Those results have involved several authors, among which Caffarelli, Urbas , and Delanoë. An important step was made recently by Ma, Trudinger and Wang [17], and Trudinger and Wang [22], who introduced a condition (named A3 and A3w in their papers) on the cost function under which they could show existence of smooth solutions to (3). Let us give right away this condition that will play a central role in the present paper. Let $A=\Omega, B=\Omega'$ be bounded domains of $\mathbb{R}^n$ on which the initial and final measures will be supported. Assume that c belongs to $C^4(\Omega\times\Omega')$. For $(x,y)\in(\Omega\times\Omega'), (\xi,\nu)\in\mathbb{R}^n\times\mathbb{R}^n$, we define

$$(5)\qquad \mathfrak{S}_{\mathfrak{c}}(x,y)(\xi,\nu) := D^2_{p_k p_l}\mathcal{A}_{ij}\,\xi_i\xi_j\,\nu_k\nu_l\,(x,p),\quad p=-\nabla_x c(x,y).$$

Whenever ξ,ν are orthogonal unit vectors, we will say that $\mathfrak{S}_{\mathfrak{c}}(x,y)(\xi,\nu)$ defines the *cost-sectional curvature from x to y in the directions* (ξ,ν). Note that this map is in general not symmetric, and that it depends on two points x and y. The reason why we use the word sectional curvature will be clear in a few lines. We will say that the cost function c has non-negative cost-sectional curvature on $(\Omega\times\Omega')$, if

$$(6)\qquad \mathfrak{S}_{\mathfrak{c}}(x,y)(\xi,\nu)\geq 0\quad \forall(x,y)\in(\Omega\times\Omega'),\forall(\xi,\nu)\in\mathbb{R}^n\times\mathbb{R}^n,\ \xi\perp\nu.$$

A cost function satisfies condition **Aw** on $(\Omega\times\Omega')$ if and only if it has non-negative cost-sectional curvature on $(\Omega\times\Omega')$, i.e. if it satisfies (6).

Under condition **Aw** and natural requirements on the domains Ω,Ω', Trudinger and Wang [22] showed that the solution to (3) is globally smooth for smooth positive measures μ_0,μ_1. They showed that **Aw** is satisfied by a large class of cost functions, that we will give as examples later on. Note that the quadratic cost satisfies assumption **Aw**. This result is achieved by the so-called continuity method, for which a key ingredient is to obtain a priori estimates on the second derivatives of the solution. At this stage, condition **Aw** was used in a crucial way. However, even if it was known that not all cost functions can lead to smooth optimal maps, it was unclear whether the condition **Aw** was necessary, or just a technical condition for the a-priori estimates to go through.

In this paper we show that the condition **Aw** is indeed the *necessary and sufficient condition for regularity:* one can not expect regularity without this condition, and more precisely, if $\mathfrak{S}_{\mathfrak{c}}(x,y)(\xi,\nu)<0$ for $(x,y)\in(\Omega\times\Omega'),\xi\perp\nu\in\mathbb{R}^n$, one can immediately build a pair of C^∞ strictly positive measures, supported on sets that satisfy the usual smoothness and convexity assumptions, so that the optimal potential is not even C^1, and the optimal map is therefore discontinuous. This result is obtained by analyzing the geometric nature of condition (6). Let us first recall that the solution ϕ of the Monge-Ampère equation is a priori known to be cost-convex (in short c-convex), meaning that at each point $x\in\Omega$, there exist $y\in\Omega'$ and a value $\phi^c(y)$ such that

$$\begin{aligned}&-\phi^c(y)-c(x,y)=\phi(x),\\ &-\phi^c(y)-c(x',y)\leq\phi(x'),\ \forall x'\in\Omega.\end{aligned}$$

The function $-\phi^c(y) - c(x,y)$ is called a supporting function, and the function $y \to \phi^c(y)$ is called the cost-transform (in short the c-transform) of ϕ, also defined by

$$\phi^c(y) = \sup_{x\in\Omega}\{-c(x,y) - \phi(x)\}.$$

(These notions will be recalled in greater details hereafter). We prove that the condition **Aw** can be reformulated as a property of cost-convex functions, which we call *connectedness of the contact set*:

$$\text{For all } x \in \Omega, \text{ the contact set } G_\phi(x) := \{y : \phi^c(y) = -\phi(x) - c(x,y)\} \text{ is connected.} \tag{7}$$

Assuming a natural condition on Ω' (namely its c-convexity, see Definition 2.9) this condition involves only the cost function since it must hold for any ϕ^c defined through a c-transform.

A case of special interest for applications is the generalization of Brenier's cost $\frac{1}{2}|x-y|^2$ to Riemannian manifolds, namely $c(x,y) = \frac{1}{2}d^2(x,y)$. Existence and uniqueness of optimal maps in that case was established by McCann [19], and further examined by several authors, with many interesting applications in geometric and functional analysis (for example [11, 20]). The optimal map takes the form $x \to \exp_x(\nabla\phi(x))$ for ϕ a c-convex potential and is called a gradient map. Then, a natural question is the interpretation of condition **Aw** and of the cost-sectional curvature in this context. We show that we have the identity:

$$\text{Cost-sectional curvature from } x \text{ to } x = 8 \cdot \text{ Riemannian sectional curvature at } x.$$

(We mean there that the equality holds for every 2-plane.) As a direct consequence of the previous result, *the optimal (gradient) map will not be continuous for arbitrary smooth positive data if the manifold does not have non-negative sectional curvature everywhere.* Although the techniques are totally different, it is interesting to notice that in recent works, Lott & Villani [25], and Sturm [21] have recovered the Ricci curvature through a property of optimal transport maps (namely through the displacement convexity of some functionals). Here, we somehow recover the sectional curvature through the continuity of optimal maps.

We next investigate the continuity of optimal maps under the stronger condition of uniformly positive sectional curvature, or condition **As**:

$$\exists C_0 > 0 : \ \mathfrak{S}_c(x,y,\xi,\nu) \geq C_0|\xi|^2|\nu|^2, \quad \forall (x,y) \in (\Omega\times\Omega'), (\xi,\nu) \in \mathbb{R}^n\times\mathbb{R}^n, \xi \perp \nu. \tag{8}$$

We obtain that the (weak) solution of (3) is C^1 or $C^{1,\alpha}$ under quite mild assumptions on the measures. Namely, we need only μ_1 to be bounded away from 0 and $\mu_0(B_r(x)) = o(r^{n-p}), p \leq 1$, ($B_r(x)$ being the ball of radius r, center x). Those conditions allow μ_0, μ_1 to be singular with respect to the Lebesgue measure and μ_0 to vanish.

This results extends the $C^{1,\alpha}$ estimate of Caffarelli [4] to a large class of cost functions and related Monge-Ampère equations. It also shows that the partial regularity results are better under **As** than under **Aw**, since Caffarelli's $C^{1,\alpha}$ regularity result required μ_0, μ_1 to have densities bounded away from 0 and infinity, and it is known to be close to optimal [26].

Finally, we prove that the quadratic cost on the sphere has uniformly positive cost-sectional curvature, i.e. satisfies **As**. We obtain therefore regularity of optimal (gradient) maps under adequate conditions.

The rest of the paper is organized as follows: in section 2 we gather all definitions and results that we will need throughout the paper. In section we state our results. Then each following section is devoted to the proof of a theorem. The reader knowledgeable to subject might skip directly to section 3.

2 Preliminaries

2.1 Notation

Hereafter dVol denotes the Lebesgue measure of $\mathbb{R}^n$ (or of the unit sphere $\mathbb{S}^{n-1}$ for Theorem 3.11), and $B_r(x)$ denotes a ball of radius r centered at x. For $\delta > 0$, we set classically $\Omega_\delta = \{x \in \Omega, d(x, \partial\Omega) > \delta)\}$. When we say that a function (resp. a measure) is smooth without stating the degree of smoothness, we assume that it is C^∞-smooth (resp. has a C^∞-smooth density with respect to the Lebesgue measure).

2.2 Kantorovitch duality and c-convex potentials

In this section, we recall how to obtain the optimal map from a c-convex potential in the general case. This allows us to introduce definitions that we will be using throughout the paper. References concerning the existence of optimal map by Monge-Kantorovitch duality is [1] for the cost $|x-y|^2$, [15] and [10] for general costs, [19] for the Riemannian case, otherwise the book [24] offers a rather complete reference on the topic.

Monge's problem (2) is first relaxed to become a problem of linear programming; one seeks now

$$\mathcal{I} = \inf\Big\{ \int_{\mathbb{R}^n\times\mathbb{R}^n} c(x,y)d\pi(x,y); \quad \pi \in \Pi(\mu_0,\mu_1)\Big\} \tag{9}$$

where $\Pi(\mu_0,\mu_1)$ is the set of positive measures on $\mathbb{R}^n \times \mathbb{R}^n$ whose marginals are respectively μ_0 and μ_1. Note that the (Kantorovitch) infimum (9) is smaller than the (Monge) infimum of the cost (2), since whenever a map T pushes forward μ_0 onto μ_1, the measure $\pi_T(x) := \mu_0(x) \otimes \delta_{T(x)}(y)$ belongs to $\Pi(\mu_1,\mu_1)$.

Then, the dual Monge-Kantorovitch problem is to find an optimal pair of potentials (ϕ,ψ) that realizes

$$\mathcal{J} = \sup\Big\{ -\int \phi(x)d\mu_0(x) - \int \psi(y)d\mu_1(y); \phi(x)+\psi(y) \geq -c(x,y)\Big\}. \tag{10}$$

The constraint on ϕ,ψ leads to the definition of c(c*)-transforms:

Definition 2.1 *Given a lower semi-continuous function* $\phi : \Omega \subset \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$*, we define its c-transform at* $y \in \Omega'$ *by*

$$\phi^c(y) = \sup_{x \in \Omega} \{-c(x, y) - \phi(x)\}. \tag{11}$$

Respectively, for $\psi : \Omega' \subset \mathbb{R}^n \to \mathbb{R}$ *also lower semi-continuous, define its c*-transform at* $x \in \Omega$ *by*

$$\psi^{c*}(x) = \sup_{y \in \Omega'} \{-c(x, y) - \psi(y)\}. \tag{12}$$

A function is said cost-convex, or, in short, c-convex, if it is the c-transform of another function* $\psi : \Omega' \to \mathbb{R}$*, i.e. for* $x \in \Omega$*,* $\phi(x) = \sup_{y \in \Omega'} \{-c(x, y) - \psi(y)\}$*, for some lower semi-continuous* $\psi : \Omega' \to \mathbb{R}$*. Moreover in this case* $\phi^{cc*} := (\phi^c)^{c*} = \phi$ *on* Ω *(see [24]).*

Our first assumption on c will be:

A0 The cost-function c belongs to $C^4(\bar{\Omega} \times \bar{\Omega}')$.

We will also always assume that Ω, Ω' are bounded. These assumptions are not the weakest possible for the existence/uniqueness theory.

Proposition 2.2 *Under assumption* **A0***, if* Ω *(resp.* Ω'*) is bounded,* ϕ^c *(resp.* ψ^{c*} *) will be locally semi-convex and Lipschitz.*

By Fenchel-Rockafellar's duality theorem, we have $\mathcal{I} = \mathcal{J}$. One can then easily show that the supremum (10) and the infimum (9) are achieved. Since the condition $\phi(x) + \psi(y) \geq -c(x, y)$ implies $\psi \geq \phi^c$, we can assume that for the optimal pair in $\mathcal{J}$ we have $\psi = \phi^c$ and $\phi = \phi^{cc*}$. Writing the equality of the integrals in (9, 10) for any optimal γ and any optimal pair (ϕ, ϕ^c) we obtain that γ is supported in $\Big\{\phi(x) + \phi^c(y) + c(x, y) = 0\Big\}$. This leads us to the following definition:

Definition 2.3 (Gradient mapping) *Let* ϕ *be a c-convex function. We define the set-valued mapping* G_ϕ *by*

$$G_\phi(x) = \Big\{y \in \Omega', \phi(x) + \phi^c(y) = -c(x, y)\Big\}.$$

For all $x \in \Omega$*,* $G_\phi(x)$ *is the contact set between* ϕ^c *and its supporting function* $-\phi(x) - c(x, \cdot)$*.*

Noticing that for all $y \in G_\phi(x)$, $\phi(\cdot) + c(\cdot, y)$ has a global minimum at x, we introduce / recall the following definitions:

Definition 2.4 (subdifferential) *For* ϕ *a semi-convex function, the subdifferential of* ϕ *at* x*, that we denote* $\partial\phi(x)$*, is the set*

$$\partial\phi(x) = \Big\{p \in \mathbb{R}^n, \phi(y) \geq \phi(x) + p \cdot (y - x) + o(|x - y|)\Big\}.$$

The subdifferential is always a convex set, and is always non empty for a semi-convex function.

Definition 2.5 (c-subdifferential) *If ϕ is c-convex, the c-sub-differential of ϕ at x, that we denote $\partial^c\phi(x)$, is the set*

$$\partial^c\phi(x) = \Big\{ - \nabla_x c(x,y), y \in G_\phi(x) \Big\}.$$

The inclusion $\emptyset \neq \partial^c\phi(x) \subset \partial\phi(x)$ always holds.

We introduce now two assumptions on the cost-function, which are the usual assumptions made in order to obtain an optimal map.

A1 For all $x \in \Omega$, the mapping $y \to -\nabla_x c(x,y)$ is injective on $\bar{\Omega}'$. Its image is a convex set of $\mathbb{R}^n$.

A2 The cost function c satisfies $\det D^2_{xy}c \neq 0$ for all $(x,y) \in \bar{\Omega} \times \bar{\Omega}'$.

This leads us to the definition of the *c-exponential map*:

Definition 2.6 *Under assumption* **A1**, *for $x \in \Omega$ we define the c-exponential map at x, which we denote by $\mathfrak{T}_x$, such that*

$$\forall (x,y) \in (\Omega' \times \Omega'), \mathfrak{T}_x(-\nabla_x c(x,y)) = y.$$

Remark. The definition c-exponential map is again motivated by the case cost=distance squared, where the c-exponential map is the exponential map. Moreover, notice that

$$[D^2_{x,y}c]^{-1} = -D_p\mathfrak{T}_x\big|_{x,p=-\nabla_x c(x,y)}.$$

Under assumptions **A1**, **A2**, G_ϕ is single valued outside of a set of Hausdorff dimension less or equal than $n-1$, hence, if μ_0 does not give mass to sets of Assort dimension less than $n-1$, G_ϕ will be the optimal map for Monge's problem while the optimal measure in (9) will be $\pi = \mu_0 \otimes \delta_{G_\phi(x)}$. So, after having relaxed the constraint that the optimal π should be supported on the graph of a map, one still obtains a minimizer that satisfy this constraint.

Let us mention at this point that Monge's historical cost was equal to the distance itself: $c(x,y) = |x-y|$. One sees immediately that for this cost function, there is not a unique y such that $-\nabla_x c(x,y) = \nabla\phi(x)$, and the dual approach fails.

We now state a general existence theorem, under assumptions that are clearly not minimal, but that will suffice for the scope of this paper, where we deal with regularity issues.

Theorem 2.7 *Let Ω, Ω' be two bounded domains of $\mathbb{R}^n$. Let $c \in C^4(\bar{\Omega} \times \bar{\Omega}')$ satisfy assumptions* **A0**-**A2**. *Let μ_0, μ_1 be two probability measures on Ω and Ω'. Assume that μ_0 does not give mass to sets of Assort dimension less or equal than $n-1$. Then there exists a $d\mu_0$ a.e. unique minimizer T of Monge's optimal transportation problem (2). Moreover, there exists ϕ c-convex on Ω such that $T = G_\phi$ (see 2.3). Finally, if ψ is c-convex and satisfies $G_{\psi\#}\mu_0 = \mu_1$, then $\nabla\psi = \nabla\phi$ $d\mu_0$ a.e.*

2.3 The Riemannian case

This approach has been extended in a natural way to compact smooth complete Riemannian manifolds by Robert McCann in [19]. We first see that all definitions can be translated into the setting of a compact Riemannian manifold, replacing the flat connexion by the Levi-Civita connexion of the manifold. The cost is allowed to be Lipschitz, semi-concave. When $c(\cdot,\cdot) = \frac{1}{2}d^2(\cdot,\cdot)$ with $d(\cdot,\cdot)$ the distance function, the c-exponential map is the exponential map, the map G_ϕ will be $x \to \exp_x(\nabla_g \phi)$, the gradient $\nabla_g \phi$ being relative to the Riemannian metric g.

2.4 Notion of c-convexity for sets

Following [17], we introduce here the notions that extend naturally the notions of convexity / strict convexity for a set.

Definition 2.8 (c-segment) *Let $p \to \mathfrak{T}_x(p)$ be the mapping defined by assumption* **A1**. *The point x being held fixed, a c-segment with respect to x is the image by $\mathfrak{T}_x$ of a segment of $\mathbb{R}^n$.*

If for $v_0, v_1 \in \mathbb{R}^n$ we have $\mathfrak{T}_x(v_i) = y_i, i = 0,1$, the c-segment with respect to x joining y_0 to y_1 will be $\{y_\theta, \theta \in [0,1]\}$ where $y_\theta = \mathfrak{T}_x(\theta v_1 + (1-\theta)v_0)$. It will be denoted $[y_0, y_1]_x$.

Definition 2.9 (c-convex sets) *Let $\Omega, \Omega' \subset \mathbb{R}^n$. We say that Ω' is c-convex with respect to Ω if for all $y_0, y_1 \in \Omega', x \in \Omega$, the c-segment $[y_0, y_1]_x$ is contained in Ω'.*

Remark. Note that this can be said in the following way: for all $x \in \Omega$, the set $-\nabla_x c(x, \Omega')$ is convex.

Definition 2.10 (uniform strict c-convexity of sets) *For Ω, Ω' two subsets of $\mathbb{R}^n$, we say that Ω' is uniformly strictly c-convex with respect to Ω if the sets $\{-\nabla_x c(x, \Omega')\}_{x\in\Omega}$ are uniformly strictly convex, uniformly with respect to x. We say that Ω is uniformly strictly c^*-convex with respect to Ω' if the dual assertion holds true.*

Remark. This definition is equivalent, although stated under a different form, with the definition given in [22].

Remarks on the sub-differential and c-sub-differential The question is to know if we have for all ϕ c-convex on Ω, for all $x \in \Omega$, $\partial\phi(x) = \partial^c\phi(x)$. Clearly, when ϕ is c-convex and differentiable at x,the equality holds. For p an extremal point of $\partial\phi(x)$, there will be a sequence x_n converging to x such that ϕ is differentiable at x_n and $\lim_n \nabla\phi(x_n) = p$. Hence, extremal points of $\partial\phi(x)$ belong to $\partial^c\phi(x)$. Then it is not hard to show the

Proposition 2.11 *Assume that Ω' is c-convex with respect to Ω. The following assertions are equivalent:*

1. *For all ϕ c-convex on Ω, $x \in \Omega$, $\partial^c\phi(x) = \partial\phi(x)$.*
2. *For all ϕ c-convex on Ω, $x \in \Omega$, $\partial^c\phi(x)$ is convex.*
3. *For all ϕ c-convex on Ω, $x \in \Omega$, $G_\phi(x)$ is c-convex with respect to x.*
4. *For all ϕ c-convex on Ω, $x \in \Omega$, $G_\phi(x)$ is connected.*

2.5 The Monge-Ampère equation

In all cases, for ϕ a C^2 smooth c-convex potential such that $G_{\phi\#}\mu_0 = \mu_1$, the conservation of mass is expressed in local coordinates by the following Monge-Ampère equation

$$\det(D^2_{xx}c(x, G_\phi(x)) + D^2\phi) = |\det D^2_{xy}c| \frac{\rho_0(x)}{\rho_1(G_\phi(x))}, \tag{13}$$

where $\rho_i = \mathrm{d}\mu_i/\mathrm{dVol}$ denotes the density of μ_i with respect to the Lebesgue measure. (See [17] for a derivation of this equation, or [11], [13].) Hence, the equation fits into the general form (3).

2.6 Generalized solutions

Definition 2.12 (Generalized solutions) *Let $\phi : \Omega \to \mathbb{R}$ be a c-convex function. Then*

- *ϕ is a weak Alexandrov solution to (13) if and only if*

  $$\text{for all } B \subset \Omega,\ \mu_0(B) = \mu_1(G_\phi(B)). \tag{14}$$

  This will be denoted by $\mu_0 = G_\phi^{\#}\mu_1$.

- *ϕ is a weak Brenier solution to (13) if and only if*

  $$\text{for all } B' \subset \Omega',\ \mu_1(B') = \mu_0(G_\phi^{-1}(B')). \tag{15}$$

  This is equivalent to $\mu_1 = G_{\phi\#}\mu_0$.

Alexandrov and Brenier solutions First notice that in the definition (15), μ_1 is deduced from μ_0, while it is the contrary in (14). As we have seen, the Kantorovitch procedure (10) yields an optimal transport map whenever μ_0 does not give mass to sets of Hausdorff dimension less than $n-1$. Moreover, the map G_ϕ will satisfy (15) by construction, and hence will be a weak Brenier solution to (13). Taking advantage of the c-convexity of ϕ one can show that whenever μ_1 is absolutely continuous with respect to the Lebesgue measure, $G_\phi^{\#}\mu_1$ is countably additive, and hence is a Radon measure (see [17, Lemma 3.4]); then a Brenier solution is an Alexandrov solutions. Note that one can consider $\mu_0 = G_\phi^{\#}\mathrm{dVol}$, this will be the Monge-Ampère measure of ϕ. Most importantly, for μ_0 supported in Ω, $G_{\phi\#}\mu_0 = \mathbf{1}_{\Omega'}\mathrm{dVol}$ does not imply that $G_\phi^{\#}\mathrm{dVol} = \mu_0$, except if Ω' is c-convex with respect to Ω (see [17]).

2.7 Cost-sectional curvature and conditions Aw, As

A central notion in the present paper will be the notion of *cost-sectional curvature* $\mathfrak{S}_\mathfrak{c}(x,y)$.

Definition 2.13 *Under assumptions* **A0**-**A2***, we can define*

$$\mathfrak{S}_\mathfrak{c}(x_0,y_0)(\xi,\nu) = D^4_{p_\nu p_\nu x_\xi x_\xi}\Big[(x,p) \to -c(x, \mathfrak{T}_{x_0}(p))\Big]\Big|_{x_0, y_0 = -\nabla_x c(x_0,y_0)}. \tag{16}$$

When ξ, ν are unit orthogonal vectors, $\mathfrak{S}_\mathfrak{c}(x_0,y_0)(\xi,\nu)$ defines the cost-sectional curvature from x_0 to y_0 in directions (ξ,ν).

Note that this definition is equivalent to (5).

Remark. The definition (16) would allow to define a 4-tensor which could be compared to the Riemannian curvature tensor, however its symmetries seem to be inconsistent with the symmetries of the Riemannian curvature tensor (the so-called Bianchi identities). Nevertheless, all the informations on the curvature tensor are contained in the sectional curvatures, i.e. two different curvature tensors can not lead to the same sectional curvature. We are now ready to introduce the conditions:

As The cost-sectional curvature is uniformly positive i.e. there exists $C_0 > 0$ such that for all $(x, y) \in (\Omega \times \Omega')$, for all $(\nu, \xi) \in \mathbb{R}^n \times \mathbb{R}^n)$ with $\xi \perp \nu$,

$$\mathfrak{S}_{\mathfrak{c}}(x, y)(\xi, \nu) \geq C_0 |\xi|^2 |\nu|^2.$$

Aw The cost-sectional curvature is non-negative: **As** is satisfied with $C_0 = 0$.

Remark . Let $c^*(y, x) := c(x, y)$. Writing $\mathfrak{S}_{\mathfrak{c}}$ in terms of (x, y) and not (x, p) as it is done in [17], one checks that if c satisfies **Aw** (resp. **As**) then c^* satisfies **Aw** (resp. **As** with a different constant).

2.8 Previous regularity results for optimal maps

The regularity of optimal maps follows from the regularity of the c-convex potential solution of the Monge-Ampère equation (13), the former being as smooth as the gradient of the latter. It falls thus into the theory of viscosity solutions of fully non-linear elliptic equations [8], however, the Monge-Ampère equation is degenerate elliptic. Two types of regularity results are usually sought for this type of equations:

Classical regularity: show that the equation has classical C^2 solutions, provided the measures are smooth enough, and assuming some boundary conditions. Due to the log-concavity of the Monge-Ampère operator, and using classical elliptic theory (see for instance [16]), C^∞ regularity of the solution of (13) follows from C^2 a priori estimates.

Partial regularity: show that a weak solution of (13) is C^1 or $C^{1,\alpha}$ under suitable conditions. We mention also that $W^{2,p}$ regularity results can be obtained.

The Euclidean Monge-Ampère equation and the quadratic cost This corresponds to the case where the cost function is the Euclidean distance squared $c(x, y) = |x - y|^2$, for which c-convexity means convexity in the usual sense, $G_\phi(x) = \nabla\phi(x)$, and equation (13) takes the following form

$$\det D^2\phi = \frac{\rho_0(x)}{\rho_1(\nabla\phi(x))}. \tag{17}$$

Here again, we have $\rho_i = d\mu_0/d\mathrm{Vol}$, $i = 0, 1$. Classical regularity has been established by Caffarelli [2, 6, 5, 7], Delanoë [12] and Urbas [23]. The optimal classical regularity result, found in [2, 7], is that for C^α smooth positive densities, and uniformly strictly convex domains, the solution of (17) is $C^{2,\alpha}(\bar{\Omega})$. Partial regularity results have been obtained by Caffarelli [3, 4, 6, 5], where it is shown that for μ_0, μ_1 having densities bounded away from 0 and infinity, the solution of (17) is $C^{1,\alpha}$. Thanks to counterexamples by Wang [26] those results are close to be optimal.

The reflector antenna The design of reflector antennas can be formulated as a problem of optimal transportation on the unit sphere with cost equal to $-\log|x-y|$. The potential (height function) $\phi : \mathbb{S}^{n-1} \to \mathbb{R}^+$ parametrizes the antenna A as follows: $A = \{x\phi(x), x \in \mathbb{S}^{n-1}\}$. Then the antenna is admissible if and only if ϕ is c-convex on $\mathbb{S}^{n-1}$ for $c(x,y) = -\log|x-y|$, and $G_\phi(x)$ yields the direction in which the ray coming in the direction x is reflected. This is the first non quadratic cost for which regularity of solutions has been established. Wang [27, 28] has shown classical C^2 (and hence C^∞) regularity of solutions of the associated Monge-Ampère equation when the densities are smooth. In a recent work, Caffarelli, Huang and Gutierrez [9] have shown C^1 regularity for the solution (i.e. continuity of the optimal map) under the condition that the measures μ_0 and μ_1 have densities bounded away from 0 and infinity.

General costs and the conditions As, Aw Recently an important step was achieved in two papers by Ma, Trudinger, and Wang . They gave in the first paper [17] a sufficient condition (**As**, called A3 in their paper) for C^2 (and subsequently C^∞) interior regularity. In the second paper [22], they could lower this condition down to **Aw** (condition A3w in their paper) to obtain a sufficient condition for global C^2 (and subsequently C^∞) regularity, assuming uniform strict c-convexity and smoothness of the domains. Note that the result under **Aw** recovers the results of Urbas and Delanöe for the quadratic cost.

Theorem 2.14 ([17, 22]) *Let Ω, Ω' be two bounded domains of $\mathbb{R}^n$. Let c satisfy* **A0**-**A2**. *Let μ_0, μ_1 be two probability measures on Ω, Ω' having densities ρ_0, ρ_1. Assume that $\rho_1 \in C^2(\bar{\Omega})$ is bounded away from 0, $\rho_2 \in C^2(\bar{\Omega}')$ is bounded away from 0.*

1. *If Ω' is c-convex with respect to Ω, and c satisfies* **As**, *then for ϕ c-convex on Ω such that $G_{\phi\#}\mu_0 = \mu_1$, $\phi \in C^3(\Omega)$.*
2. *If Ω, Ω' are strictly uniformly c,c*-convex with respect to each other and c satisfies* **Aw**, *for ϕ c-convex on Ω such that $G_{\phi\#}\mu_0 = \mu_1$, $\phi \in C^3(\bar{\Omega})$.*

3 Results

We present some answers to the following four questions:

1. Is there a sharp necessary and sufficient condition on the cost function which would guarantee that when both measures have C^∞ smooth densities, and their supports satisfy usual convexity assumptions, the solution of (13) (and hence the optimal map) is C^∞ smooth ?
2. Is there a similar partial regularity result available under a general condition ?
3. What are the cost-functions for which connectedness of the contact set holds (7) ?
4. When the cost is set to be the squared distance of a Riemannian manifold, what is the meaning of conditions **Aw**, **As** in terms of the Riemannian metric ?

3.1 Condition Aw, connectedness of the contact set and regularity issues

Answer to questions 1 and 3: Condition **Aw** *is necessary and sufficient for regularity of optimal maps. Moreover* **Aw** *is equivalent to the connectedness of the contact set.*

Theorem 3.1 *Let c satisfy* **A0**, **A1**, **A2**, **As**, *Ω' being c-convex with respect to Ω. Let ϕ be c-convex on Ω. Then at each point $x \in \Omega$, we have*

$$\partial\phi(x) = \partial^c\phi(x).$$

Moreover, the contact set

$$G_\phi(x) = \{y : \phi(x) + \phi^c(y) = -c(x, y)\}$$

is c-convex with respect to x, it is equal to $\mathfrak{T}_x(\partial\phi)$. If Ω is assumed c-convex with respect to Ω', then*

$$G_{\phi^c}(y) = \{x : \phi(x) + \phi^c(y) = -c(x, y)\}$$

is c-convex with respect to y*

Then we have the second theorem under **Aw**.

Theorem 3.2 *Let c be a cost function that satisfies* **A0**, **A1**, **A2**, *Ω, Ω' being smooth uniformly strictly c-convex (resp. c*-convex) with respect to each other. The following assertions are equivalent.*

1. *The cost function c satisfies* **Aw** *in $\Omega \times \Omega'$.*
2. *For μ_0, μ_1 smooth strictly positive probability measures in $\bar{\Omega}, \bar{\Omega}'$ there exists a c-convex potential $\phi \in C^1(\Omega)$ such that $G_{\phi\#}\mu_0 = \mu_1$.*
3. *For μ_0, μ_1 smooth satirically positive probability measures in $\bar{\Omega}, \bar{\Omega}'$ there exists a c-convex potential $\phi \in C^\infty(\bar{\Omega})$ such that $G_{\phi\#}\mu_0 = \mu_1$.*
4. *For all ϕ c-convex in Ω, for all $x \in \Omega$, $\partial^c\phi(x) = \partial\phi(x)$.*
5. *For all ϕ c-convex in Ω, for all $y \in \Omega'$, the set $\{x : \phi(x) + \phi^c(y) = -c(x, y)\}$ is c*-convex with respect to y.*
6. *C^1 c-convex potential are dense among c-convex potentials for the topology of local uniform convergence.*

Hence, if condition **Aw** *is violated at some points $(x_0, y_0) \in (\Omega \times \Omega')$, there exist smooth positive measures μ_0, μ_1 on Ω, Ω' such that there exists no C^1 c-convex potential satisfying $G_{\phi\#}\mu_0 = \mu_1$.*

REMARK. Setting $c^*(y,x) = c(x,y)$ we have seen that $\mathfrak{S}_{\mathfrak{c}} \geq 0$ implies $\mathfrak{S}_{\mathfrak{c}^*} \geq 0$. Hence all of those assertions are equivalent to their dual counterpart.

We can add the following equivalent condition for **Aw**:

Theorem 3.3 *Under the assumptions of Theorem 3.2, condition* **Aw** *holds if and only if, for any $x_0 \in \Omega$, $(y_0, y_1) \in \Omega'$, letting $\bar{\phi}$ be defined by*

$$\bar{\phi}(x) = \max\{-c(x,y_0) + c(x_0,y_0), -c(x,y_1) + c(x_0,y_1)\},$$

for any $y_\theta \in [y_0, y_1]_{x_0}$ (see Definition 2.8),

$$\bar{\phi}(x) \geq -c(x,y_\theta) + c(x_0,y_\theta)$$

holds in Ω.

In other words, $f_\theta(x) = -c(x,y_\theta) + c(x_0,y_\theta)$ which the supporting function that interpolates at x_0 (nonlinearly) between $f_0(x) = -c(x,y_0) + c(x_0,y_0)$ and $f_1(x) = -c(x,y_1) + c(x_0,y_1)$, has to remain below $\max\{f_0, f_1\}$.

REMARK 1. The function $\bar{\phi}$ furnishes the counter-example to regularity when **Aw** is not satisfied, since for a suitable choice of x_0, y_0, y_1 $\bar{\phi}$ can not be approximated by C^1 c-convex potentials.

REMARK 2. As shown by Propositions 5.1, 5.7, a quantitative version of Theorem 3.3 holds to express condition **As**.

Of course condition **Aw** looks like a good candidate, since in [22] it is shown shown to be a sufficient condition for classical regularity. We show here that it is necessary: if it is violated at some point, one can always build a counterexample where the solution to (13) is not C^1 even with C^∞ smooth positive measures and good boundary conditions (hence the optimal is not continuous). Moreover condition **Aw** is equivalent to a very natural geometric property of c-convex functions.

3.2 Improved partial regularity under As

Partial answer to question 2: There is partial (i.e. C^1 and $C^{1,\alpha}$) regularity under **As***, requiring much lower assumptions on the measures than what is needed in the quadratic case. There can not be C^1 regularity without* **Aw***. When only* **Aw** *is satisfied, we don't know yet, except for the case $c(x,y) = |x-y|^2$.*

Let us begin by giving the two integrability conditions that will be used in this result. The first reads

$$\text{For some } p \in]n, +\infty], C_{\mu_0} > 0,$$
$$\mu_0(B_\epsilon(x)) \leq C_{\mu_0} \epsilon^{n(1-\frac{1}{p})} \text{ for all } \epsilon \geq 0, x \in \Omega. \tag{18}$$

The second condition reads

$$\text{For some } f : \mathbb{R}^+ \to \mathbb{R}^+ \text{ with } \lim_{\epsilon \to 0} f(\epsilon) = 0,$$
$$\mu_0(B_\epsilon(x)) \leq f(\epsilon) \epsilon^{n(1-\frac{1}{n})} \text{ for all } \epsilon \geq 0, x \in \Omega. \tag{19}$$

In order to appreciate the forthcoming theorem, let us mention a few facts on these integrability conditions (the proof of this proposition is given at the end of the paper).

Proposition 3.4 *Let μ_0 be a probability measure on $\mathbb{R}^n$.*

1. *If μ_0 satisfies (18) for some $p > n$, μ_0 satisfies (19).*
2. *If $\mu_0 \in L^p(\Omega)$ for some $p > n$, μ_0 satisfies (18) with the same p.*
3. *If $\mu_0 \in L^n(\Omega)$, μ_0 satisfies (19).*
4. *If μ_0 satisfies (19), μ_0 does not give mass to set of Hausdorff dimension less or equal than $n-1$, hence (19) guarantees the existence of an optimal map.*
5. *There are probability measures on Ω that satisfy (18) (and hence (19)) and that are not absolutely continuous with respect to the Lebesgue measure.*

Then our result is

Theorem 3.5 *Let c be a cost function that satisfies assumptions* **A0**, **A1**, **A2**, **As** *on $(\Omega \times \Omega')$, Ω' being c-convex with respect to Ω. Let μ_0, μ_1 be probability measures on Ω and Ω', and ϕ be a c-convex potential on Ω such that $G_{\phi\#}\mu_0 = \mu_1$. Assume that $\mu_1 \geq m\,\mathrm{dVol}$ on Ω' for some $m > 0$*

1. *Assume that μ_0 satisfies (18) for some $p > n$. Let $\alpha = 1 - \frac{n}{p}$, $\beta = \frac{\alpha}{4n-2+\alpha}$. Then for any $\delta > 0$ we have*

   $$\|\phi\|_{C^{1,\beta}(\Omega_\delta)} \leq \mathcal{C},$$

   and $\mathcal{C}$ depends only on $\delta > 0$, C_{μ_0} in (18), on the constant $C_0 > 0$ in condition **As**, *on m and on $\|D^2_{xy}c(\cdot,\cdot)^{-1}\|_{L^\infty(\Omega\times\Omega')}, \|D^3c(\cdot,\cdot)\|_{L^\infty(\Omega\times\Omega')}$.*
2. *If μ_0 satisfies (19), then ϕ belongs to $C^1(\Omega_\delta)$ and the modulus of continuity of $\nabla\phi$ is controlled by f in (19).*

As an easy corollary of Theorem 3.5, we can extend the C^1 estimates to the boundary if the support of the measure μ_0 is compactly contained in Ω.

Theorem 3.6 *Assume in addition to the assumptions of Theorem 3.5 that μ_0 is supported in $\bar{\Omega}_2$, with Ω_2 compactly contained in Ω. Then, if μ_0 satisfies (19), $\phi \in C^1(\bar{\Omega}_2)$ and if μ_0 satisfies (18), $\phi \in C^{1,\beta}(\bar{\Omega}_2)$, with β as in Theorem 3.5.*

We mention also the consequence of this result for the reflector antenna problem (see paragraph 2.8 for a brief description of the problem).

Theorem 3.7 *let μ_0, μ_1 be probability measures on $\mathbb{S}^{n-1}$. Let $\phi : \mathbb{S}^{n-1} \to \mathbb{R}^+$ be the height function of the admissible antenna that reflects the incoming intensity μ_0 into the outcoming intensity μ_1. Then the conclusions of Theorems 3.5, 3.6 hold for ϕ.*

As in the case of the sphere hereafter, we need an a priori estimate to bound the distance $d(x, G_\phi(x))$ away from 0, otherwise the cost function $\log|x-y|$ becomes singular. The strategy for the proof of this Theorem will follow the same lines as the proof of Theorem 3.11.

The integrability conditions on μ_0, μ_1 are really mild: we only ask that μ_1 be bounded by below, and that $\mu_0(B_r) \le r^{n-p}$ for $p \ge 1$ ($p > 1$ yields $C^{1,\alpha}$ regularity) (see conditions (18) and (19) and the subsequent discussion). The continuity of the optimal map is also asserted in the case $\mu_0 \in L^n$ (that implies (19)), which is somehow surprising: indeed $D^2\phi \in L^n$ does not imply $\phi \in C^1$, but here $\det(D^2\phi - \mathcal{A}(x, \nabla\phi)) \in L^n$ implies $\phi \in C^1$. Then, as a consequence with obtain Theorem 3.7 that improves the result obtained independently by Caffarelli, Gutierrez and Huang [9] on reflector antennas. Moreover our techniques yield quantitative $C^{1,\alpha}$ estimates: the exponent α can explicitly computed. Finally, our continuity estimates extends up to the boundary (Theorem 3.6). This is achieved through a geometric formulation of condition **As**.

A full satisfactory answer would include a general result of partial regularity under condition **Aw**. This result is expected in view of the Euclidean case (since the quadratic cost is really the limit case for condition **Aw**). Note that, in view of counterexamples given in [26], the results under **Aw** can not be as good as under **As**, and can not be much better than Caffarelli's results [6] that require densities bounded away from 0 and infinity.

3.3 Conditions Aw, As for the quadratic cost of a Riemannian manifold

One sided answer to question 4: The cost-sectional curvature yields the sectional curvature

Theorem 3.8 *Let M be a C^4 Riemannian manifold. Let $c(x,y) = d^2(x,y)/2$ for all $(x,y) \in M \times M$. Let $\mathfrak{S}_\mathfrak{c}$ be given by (16). Then, for all $\xi, \nu \in T_xM$,*

$$\frac{\mathfrak{S}_\mathfrak{c}(x,x)(\nu,\xi)}{|\xi|_g^2|\nu|_g^2 - (\xi\cdot\nu)_g^2} = 8 \cdot \text{ Sectional Curvature of } M \text{ at } x \text{ in the 2-plane } (\xi,\nu).$$

Hence if **Aw** *(resp,* **As***) is satisfied at (x,x), the sectional curvature of M at x is non-negative (resp. strictly positive).*

Corollary 3.9 *Let M be a compact Riemannian manifold. If the sectional curvature of M is not everywhere non-negative, there are smooth positive measures on M such that the optimal map (for the cost function $c(x,y) = d^2(x,y)/2$) is not continuous.*

At the end of the proof of Theorem 3.8, we give a counterexample to regularity for a two-dimensional manifold with negative sectional curvature.

This observation closes (with a negative answer) the open problem of the regularity of optimal gradient maps when the manifold does not have non-negative sectional curvature everywhere. Of course, one wonders whether non-negative sectional curvature implies **Aw**, and I could not answer to this question. It seems that when $x \neq y$ the cost-curvature involves derivatives of the curvature tensor along the geodesic linking x to y. However, there is a partial converse assertion in the special case of constant sectional curvature:

The quadratic cost on the round sphere $\mathbb{S}^{n-1}$ satisfies **As**

Theorem 3.10 *Let $\mathbb{S}^{n-1}$ be the unit sphere of $\mathbb{R}^n$ equipped with the round metric g, and Riemannian distance d. Let $c(x,y) = \frac{1}{2}d^2(x,y)$. Then c satisfies* **As** *on $\mathbb{S}^{n-1} \times \mathbb{S}^{n-1} \setminus \{(x,x), x \in \mathbb{S}^{n-1}\}$.*

This Theorem is a corollary of the Proposition 6.2, which shows that condition **As** is satisfied for any choice of local coordinates. Here the cost function is only locally smooth on $\mathbb{S}^{n-1} \times \mathbb{S}^{n-1} \setminus \{(x,x), x \in \mathbb{S}^{n-1}\}$, and it is not enough to show that **As** is satisfied: we have to prove a priori that $G_\phi(x)$ stays away from the cut locus of x, otherwise the Monge-Ampère equation might become singular (namely the term $|D_{xy}c|$ blows up). We use for that a previous a-priori estimate obtained with Delanöe [14], and we can obtain the

Theorem 3.11 *Let $\mathbb{S}^{n-1}$ be the unit sphere of $\mathbb{R}^n$ equipped with the round metric g, and Riemannian distance d, and let $c(x,y) = \frac{1}{2}d^2(x,y)$. Let ϕ be a c-convex potential on $\mathbb{S}^{n-1}$, then for all $x \in \mathbb{S}^{n-1}$, $\partial\phi(x) = \partial^c\phi(x)$, and the set $G_\phi(x) = \{y \in \mathbb{S}^{n-1}, \phi(x) + \phi^c(y) = -c(x,y)\}$ is c-convex with respect to x.*

Let μ_0, μ_1 be two probability measures on $\mathbb{S}^{n-1}$. Assume that $G_{\phi\#}\mu_0 = \mu_1$. Assume that $\mu_1 \geq m\mathrm{dVol}$, for some $m > 0$. Then

1. *If μ_0 satisfies (19), then $\phi \in C^1(\mathbb{S}^{n-1})$, and as in Theorem 3.5, the modulus of continuity of $\nabla\phi$ depends on f in (19).*
2. *If μ_0 satisfies (18) for some $p > n$, then $\phi \in C^{1,\beta}(\mathbb{S}^{n-1})$ with $\beta = \beta(n,p)$ as in Theorem 3.5.*
3. *If μ_0, μ_1 have positive $C^{1,1}$ (resp. C^∞) densities with respect to the Lebesgue measure, then $\phi \in C^{3,\alpha}(\mathbb{S}^{n-1})$ for every $\alpha \in [0,1[$ (resp. $\phi \in C^\infty(\mathbb{S}^{n-1})$).)*

3.4 Examples of costs that satisfy As or Aw

We repeat the collection of cost that was given in [17], and [22].

- $c(x,y) = \sqrt{1 + |x-y|^2}$ satisfies **As**.
- $c(x,y) = -\sqrt{1 - |x-y|^2}$ satisfies **As**.
- $c(x,y) = (1 + |x-y|^2)^{p/2}$ satisfies **As** for $1 \leq p < 2$, $|x-y|^2 < \frac{1}{p-1}$.
- $c(x,y) = |x-y|^2 + |f(x) - g(y)|^2$ $f, g : \mathbb{R}^n \to \mathbb{R}$ convex (resp. uniformly strictly convex) with $|\nabla f|, |\nabla g| < 1$ satisfies **Aw** (resp. **As**).
- $c(x,y) = \pm\frac{1}{p}|x-y|^p, p \neq 0$ and satisfies **Aw** for $p = \pm 2$ or $p = -\frac{1}{2}$ ($-$ only) and **As** for $-\frac{1}{2} < p < 1$ ($-$ only).

- $c(x,y) = -\log|x-y|$ satisfies **As** on $\mathbb{R}^n \times \mathbb{R}^n \setminus \{(x,x), x \in \mathbb{R}^n\}$.
- As pointed out in [22], the reflector antenna problem ([27]) corresponds to the case $c(x,y) = -\log|x-y|$ restricted to $\mathbb{S}^n$. This cost satisfies **As** on $\mathbb{S}^{n-1} \times \mathbb{S}^{n-1} \setminus \{x = y\}$.
- As proved in Theorem 3.10, the Riemannian distance on the sphere satisfies **As**. Note that it is the restriction to $\mathbb{S}^{n-1}$ of the cost $c(x,y) = \theta^2(x,y)$, where θ is the angle formed by x and y.

Acknowledgments At this point I wish to express my gratitude to Neil Trudinger and Xu-Jia Wang, for many discussions, and for sharing results in progress while we were all working on this subject. I also thank Cédric Villani for fruitful discussions, and Philippe Delanoë with whom we started to think about the problem of regularity for optimal transportation on the sphere. I also wish to thank Robert McCann, who first raised to me the issue of the connectedness of the contact set, in 2003.

4 Proof of Theorems 3.1, 3.2

We begin with the following uniqueness result of independent interest:

Proposition 4.1 *Let μ, ν be two probability measures on Ω, Ω', with Ω and Ω' connected domains of $\mathbb{R}^n$. Assume that either μ or ν is positive Lebesgue almost everywhere in Ω (resp. in Ω'). Then, among all pairs of functions (ϕ, ψ) such that ϕ is c-convex, ψ is c^*-convex, the problem (10) has at most one minimizer up to an additive constant.*

The proof of this proposition is deferred to the end of the paper.

4.1 Proof of Theorem 3.1

We will begin with the following lemma:

Lemma 4.2 *Let ϕ be c-convex. Let $(\phi_\epsilon)_{\epsilon>0}$ be a sequence of c-convex potentials that converges uniformly to ϕ on compact sets of Ω. Then, if $p = -\nabla_x c(x_0, y) \in \partial\phi(x_0)$, $x_0 \in \Omega, y \in \Omega'$, there exists a sequence $(x_\epsilon)_{\epsilon>0}$ that converges to x_0, a sequence $(y_\epsilon)_{\epsilon>0}$ that converges to y such that $p_\epsilon = -\nabla_x c(x_\epsilon, y_\epsilon) \in \partial\phi_\epsilon(x_\epsilon)$. Finally, p_ϵ converges to p.*

Proof. Let $y = \mathfrak{T}_{x_0}(p)$, i.e. $p = -\nabla_x c(x_0, y)$. Assuming that Ω, Ω' are bounded, one can choose K large enough so that

$$\tilde{\phi}(x) := \phi(x) + K|x - x_0|^2/2 + c(x,y),$$
$$\tilde{\phi}_\epsilon(x) := \phi_\epsilon(x) + K|x - x_0|^2/2 + c(x,y),$$

are convex. One can also assume, by subtracting a constant that $\tilde{\phi}(x_0) = 0$, and that $\tilde{\phi}(x) \geq 0$ on Ω. Finally, one can assume (by relabelling the sequence) that on $B_r(x_0)$ compactly contained in Ω we have $|\phi_\epsilon - \phi| \leq \epsilon$.

Consider then $\tilde{\phi}_\epsilon^\delta = \tilde{\phi}_\epsilon + \delta|x-x_0|^2/2 - \epsilon$. We have $\tilde{\phi}_\epsilon^\delta(x_0) \leq 0$, and on $\partial B_\mu(x_0)$, with $\mu \leq r$,

$$\begin{aligned} \tilde{\phi}_\epsilon^\delta(z) &\geq \tilde{\phi}(z) + \delta\mu^2/2 - 2\epsilon \\ &\geq \delta\mu^2/2 - 2\epsilon. \end{aligned}$$

By taking $\mu = \epsilon^{1/3}, \delta = 4\epsilon^{1/3}$, we get that $\tilde{\phi}_\epsilon^\delta$ has a local minimum in $B_\mu(x_0)$, hence at some point $x_\epsilon \in B_\mu(x_0)$, we have

$$\partial\phi_\epsilon(x_\epsilon) \ni -\nabla_x c(x_\epsilon, y) - K(x_\epsilon - x_0) - \delta(x_\epsilon - x_0).$$

Then we have $|(K+\delta)(x_\epsilon - x_0)|$ small, and thanks to **A1**, **A2**, there exists y_ϵ close to y such that $\nabla_x c(x_\epsilon, y_\epsilon) = \nabla_x c(x_\epsilon, y) + K(x_\epsilon - x_0) + \delta(x_\epsilon - x_0)$. Thus $-\phi_\epsilon(x) - c(x, y_\epsilon)$ has a critical point at x_ϵ. This implies that $p_\epsilon = -\nabla_x c(x_\epsilon, y_\epsilon) \in \partial\phi_\epsilon(x_\epsilon)$. Finally, since $x_\epsilon \to x, y_\epsilon \to y$, we conclude $p_\epsilon \to p$.

□

Now we prove that $\partial^c\phi = \partial\phi$. In order to do this, we must show that if ϕ is c-convex, if $-\phi(\cdot) - c(\cdot, y)$ has a critical point at x_0, this is a global maximum.

We first have the following observation:

Lemma 4.3 *Let ϕ be c-convex. Assume that $-\phi - c(\cdot, y)$ has a critical point at x_0 (i.e. $0 \in \partial\phi(x_0) + \nabla_x c(x_0, y)$), and that it is not a global maximum. Then ϕ is not differentiable at x_0.*

Proof. Indeed, $-\phi(\cdot) - c(\cdot, y)$ has a critical point at x_0, but we don't have $\phi(x_0) + \phi^c(y) = -c(x_0, y)$. However, there is a point y' such that $\phi(x_0) + \phi^c(y') = -c(x_0, y')$. Hence, $\{-\nabla_x c(x_0, y), -\nabla_x c(x_0, y')\} \in \partial\phi(x_0)$, and we have $\nabla_x c(x_0, y) \neq \nabla_x c(x_0, y')$ from assumption **A1**.

□

We show the following:

Proposition 4.4 *Assume* **D** *holds. Let $p = -\nabla_x c(x_0, y) \in \partial\phi(x_0)$ with ϕ c-convex. Then $-\phi(\cdot) - c(\cdot, y)$ reaches a global maximum at x_0.*

D C^1 c-convex functions are dense in the set $\Big\{\phi$ c-convex on Ω, $G_\phi(\Omega) \subset \Omega'\Big\}$ for the topology of uniform convergence on compact sets of $\bar{\Omega}$.

Proof. Assume the contrary, i.e. that $-\phi(x_1) - c(x_1, y) > -\phi(x_0) - c(x_0, y)$ for some $x_1 \in \Omega$. We use **D**: there exists a sequence of C^1 c-convex potentials $(\phi_\epsilon)_{\epsilon>0}$ that converges to ϕ. We use Lemma 4.2: there will exist a sequence $(x_\epsilon)_{\epsilon>0}$ such that $x_\epsilon \to x_0$ and $\nabla\phi_\epsilon(x_\epsilon) \to -\nabla_x c(x_0, y)$. Let y_ϵ be such that $\nabla\phi_\epsilon(x_\epsilon) = -\nabla_x c(x_\epsilon, y_\epsilon)$. Then $y_\epsilon \to y$. Since ϕ_ϵ is C^1, by Lemma 4.3, x_ϵ, the critical point of $-\phi_\epsilon(\cdot) - c(\cdot, y_\epsilon)$ is necessarily a global maximum. Finally, since ϕ_ϵ converges uniformly to ϕ, we see that $-\phi(\cdot) - c(\cdot, y)$ reaches at x_0 a global maximum.

Lemma 4.5 *1. Let c satisfy* **A0**-**As***, let Ω' be c-convex with respect to Ω. Then* **D** *holds.*

2. Assume Ω, Ω' are uniformly strictly c-(c-) convex with respect to each other, and* **As** *is replaced by condition* **Aw***. Then* **D** *holds.*

Proof. As we will see, this result is implied immediately by [17] (for point 1.) and [22] (for point 2.) combined with Proposition 4.1. Let ϕ be c-convex. Denote $\mu_1 = G_{\phi\#}\mathbf{1}_\Omega \mathrm{dVol}$. Note that from Proposition 4.1, ϕ is the unique up to a constant c-convex potential such that $G_{\phi\#}\mathbf{1}_\Omega \mathrm{dVol} = \mu_1$. Consider a sequence of smooth positive densities $(\mu_1^\epsilon)_{\epsilon>0}$ in Ω' such that $\mu_1^\epsilon \mathrm{dVol}$ converges weakly-$*$ to μ_1, and has same total mass than μ_1. Consider ϕ_ϵ such that $G_{\phi_\epsilon\#}\mathbf{1}_\Omega \mathrm{dVol} = \mu_1^\epsilon \mathrm{dVol}$. From [17] in case 1 and [22] in case 2, ϕ_ϵ is C^2 smooth inside Ω. Then, by Proposition 4.1, up to a normalizing constant, ϕ_ϵ is converging to ϕ, and $\nabla\phi_\epsilon$ is converging to $\nabla\phi$ on the points where ϕ is differentiable. □

Hence, under the assumptions of Lemma 4.5, $\partial\phi(x) = \partial^c\phi(x)$. In view of Proposition 2.11, the equality $\partial\phi(x) = \partial^c\phi(x)$ for all ϕ, x is equivalent to the c-convexity of the set

$$G_\phi(x) = \Big\{ y : \phi(x) + \phi^c(y) = -c(x,y) \Big\}.$$

The proof of Theorem 3.1 is now complete.

□

4.2 Proof of Theorem 3.2

We now show that if **Aw** is violated somewhere in $(\Omega \times \Omega')$, there will exist a c-convex potential for which we don't have $\partial\phi = \partial^c\phi$. In view of Lemma 4.5 and Proposition 4.4, this will imply that this potential can not be a limit of C^1-smooth c-convex potentials. Moreover, using the construction done in the proof of Lemma 4.5, this implies that there exists positive densities μ_0, μ_1 in Ω, Ω' such that the c-convex potential ϕ satisfying $G_{\phi\#}\mu_0 = \mu_1$ is not C^1 smooth.

Assume that for some $x_0 \in \Omega, y \in \Omega', p = -\nabla_x c(x_0, y)$, for some ξ, ν unit vectors in $\mathbb{R}^n$ with $\xi \perp \nu$, one has

$$D^2_{p_\nu p_\nu} \Big[p \to D^2_{x_\xi x_\xi} c(x, \mathfrak{T}_x(p)) \Big] \geq N_0 > 0. \tag{20}$$

Let $y_0 = \mathfrak{T}_{x_0}(p - \epsilon\nu), y_1 = T_{x_0}(p + \epsilon\nu)$, with ϵ small, and recall that $y = \mathfrak{T}_{x_0}(p)$. Hence y is the 'middle' of the c-segment $[y_0, y_1]_x$. Let us define

$$\bar{\phi}(x) = \max \Big\{ -c(x,y_0) + c(x_0,y_0), -c(x,y_0) + c(x_0,y_0) \Big\}. \tag{21}$$

Remark. This function will be used often in the geometric interpretation of **As**, **Aw**. It is the "second simplest" c-convex function, as the supremum of two supporting functions. It plays the role of $(x_1, ..., x_n) \to |x_1|$ in the Euclidean case.

Note first that $\xi \perp \nu$ implies that $\xi \perp (\nabla_x c(x_0, y_1) - \nabla_x c(x_0, y_0))$. Consider near x_0 a smooth curve $\gamma(t)$ such that $\gamma(0) = x_0$, $\dot\gamma(0) = \xi$, and such that for $t \in [-\delta, \delta]$, one has

$$f_0(\gamma(t)) := -c(\gamma(t), y_0) + c(x_0, y_0) = -c(\gamma(t), y_1) + c(x_0, y_1) =: f_1(\gamma(t)).$$

Such a curve exists by the implicit function theorem, and it is C^2 smooth. On γ, we have

$$\bar{\phi} = \frac{1}{2}(f_0 + f_1)$$

since $f_0 = f_1$ on γ. Then we compare $\frac{1}{2}(f_0 + f_1)$ with $-c(x, y) + c(x_0, y)$. By (20) we have

$$\frac{1}{2}\left[D^2_{x_\xi x_\xi} c(x_0, y_0) + D^2_{x_\xi x_\xi} c(x_0, y_1)\right] \geq D^2_{x_\xi x_\xi} c(x_0, y) + c(\epsilon, N_0),$$

where $c(\epsilon, N_0)$ is positive for ϵ small enough. Then of course $\nabla_x c(x_0, y) = \frac{1}{2}[\nabla_x c(x_0, y_0) + \nabla_x c(x_0, y_1)]$. Hence we have, for ϵ small enough,

$$\begin{aligned}
& [-c(\gamma(t), y) + c(x_0, y)] - \bar{\phi}(\gamma(t)) \\
= \; & [-c(\gamma(t), y) + c(x_0, y)] - \frac{1}{2}(f_0 + f_1)(\gamma(t)) \\
= \; & \left[\frac{1}{2}\left[D^2_{xx} c(x_0, y_0) + D^2_{xx} c(x_0, y_1)\right] - D^2_{xx} c(x_0, y)\right] \cdot (\gamma(t) - x_0) \cdot (\gamma(t) - x_0)/2 + o(t^2) \\
= \; & \left[\frac{1}{2}\left[D^2_{x_\xi x_\xi} c(x_0, y_0) + D^2_{x_\xi x_\xi} c(x_0, y_1)\right] - D^2_{x_\xi x_\xi} c(x_0, y)\right] t^2/2 + o(t^2) \\
\geq \; & c(\epsilon, N_0) t^2/2 + o(t^2).
\end{aligned}$$

This will be strictly positive for $t \in [-\delta, \delta] \setminus \{0\}$ small enough, and of course the difference $-\bar{\phi} - [c(x, y) - c(x_0, y)]$ vanishes at x_0. Obviously, the function $\bar{\phi}$ is c-convex, $-\bar{\phi}(\cdot) - c(\cdot, y)$ has a critical point at x_0, and this is not a global maximum. Hence **D** can not hold true, since we have seen that **D** would imply that this situation is impossible.

The proof of Theorems 3.2, 3.3 is complete.

$\square$.

5 Proof of Theorem 3.5

5.1 Sketch of the proof

The key argument of the proof is the geometrical translation of condition **As**, and how it will imply C^1 regularity for ϕ: assume that for c-convex ϕ, both $-\phi(\cdot) - c(\cdot, y_0)$ and $-\phi(\cdot) - c(\cdot, y_1)$ reach a local maximum at $x = 0$. Hence, $-\nabla c(0, y_0)$ and $-\nabla c(0, y_1)$ both belong to $\partial\phi(0)$. From assumption **A1**, we have $\nabla c(0, y_1) \neq \nabla c(0, y_0)$, hence ϕ is not differentiable at 0. Consider y_θ the c-segment with respect to $x = 0$ joining y_0 to y_1. Then, as we will see in Proposition 5.1, condition **As** implies that the functions $\{-\phi(x) - c(x, y_\theta)\}, \theta \in [\epsilon, 1 - \epsilon]$ will also have a local maximum at $x = 0$, and moreover this maximum will be strict in the following sense: we will have

$$-\phi(x) - c(x, y_\theta) \leq -\phi(0) - c(0, y_\theta) - \delta|x|^2/2 + o(|x|^2),$$

with $\delta > 0$ depending on $|y_1 - y_0|$ and $C_0 > 0$ in condition **As**, and bounded by below for θ away from 0 and 1.

Then, by estimating all supporting functions to ϕ on a small ball centered at 0, we will find that $G_\phi(B_\epsilon(0))$ contains a $C\epsilon$ neighborhood of $\{y_\theta, \theta \in [1/4, 3/4]\}$, $C > 0$ depending on C_0 in condition **As**. This is the Proposition 5.3.

Actually, Proposition 5.3 shows that one can find locally supporting functions on some small ball around 0. Hence we show that $\partial\phi$ is big. This is where we use Theorem 3.1 to show that $G_\phi(B_\epsilon(0))$ is large.

One this is shown, we can contradict the bound on the Jacobian determinant of G_ϕ.

We now enter into the rigorous proof of Theorem 3.5.

5.2 Geometric interpretation of condition As

The core of the proof of Theorem 3.5 is the following proposition, which a geometrical translation of assumption **As**. Actually, (this will be an important remark for section 6.3), as we will see in Proposition 5.7, the result of this proposition is equivalent to assumption **As** for a smooth cost function.

Proposition 5.1 *For $y_0, y_1 \in \Omega'$, let $(y_\theta)_{\theta\in[0,1]}$ be the c-segment with respect to $x_0 \in \Omega$ joining y_0 to y_1, in the sense of Definition 2.8. Let*

$$\bar{\phi}(x) = \max\{-c(x, y_0) + c(x_0, y_0), -c(x, y_1) + c(x_0, y_1)\}.$$

Then for $\epsilon \in]0, \frac{1}{2}[$, for all $\theta \in [\epsilon, 1-\epsilon]$, for all $|x - x_0| \le C\epsilon$, we have

$$\bar{\phi}(x) \ge -c(x, y_\theta) + c(x_0, y_\theta) + \delta_0 |y_1 - y_0|^2 \theta(1-\theta)|x - x_0|^2 - \gamma|x - x_0|^3,$$

*where δ_0 depends on $C_0 > 0$ in assumption **As**, the bound in assumption **A2**, γ depends on $\|c(\cdot,\cdot)\|_{C^3}$, and C is bounded away from 0 for $|y_0|, |y_1|$ bounded.*

Proof of Proposition 5.1. Rotating and shifting the coordinate, and subtracting an affine function, we can assume that $x_0 = 0$ and that

$$\begin{aligned}
f_0(x) &:= -c(x, y_0) + c(0, y_0) = ax^1 - D^2_{xx}c(0, y_0)\cdot x \cdot x/2 + o(|x|^2),\\
f_1(x) &:= -c(x, y_1) + c(0, y_1) = bx^1 - D^2_{xx}c(0, y_1)\cdot x \cdot x/2 + o(|x|^2),\\
-c(x, y_\theta) + c(0, y_\theta) &= [\theta b + (1-\theta)a]x^1 - D^2_{xx}c(0, y_\theta)\cdot x \cdot x/2 + o(|x|^2),
\end{aligned}$$

with $a < b$ and where x^i is the coordinate of x in the direction e_i. This means that $y_0 = T_0(ae_1), y_1 = T_0(be_1)$.

Using the general fact that $\max\{f_0, f_1\} \ge \theta f_1 + (1-\theta)f_0$ for $0 \le \theta \le 1$, we have

$$\begin{aligned}
\bar{\phi}(x) \ \ge\ & (\theta b + (1-\theta)a)x^1\\
& -(\theta D^2_{xx}c(0, y_1) + (1-\theta)D^2_{xx}c(0, y_0))\cdot x \cdot x/2 + o(|x|^2).
\end{aligned}$$

Then we use the assumption **As**:

Lemma 5.2 *Under assumption* **As***,*

$$\begin{aligned}-D^2_{xx}c(0,y_\theta)\cdot x\cdot x \;\;&\leq\;\; -((1-\theta)D^2_{xx}c(0,y_0)+\theta D^2_{xx}c(0,y_1))\cdot x\cdot x\\ &\qquad -\delta|x|^2+\Delta|x_1||x|,\end{aligned}$$

where

$$\begin{aligned}&\delta=\frac{1}{2}C_0|y_1-y_0|^2\theta(1-\theta), C_0 \textit{ is given in assumption } \mathbf{As}\ ,\\ &\Delta=\Delta_0|y_1-y_0|^2\theta(1-\theta), \Delta_0 \textit{ depends on } \|c(\cdot,\cdot)\|_{C^4(\Omega\times\Omega')},\\ &\quad \textit{and } \|[\det D_{xy}c]^{-1}\|_{L^\infty(\Omega\times\Omega')}.\end{aligned}$$

Proof. We first observe the following property on real functions: Let $f:\mathbb{R}\to\mathbb{R}$ be C^2, with $f''\geq\alpha>0$. Then we have, for all $t_0,t_1\in\mathbb{R}$,

$$\theta f(t_0)+(1-\theta)f(t_1)\geq f(\theta t_0+(1-\theta)t_1)+\frac{\alpha}{2}\theta(1-\theta)|t_1-t_0|^2.$$

We apply this observation to the function

$$f:t\to -D^2_{xx}c(x,\mathfrak{T}_x(te_1))\cdot x'\cdot x'$$

where x' is equal to $(0,x^2,..,x^n)$, and hence $x'\perp e_1$. ¿From assumption **As**, this function satisfies $f''\geq C_0|x'|^2$. Then, take $t_0=a, t_1=b$ (note that $y_\theta=\mathfrak{T}_{x=0}(\theta b+(1-\theta)a)$), to obtain that

$$\begin{aligned}-D^2_{xx}c(0,y_\theta)\cdot x'\cdot x' \;\;&\leq\;\; -((1-\theta)D^2_{xx}c(0,y_0)+\theta D^2_{xx}c(0,y_1))\cdot x'\cdot x'\\ &\qquad -\frac{1}{2}C_0|x'|^2\theta(1-\theta)|b-a|^2.\end{aligned}$$

Then note that

$$\frac{|b-a|}{|y_0-y_1|}=\frac{|be_1-ae_1|}{|\mathfrak{T}_{x=0}(be_1)-\mathfrak{T}_{x=0}(ae_1)|} \tag{22}$$

is uniformly bounded away from 0 and infinity from assumptions **A1**, **A2**. To conclude the lemma, we now have to get rid of the terms where x^1 appears. First we note that for a C^2 function f, we have

$$\Big|\theta f(t_0)+(1-\theta)f(t_1)-f(\theta t_0+(1-\theta)t_1)\Big|\leq\frac{1}{2}\|f\|_{C^2}\theta(1-\theta)|t_1-t_0|^2.$$

We apply this observation to

$$f:t\to D^2_{xx}c(x,\mathfrak{T}_x(te_1))\cdot x\cdot x-D^2_{xx}c(x,\mathfrak{T}_x(te_1))\cdot x'\cdot x',$$

for which we have $|f''|\leq 2\Delta_0|x^1||x|$, where Δ_0 depends on $\|c(\cdot,\cdot)\|_{C^4}$ and $D_p(p\to\mathfrak{T}_x(p))=[D_{xy}c]^{-1}$. Following then the same lines as above, we conclude the lemma.

□

Using this lemma, we now have

$$\begin{array}{rcl} \bar\phi(x) & \geq & (\theta b + (1-\theta)a)x_1 - D^2_{xx}c(0,y_\theta)\cdot x\cdot x/2 \\ & + & \delta|x|^2 - \Delta|x_1||x| + o(|x|^2), \end{array} \tag{23}$$

with δ, Δ given in Lemma 5.2. We need to eliminate the term $-\Delta|x_1||x|$. In order to do so, notice that

$$|D^2_{xx}c(0,y_\theta)\cdot x\cdot x/2 - D^2_{xx}c(0,y_{\theta'})\cdot x\cdot x/2| \leq C_1|\theta-\theta'||x|^2,$$

where C_1 depends on $|b-a|$, $[D_{xy}c]^{-1}$ and $\|c\|_{C^3}$. Then in (23), using that $\Delta_0|x||x^1| \leq C_0|x|^2/4 + |x^1|^2\Delta_0^2/C_0$, we obtain

$$\begin{array}{rcl} \delta|x|^2 - \Delta|x||x_1| & = & |y_1-y_0|^2\theta(1-\theta)(C_0|x|^2/2 - \Delta_0|x||x_1|) \\ & \geq & |y_1-y_0|^2\theta(1-\theta)(C_0|x|^2/4 - (\Delta_0^2/C_0)|x^1|^2). \end{array}$$

Hence for $\theta, \theta' \in [0,1]$, (23) can be changed in

$$\begin{array}{rl} \bar\phi(x) \geq & (\theta' b + (1-\theta')a)x_1 - D^2_{xx}c(0,y'_\theta)\cdot x\cdot x/2 \\ & +\delta|x|^2/2 + ((b-a)(\theta-\theta') - C_2x^1)x^1 - C_1|\theta-\theta'||x|^2 + o(|x|^2), \end{array}$$

where $C_2 = \theta(1-\theta)|y_1-y_0|^2\Delta_0^2/C_0$. We chose $\epsilon > 0$ small. Taking $\theta' \in [\epsilon, 1-\epsilon]$, we choose θ such that $\theta = \theta' + x^1C_2/(b-a)$. This is possible if we restrict to $|x^1| \leq (b-a)C_2^{-1}\epsilon$, (using (22), note that $(b-a)C_2^{-1}$ is bounded away from 0 for $|y_1-y_0|$ bounded). Note also that θ will depend on x_1 but not θ'. For the term $-C_1|\theta-\theta'||x|^2$, we have $C_1 \leq \bar C_1|b-a|$ where $\bar C_1$ depends on $[D_{xy}c]^{-1}$ and $\|c\|_{C^3}$. Thus, $C_1|\theta-\theta'||x|^2 \leq \bar C_1|b-a||x|^2|x^1|C_2/|b-a| \leq \bar C_1 C_2|x|^3$. Hence we find

$$\begin{array}{l} \forall\theta' \in [\epsilon, 1-\epsilon], \\ \bar\phi(x) \geq (\theta' b + (1-\theta')a)x_1 - D^2_{xx}c(0,y_{\theta'})\cdot x\cdot x/2 + \delta|x|^2/2 + o(|x|^2). \end{array}$$

Noticing that all the terms $o(|x|^2)$ are in fact bounded by $C|D^3c(\cdot,\cdot)|\,|x|^3$, and that $\theta' b + (1-\theta')a = -\nabla_x c(0,y_{\theta'})$, we conclude the Proposition.

□

5.3 Construction of supporting functions

We let $\mathcal{N}_\mu(B)$ denote the μ-neighborhood of a set B, and we use the Proposition 5.1 to prove the following:

Proposition 5.3 *Let ϕ be c-convex. There exists $C > 0$ such that if*

$$|G_\phi(x_1) - G_\phi(x_0)| \geq C|x_1-x_0|^{1/5} > 0,$$

there exists $x_m \in [x_0, x_1]$*, such that, if* $\mathcal{N}_\epsilon([x_0, x_1]) \subset \Omega$*, then we have*

$$\mathcal{N}_\mu(\{y_\theta, \theta \in [\kappa, 1-\kappa]\}) \cap \Omega' \subset G_\phi(B_\epsilon(x_m)),$$

where

$$\begin{aligned} \epsilon &= C\left(\frac{|x_1 - x_0|}{|y_1 - y_0|}\right)^{1/2}, \\ \mu &= C\epsilon|y_1 - y_0|^2. \end{aligned}$$

Here $y_0 = G_\phi(x_0), y_1 = G_\phi(x_1)$*,* $\{y_\theta\}_{\theta\in[0,1]} = [y_0, y_1]_{x_m}$ *denotes the c-segment from* y_0 *to* y_1 *with respect to* x_m*. The constant* C *depends only on* κ *and on the constants in the assumptions* **A0**-**As***, and is bounded away from 0 and infinity for* κ *bounded away from 0.*

PROOF. We assume here that we can find x_0 and x_1 in Ω close such that $\nabla\phi(x_1) - \nabla\phi(x_0)$ is as large as we want compared to $x_1 - x_0$. If this does not happen, then ϕ is $C^{1,1}$. We can assume that $\phi(x_0) = \phi(x_1)$. Then we note $y_0 = G_\phi(x_0), y_1 = G_\phi(x_1)$. From (11), we have

$$\begin{aligned} &\forall x \in \Omega, -c(x_0, y_0) - \phi(x_0) \geq -c(x, y_0) - \phi(x), \\ &\forall x \in \Omega, -c(x_1, y_1) - \phi(x_1) \geq -c(x, y_1) - \phi(x). \end{aligned}$$

Hence

$$\begin{aligned} -c(x, y_0) + c(x_0, y_0) + \phi(x_0) \leq \phi(x), \\ -c(x, y_1) + c(x_1, y_1) + \phi(x_1) \leq \phi(x), \end{aligned}$$

with equality at $x = x_0$ in the first line, at $x = x_1$ in the second line. Using that $\phi(x_0) = \phi(x_1)$, the supporting functions $-c(x, y_0) + c(x_0, y_0) + \phi(x_0)$ and $-c(x, y_1) + c(x_1, y_1) + \phi(x_1)$ will cross at some point x_m on the segment $[x_0, x_1]$. We might suppose that at this point they are equal to 0. Hence

$$-c(x_m, y_0) + c(x_0, y_0) + \phi(x_0) = 0 \tag{24}$$

$$-c(x_m, y_1) + c(x_1, y_1) + \phi(x_1) = 0. \tag{25}$$

Lemma 5.4 *Under the assumptions made above, we have*

$$\phi \leq C_3|x_1 - x_0||y_1 - y_0|,$$

in the segment $[x_0, x_1]$*, where* C_3 *depends on* $\|c(\cdot,\cdot)\|_{C^2(\Omega\times\Omega')}$*.*

PROOF. Using (24,25), we have

$$\begin{aligned} H &= \phi(x_0) \leq -\nabla_x c(x_m, y_0) \cdot (x_0 - x_m) + \|c\|_{C^2}|x_0 - x_m|^2/2 \\ &= \phi(x_1) \leq -\nabla_x c(x_m, y_1) \cdot (x_1 - x_m) + \|c\|_{C^2}|x_1 - x_m|^2/2. \end{aligned}$$

By semi-convexity, on $[x_0, x_1]$ we have $\phi(x) \le H + C|x_1 - x_0|^2$, where C is also controlled by $\|c\|_{C^2(\Omega\times\Omega')}$ Then we assume that $-\nabla_x c(x_m, y_0)\cdot(x_0 - x_m)$ and $-\nabla_x c(x_m, y_1)\cdot(x_1 - x_m)$ are both positive, otherwise we are done, since we have assumed that $|x_1 - x_0|^2$ is small compared to $|x_1 - x_0||y_1 - y_0|$. This implies that

$$\begin{aligned}
-\nabla_x c(x_m, y_0)\cdot(x_0 - x_m) &\le -\nabla_x c(x_m, y_0)\cdot(x_0 - x_1)\\
-\nabla_x c(x_m, y_1)\cdot(x_1 - x_m) &\le -\nabla_x c(x_m, y_1)\cdot(x_1 - x_0).
\end{aligned}$$

Then we have

$$\begin{aligned}
2H &\le -\nabla_x c(x_m, y_0)\cdot(x_0 - x_1) - \nabla_x c(x_m, y_1)\cdot(x_1 - x_0) + \|c\|_{C^2}|x_0 - x_1|^2\\
&\le |\nabla_x c(x_m, y_0) - \nabla_x c(x_m, y_1)||x_0 - x_1| + \|c\|_{C^2}|x_0 - x_1|^2\\
&\le \|c\|_{C^2}\Big(|x_1 - x_0||y_1 - y_0| + |x_0 - x_1|^2\Big),
\end{aligned}$$

Recall that we assume that $|x_1 - x_0|$ is small compared to $|y_1 - y_0|$, otherwise there is nothing to prove; this means that $|x_1 - x_0|^2$ is small compared to $|x_1 - x_0||y_1 - y_0|$, and we conclude.

□

We now use Proposition 5.1 (centered at x_m) that will yield

$$\begin{aligned}
\phi(x) &\ge \max\{-c(x, y_0) + c(x_m, y_0), -c(x, y_1) + c(x_m, y_1)\}\\
&\ge -c(x, y_\theta) + c(x_m, y_\theta) + \delta_0\theta(1-\theta)|y_0 - y_1|^2|x - x_m|^2 - \gamma|x - x_m|^3. \qquad (26)
\end{aligned}$$

for all $\theta \in [\epsilon, 1-\epsilon]$, $|x - x_m| \le C\epsilon$, and with $[y_\theta]_{\theta\in[0,1]}$ the c-segment with respect to x_m joining y_0 to y_1. Note that ϵ is small but fixed once for all.

We want to find supporting functions to ϕ on a ball of suitable radius. For that we consider a function of the form

$$f_y(x) = -c(x, y) + c(x_m, y) + \phi(x_m).$$

Of course, this function coïncides with ϕ at x_m. If we have $f_y \le \phi$ on $\partial B_r(x_m)$, then $-\phi + f_y$ will have a local maximum inside $B_r(x_m)$, hence, for some point $x \in B_r(x_m)$, we will have $-\nabla_x c(x, y) \in \partial\phi(x)$. In view of Theorem 3.1, this implies that $y \in G_\phi(B_r(x_m))$.

First we have

$$\begin{aligned}
& -c(x, y) + c(x_m, y)\\
= \ & -c(x, y_\theta) + c(x_m, y_\theta)\\
+ \ & \int_{s=0}^1 [\nabla_y c(x_m, y_\theta + s(y - y_\theta)) - \nabla_y c(x, y_\theta + s(y - y_\theta))]\cdot(y - y_\theta)\, ds\\
\le \ & -c(x, y_\theta) + c(x_m, y_\theta) + \|D^2_{xy}c\|_{L^\infty(\Omega\times\Omega')}|y - y_\theta||x - x_m|.
\end{aligned}$$

We then have

$$\begin{aligned}
& -c(x, y) + c(x_m, y) + \phi(x_m)\\
\le \ & -c(x, y_\theta) + c(x_m, y_\theta) + C_4|y - y_\theta||x - x_m| + C_3|x_1 - x_0||y_1 - y_0|,
\end{aligned}$$

where $C_4 = \|D^2_{xy}c\|_{L^\infty(\Omega\times\Omega')}$, and we have used Lemma 5.4 to estimate $\phi(x_m)$. Using (26), we want this to be bounded by

$$-c(x, y_\theta) + c(x_m, y_\theta) + \delta_0\theta(1-\theta)|y_0 - y_1|^2|x - x_m|^2 - \gamma|x - x_m|^3$$

on the set $\{|x - x_m| = r\}$, for some $r > 0$. First we restrict θ to $[1/4, 3/4]$, (i.e. we take $\epsilon = 1/4$ in (26)). Then we want

$$\frac{3}{16}\delta_0|y_0 - y_1|^2r^2 - \gamma r^3 \geq C_4|y - y_\theta|r + C_3|x_1 - x_0||y_1 - y_0|.$$

We choose $|y - y_\theta| \leq C_5 r|y_1 - y_0|^2$ for C_5 small enough (for example $C_5 = \delta_0/(16C_4)$), and the above inequality will be satisfied for

$$r^2 = 16C_3\frac{|x_1 - x_0|}{|y_1 - y_0|},$$

if for this value of r, we have indeed $\gamma r \leq (\delta_0/16)|y_0 - y_1|^2$. If not then it means that $|y_1 - y_0|^5 \leq \left(\frac{16\gamma}{\delta_0}\right)^2 C_4|x_1 - x_0|$ and we have therefore a $C^{1,1/5}$ estimate for ϕ.

Now we assume that this is not the case, and therefore the ratio $\frac{|x_1 - x_0|}{|y_1 - y_0|}$ is small. Hence we consider a ball of radius

$$r = C_6\left(\frac{|x_1 - x_0|}{|y_1 - y_0|}\right)^{1/2}$$

centered at x_m, where $C_6 = (16C_3)^{1/2}$.

We denote $\mu = C_5 r|y_1 - y_0|^2$. Note that $C_5 = \delta_0/(16C_4)$, hence C_5 is bounded away from 0. We denote $N_\mu(S)$ the μ neighborhood of a set S. The functions $-c(x, y) + c(x_m, y) + \phi(x_m)$, for $y \in N_\mu\{y_\theta, \theta \in [1/4, 3/4]\}$ will be equal to ϕ at x_m, and will be below ϕ on the boundary of the ball $B_r(x_m)$. This proves Proposition 5.3.

□

5.4 Continuity estimates

Proposition 5.5 *Let ϕ be c-convex with $G_\phi(\Omega) \subset \Omega'$. Assume that c satisfies* **A0**-**As** *in $\Omega\times\Omega'$, and that Ω' is c-convex with respect to Ω. Then,*

- *if $G^\#_\phi$dVol satisfies (19), then $\phi \in C^1_{loc}(\Omega)$;*
- *if $G^\#_\phi$dVol, satisfies (18), for some $p > n$, then $\phi \in C^{1,\beta}_{loc}(\Omega)$, with $\beta(n, p)$ as in Theorem 3.5.*

Remark. We recall that $G^\#_\phi$ is defined in Definition 2.12.

Proof. Consider $\Omega_\delta = \{x \in \Omega, d(x, \partial\Omega) > \delta\}$. In order to have $\mathcal{N}_\epsilon([x_0, x_1]) \subset \Omega$, it is enough to have

1. $x_0, x_1 \in \Omega_\delta$,

2. $|x_0 - x_1| < \delta/2$,

3. $\epsilon < \delta/2$.

Note that, in Proposition 5.3, either $|G_\phi(x_0) - G_\phi(x_1)| \leq C|x_1 - x_0|^{1/5}$, or $\epsilon < D|x_1 - x_0|^{2/5}$, where D depends on C. Hence, for any $\delta > 0$ small, for all $x_0, x_1 \in \Omega_\delta$ such that $|x_0 - x_1| < \delta/2$ and $|x_1 - x_0|^{2/5} < \delta/(2E)$ where E depends on C, if $|G_\phi(x_0) - G_\phi(x_1)| \geq C|x_1 - x_0|^{1/5}$, then $\mathcal{N}_\epsilon([x_0, x_1]) \subset \Omega$ and Proposition 5.3 applies. We now set

$$R_\delta = \inf\{\delta/2, (\delta/(2E))^{5/2}\}. \tag{27}$$

From Proposition 4.4, we will have $N_\mu\{y_\theta, \theta \in [1/4, 3/4]\} \cap \Omega' \subset G_\phi(B_\epsilon(x_m))$. The volume of $N_\mu\{y_\theta, \theta \in [1/4, 3/4]\} \cap \Omega'$ is of order

$$[y_1 - y_0]\mu^{n-1} \sim [y_1 - y_0]\epsilon^{n-1}|y_1 - y_0|^{2(n-1)},$$

(remember that $[y_0, y_1]_{x_m} \subset \Omega'$ by c-convexity of Ω' with respect to Ω) while $B_\epsilon(x_m)$ has a volume comparable to ϵ^n.

$C^{1,\beta}$ **estimates** If the Jacobian determinant of the mapping G_ϕ is bounded, (in other words, if $G_\phi^{\#}\mathrm{dVol}$ has a density bounded in L^∞ with respect to the Lebesgue measure) then the volume of $G_\phi(B_\epsilon(x_m))$ must be of order ϵ^n, hence we get that $|y_1 - y_0|^{2n-1} \leq C\epsilon$, where C depends on all the previous constants C_i. This implies

$$|y_1 - y_0|^{2n-1} \leq C\left(\frac{|x_1 - x_0|}{|y_1 - y_0|}\right)^{1/2},$$

hence

$$|G_\phi(x_1) - G_\phi(x_0)| \leq C|x_1 - x_0|^{\frac{1}{4n-1}},$$

$G_\phi \in C_{loc}^{\frac{1}{4n-1}}(\Omega)$, and we conclude, using $-\nabla_x c(x, y_i) = \nabla\phi(x_i), i = 0, 1$, that

$$|\nabla\phi(x_1) - \nabla\phi(x_0)| \leq C|x_1 - x_0|^{\frac{1}{4n-1}}.$$

We can refine the argument: Let F be defined by

$$F(V) = \sup\Big\{\mathrm{Vol}(G_\phi(B)), B \subset \Omega \text{ a ball of volume } V\Big\}.$$

Let $\mu = G_\phi^{\#}\mathrm{dVol}$, we have $\mathrm{Vol}(G_\phi(B)) = \mu(B)$, hence we have $F(V) \leq \sup_{\mathrm{Vol}(B)=V}\mu(B)$. Then, by Proposition 5.3, we have $F(|B_r(x_m)|) \geq \mathrm{Vol}(N_\mu\{y_\theta, \theta \in [1/4, 3/4]\})$, which gives

$$F\left(\omega_n C_5^n \frac{|x_1 - x_0|^{n/2}}{|y_1 - y_0|^{n/2}}\right) \geq C_7|x_1 - x_0|^{(n-1)/2}|y_1 - y_0|^{(3n-1)/2} \tag{28}$$

with ω_n the volume of the n-dimensional unit ball, and C_7 depends on all the constants C_i above, and it can be checked that C_7 is bounded away from 0. Assume that $F(V) \leq CV^{\kappa}$ for some $\kappa \in \mathbb{R}$. Note that $\mu \in L^p$ implies the (stronger) bound $F(V) = o(V^{1-1/p})$, hence we might write $\kappa = 1 - 1/p$ for some $p \in]1, +\infty]$, and the condition

$$F(V) \leq CV^{1-1/p}$$

is then equivalent to condition (18) for μ. Then we find

$$|y_1 - y_0|^{2n-1+\frac{1}{2}(1-\frac{n}{p})} \leq C_8 |x_1 - x_0|^{\frac{1}{2}(1-\frac{n}{p})}.$$

We see first that we need $p > n$, then we get, setting $\alpha = 1 - n/p$,

$$\begin{aligned}&\forall x_1, x_0 \in \Omega_\delta \text{ such that } |x_1 - x_0| \leq R_\delta,\\ &|G_\phi(x_1) - G_\phi(y_0)| \leq C_9 |x_1 - x_0|^{\frac{\alpha}{4n-2+\alpha}},\end{aligned}$$

where R_δ is given in (27). This yields Hölder continuity for G_ϕ. Then we use that $\nabla\phi(x) = -\nabla_x c(x, G_\phi(x))$ and the smoothness of c to obtain a similar Hölder estimate for $\nabla\phi$.

C^1 estimates If we only assume condition (19) for $\mu = G_{\phi\#}\mathrm{dVol}$, in other words, if $F(V) = o(V^{1-1/n})$, we can write $F(V) \leq \left[f(V^{2/n})\right]^{2n-1} V^{1-1/n}$, for some increasing $f : [0,1] \to \mathbb{R}^+$, with $\lim_{V\to 0} f(V) = 0$. We then have, as $|x_1 - x_0|$ goes to 0, $\frac{|x_1-x_0|}{|y_1-y_0|}$ that goes also to 0 (otherwise there is nothing to prove). Using the special form of F in (28), we get

$$f^{2n-1}\left(C_{10}\frac{|x_1 - x_0|}{|y_1 - y_0|}\right) \geq (C_{11}|y_1 - y_0|)^{2n-1},$$

hence we get that $|y_1 - y_0|$ goes to 0 when $|x_1 - x_0|$ goes to 0. Let g be the modulus of continuity of G_ϕ in Ω_δ, then g satisfies

$$\forall x_1, x_0 \in \Omega_\delta \text{ such that } |x_1 - x_0| \leq R_\delta, \quad f\left(C_{10}\frac{u}{g(u)}\right) \geq C_{11} g(u).$$

This yields a uniform control on the modulus of continuity of G_ϕ: Indeed f can be chosen increasing, thus invertible, and we write

$$u \geq f^{-1}(C_{11} g(u)) \frac{g(u)}{C_{10}}.$$

Note that C_{10} was obtained from C_5, which is bounded away from 0. Then $g(u) \leq \omega(u)$ where ω is the inverse of $z \to f^{-1}(C_{11}z)\dfrac{z}{C_{10}}$. It is easily checked that $\lim_{r\to 0^+} \omega(r) = 0$. This shows the continuity of G_ϕ. Finally we have $\nabla\phi(x) = -\nabla_x c(x, G_\phi(x))$, and the continuity of $\nabla\phi$ is asserted.

$\square$

REMARK 1. The power $\beta = \frac{\alpha}{4n-2+\alpha}$ is not optimal for example if $n = 1, p = +\infty$, for which the $C^{1,1}$ regularity is trivial, but note that in order to obtain this bound, we had to assume that $|y_1 - y_0| \geq |x_1 - x_0|^{1/5}$, and, before, that $|x_1 - x_0| = o(|y_1 - y_0|)$. Hence the conclusion should be: either ϕ is $C^{1,1}$, or ϕ is $C^{1,1/5}$ or ϕ is $C^{1,\beta}$. Note that $\beta \leq 1/7$ for $n \geq 2$.

Proof of Theorem 3.5 In Theorem 3.5, we do not want to assume that $\mu_1 \in L^1(\mathbb{R}^n)$, hence we do not have that $G_\phi^\# \mu_1 = \mu_0$, which would imply $\mu_0(B) = \mu_1(G_\phi(B))$. Hence we need the following proposition to finish the proof:

Proposition 5.6 *Let ϕ be c-convex on Ω, with $G_\phi(\Omega) \subset \Omega'$. Assume that $G_{\phi\#}\mu_0 = \mu_1$. Assume that $\mu_1 \geq m\mathrm{dVol}$ on Ω'. Then for all $\omega \subset \Omega$, we have*

$$\mu_0(\omega) \geq m\mathrm{Vol}(G_\phi(\omega)), \text{ and hence, } G_\phi^\# \mathrm{dVol} \leq \frac{1}{m}\mu_0.$$

Proof. In Ω' we consider $N = \{y \in \Omega', \exists x_1 \neq x_2 \in \Omega, G_\phi(x_1) = G_\phi(x_2) = y\}$. Then $N = \{y \in \Omega', \phi^c \text{ is not differentiable at } y\}$. Hence $\mathrm{Vol}(N) = 0$, and $\mathrm{Vol}(G_\phi(\omega) \setminus N) = \mathrm{Vol}(G_\phi(\omega))$. Moreover, $G_\phi^{-1}(G_\phi(\omega) \setminus N) = G_{\phi^c}(G_\phi(\omega) \setminus N) \subset \omega$. Hence,

$$\mu_0(\omega) \geq \mu_1(G_\phi(\omega) \setminus N) \geq m\mathrm{Vol}(\omega).$$

□

Proof of the boundary regularity This part is easy: under the assumptions of Theorem 3.6, the density μ_0 satisfies (18) with $p > n$ (resp. satisfies (19)). Hence Theorem 3.5 applies and $\phi \in C^{1,\beta}_{loc}(\Omega)$ (resp. $\phi \in C^1_{loc}(\Omega)$). Since Ω_2 is compactly contained in Ω, we conclude the boundary regularity on Ω_2. This proves Theorem 3.6. □

Remark. This proof of the boundary regularity is very simple because we have interior regularity even when μ_0 vanishes. This is not the case for the classical Monge-Ampère equation, and the boundary regularity requires that both Ω and Ω' are convex, and is more complicated to establish (see [5]).

We now show that there is indeed equivalence between assumption **As** at a point x and the conclusion of Proposition 5.1. This is a quantitative version of Theorem 3.3.

Proposition 5.7 *Assume that at a point x_0 for all y_0, y_1, for $y_{1/2}$ the 'middle' point of $[y_0, y_1]_{x_0}$, we have*

$$\bar{\phi}(x) \geq -c(x, y_{1/2}) + c(x_0, y_{1/2}) + \delta_0 |y_0 - y_1|^2 |x - x_0|^2 + O(|x - x_0|^3)$$

with $\bar{\phi}$ as above. Then the cost function satisfies assumption **As** *at x_0 with $C_0 = C\delta_0$, for some constant $C > 0$ that depends on the bound in* **A2**.

Proof. The proof follows the same lines as the proof of Theorem 3.2, and is omitted here.

□

6 Proof of Theorem 3.11

6.1 Monge-Kantorovitch problem on a Riemannian manifold

The reader interested in details should refer to [19]. For M a complete Riemannian manifold, compact and without boundary, with distance function $d(\cdot, \cdot)$. From now, we restrict to the case $c(x, y) = \frac{1}{2}d^2(x, y)$. Then optimal maps for Monge's problem are gradient maps, i.e.

$$T_{opt}(x) = G_\phi(x) := \exp_x(\nabla_g \phi(x)),$$

where $\nabla_g \cdot$ is the gradient with respect to the Riemannian metric g on M, and ϕ is some c-convex potential obtained by the Monge-Kantorovitch dual approach. With intrinsic notations, it has been established [14] that the Monge-Ampère equation reads

$$\mathrm{Jac}\,(\exp_x)\,(\nabla\phi)\,\det\left[\mathcal{H}(c(p,\cdot)+\phi)\right]_{p=G_\phi(x)} = \frac{\mathrm{d}\mu_0/\mathrm{dVol}(x)}{\mathrm{d}\mu_1/\mathrm{dVol}(G_\phi(x))}. \tag{29}$$

where $\mathcal{H}$ is the Hessian endomorphism operator. Note that for any point x, and $p = G_\phi(x)$, $q \to c(q,p) + \phi(q)$ has a critical point at $q = x$, therefore we have in local coordinates

$$\mathcal{H}(c(p,\cdot)+\phi) = g^{ij} D_{ij}(c(p,\cdot)+\phi).$$

We will see hereafter the particular form that $D^2_{xx}c(x,y)$ takes in the case of the round sphere. Finally, in the Riemannian framework, the definitions of sub-differential and c-sub-differential take the following form:

$$\partial\phi(x) = \Big\{p \in \mathfrak{T}_x(M), \phi(\exp_x(v)) \geq \phi(x) + (v,p)_g + o(|v|_g)\Big\},$$
$$\partial^c\phi(x) = \Big\{p \in \mathfrak{T}_x(M), \phi(x) + \phi^c(\exp_x(p)) = -c(x, \exp_x(p))\Big\},$$

where $(v,p)_g$ denotes the scalar product on $\mathfrak{T}_x(M)$ with respect to the metric g, $|v|^2_g = (v,v)_g$.

Strategy of the proof Most of the proof is contained in the following points:

1- Given μ_0, μ_1 satisfying the assumptions of Theorem 3.11, there exists a constant σ such that $d(x, G_\phi(x)) \leq \pi - \sigma$ for all $x \in \mathbb{S}^{n-1}$. Hence, we can reduce locally the problem to an Euclidean problem, and the distance function does not become singular on that set.

2- The assumption **As** is satisfied by the cost function distance squared on the sphere.

Once this is established, we proceed as follows:

Given $x_0 \in \mathbb{S}^{n-1}$ we can build around x_0 a system of geodesic coordinates on the set $\{x, d(x,x_0) \leq R\}$ for $R < \pi$. From point 1-, for r small enough, the graph $\{(x, G_\phi(x)), x \in B_r(x_0)\}$ is included in the set $B_r(x_0) \times B_{\pi-2r}(x_0)$ on which the cost function is C^∞. From point 2- and using [17], a C^4 smooth solution to (13) on $B_r(x_0)$ will enjoy a C^2 a priori estimate at x_0. This estimate will depend only on the smoothness of μ_0, μ_1, on r, and r is small but can be chosen once for all. Then the method of continuity allows to build smooth solutions for any smooth positive densities.

Then, the Theorem 3.1 follows, in particular that $\partial\phi = \partial^c\phi$.

Then Proposition 5.3 holds on $B_r(x_0) \times B_{\pi-2r}(x_0)$. With a straightforward adaption of Proposition 5.6 to the sphere, this yields C^1 estimates in, say, $B_{r/2}(x_0)$, and the Theorem is proved.

6.2 Reduction of the problem to an Euclidean problem

We denote by B_r (resp. $B_r(x)$) a Riemannian ball of radius r (resp. centered at x), $\mathbb{S}^{n-1}$ denotes the unit sphere of $\mathbb{R}^n$.

Uniform distance to the cut locus We show that there exists a subset S^2_σ of $\mathbb{S}^{n-1}\times\mathbb{S}^{n-1}$ on which **A0-A2** are satisfied, and such that the graph of G_ϕ: $\{(x,G_\phi(x)), x\in\mathbb{S}^{n-1}\}$, is contained in S^2_σ. This subset S^2_σ is defined by

$$S^2_\sigma = \Big\{(x,y)\in\mathbb{S}^{n-1}\times\mathbb{S}^{n-1}, d(x,y)\leq \pi-\sigma\Big\} \tag{30}$$

where $\sigma>0$ depends on some condition on μ_0,μ_1.

Proposition 6.1 *Let μ_0,μ_1 be two probability measures on $\mathbb{S}^{n-1}$, let ϕ be a c-convex potential such that $G_{\phi\#}\mu_0=\mu_1$. Assume that there exists $m>0$ such that $\mu_1\geq m\mathrm{dVol}$ and that μ_0 satisfies (19). Then there exists $\sigma>0$ depending on m and , on f in (19), such that $\{(x,G_\phi(x)), x\in\mathbb{S}^{n-1}\}\subset S^2_\sigma$, where S^2_σ is defined in (30).*

Proof: We use [14]; in that paper, it was shown, for ϕ satisfying $G_{\phi\#}\mu_0=\mu_1$, that we have $|d\phi|\leq\pi-\epsilon$, and $\epsilon>0$ depends on $\|\mathrm{d}\mu_1/\mathrm{dVol}\|_{L^\infty}\|\left(\mathrm{d}\mu_0/\mathrm{dVol}\right)^{-1}\|_{L^\infty}$. Considering ϕ^c and G_{ϕ^c} that pushes forward μ_1 onto μ_0, we also have a bound such as $|d\phi|\leq\pi-\epsilon$, where $\epsilon>0$ depends on $\|\mathrm{d}\mu_0/\mathrm{dVol}\|_{L^\infty}\|\left(\mathrm{d}\mu_1/\mathrm{dVol}\right)^{-1}\|_{L^\infty}$.

Here we slightly extend this bound to the case where μ_0 satisfies (19), and $\mu_1\geq m\mathrm{dVol}$. It was shown in [14, Lemma 2] that for all $x_1,x_2\in\mathbb{S}^{n-1}$,

$$d(G_\phi(x_2),x_1')\leq 2\pi\frac{d(G_\phi(x_1),x_1')}{d(x_1,x_2)},$$

where x_1' is the antipodal point to x_1. Hence the set $E_\delta=\{x,d(x,x_1)\geq\delta\}$ is sent by G_ϕ in $B_\epsilon(x_1')$ where

$$\epsilon=2\pi\frac{d(G_\phi(x_1),x_1')}{\delta}.$$

This implies

$$\mu_1(B_\epsilon(x_1'))\geq\mu_0(E_\delta).$$

Taking $\delta>0$ fixed (for example $\delta=\pi/2$), we have

$$\mu_1(B_{4d(G_\phi(x_1),x_1')}(x_1))\geq\inf_z\mu_0(D_z),$$

where D_z is the half-sphere centered at z. Hence

$$\begin{aligned}&\inf_{x\in\mathbb{S}^{n-1}}\Big\{|\pi-d\phi(x)|,|\pi-d\phi^c(x)|\Big\}\\ \geq\ &\inf\Big\{r,\exists x\in\mathbb{S}^{n-1},\ \mu_1(B_{4r}(x))\geq\inf_z\mu_0(D_z)\Big\}\\ :=\ &\sigma.\end{aligned}$$

In particular, if μ_1 satisfies (19), and $\mu_0\geq m\mathrm{dVol}, m>0$, we have

$$\mu_1(B_{4r}(x))\leq\left(\mathrm{Vol}(B_{4r}(x))\right)^{1-1/n}$$

for r small enough, hence $\sigma>0$ and the uniform distance between $G_\phi(x)$ and the cut locus of x is asserted. Considering ϕ^c, the same conclusion holds. □

Construction of a local system of coordinates Given $x_0 \in \mathbb{S}^{n-1}$, we consider a system of geodesic coordinates around x_0, i.e. given a system of orthonormal coordinates at x_0 and the induced system of coordinates on $T_{x_0}(\mathbb{S}^{n-1})$, we consider the mapping $p \in T_{x_0}(\mathbb{S}^{n-1}) \to \exp_{x_0}(p)$. This mapping is a diffeomorphism from $B_R(0) \subset \mathbb{R}^{n-1}$ to $B_R(x_0)$ as long as $R < \pi$. Then, for r small enough, we have

$$S^2_\sigma \cap \{(x,y), x \in B_r(x_0)\} \subset B_r(x_0) \times B_{\pi-\sigma+r} \subset S^2_{\sigma-2r}.$$

We now take $r = \sigma/3$ Hence, $B_r(0)$ is sent in $B_{\pi-2r}(0)$, and the cost function is C^∞ on $B_r(0) \times B_{\pi-2r}(0)$for r small enough.

In this case, we have $\mathfrak{T}_x(p) = \exp_x(p)$, and also

$$D^2_{xy}c(x,y) = [D_v \exp_x]^{-1}(\exp_x^{-1}(y)), \tag{31}$$

where $D_v \exp_x$ is the derivative with respect to v of $v \to \exp_x(v)$. Assumption **A1** is trivial, since y is indeed uniquely defined by $y = \exp_x(p)$. From (31), assumption **A2** is true on any smooth compact Riemannian manifold, since in this case $\mathrm{Jac}(v \to \exp_x v)$ is bounded by above.

6.3 Verification of assumption As: Proof of Theorem 3.10

Here we prove the following, which implies Theorem 3.10.

Proposition 6.2 *Let $x_1, .., x_{n-1} \in \Omega \subset \mathbb{R}^{n-1}$ be a system of local coordinates on $\mathbb{S}^{n-1}$, with $\varphi : \Omega \to \mathbb{S}^{n-1}$. For any $y \in \mathbb{S}^{n-1}$, $x \in \Omega$, let $c(x,y) = \frac{1}{2}d^2(\varphi(x), y)$. Then $c(\cdot,\cdot)$ satisfies assumption* **As** *on Ω, with a constant $C_0 > 0$, with C_0 bounded away from 0 for $\{g_{ij}(x), g^{ij}(x), 1 \le i,j \le n-1, x \in \Omega\}$ bounded, where $[g^{ij}]$=$[g_{ij}]^{-1}$ and g_{ij} are the coefficients of the round Riemannian metric on $\mathbb{S}^{n-1}$ in the local coordinates $x_1, .., x_{n-1}$.*

Remark. It can seem surprising that we do not chose any system of coordinates around y. Indeed, $y = \mathfrak{T}_x(p) = \exp_x(p)$ for $p \in T_x\mathbb{S}^{n-1}$, and p is written with the induced system of coordinates on $T_x\Omega$. Moreover, condition **As** is written in terms of x and p only. Indeed, we only need to know $D^2_{xx}c(x, \exp_x(p))$ to check condition **As**.

Computations in normal coordinates It has been established in [14] that in a system of normal coordinates $e_1, ..e_{n-1}$ at x with $e_1 = p/|p|$, we have

$$D^2_{x_ix_j}c|_{x,\exp_x(p)} = \begin{cases} 1 \text{ if } i = j = 1 \\ \delta_{ij}\dfrac{|p|\cos(|p|)}{\sin(|p|)} \text{ otherwise.} \end{cases}$$

This relies on the fact that $D^2_{xx}c(x, \exp_x(p)) = [D_p(\exp_x(p))]^{-1}D_x(\exp_x(p))$, and follows by computations of Jacobi fields. Hence, we have, in this system of coordinates,

$$M(x,p) := D^2_{xx}c(x, \mathfrak{T}_x(p)) = I - (1 - \frac{r\cos r}{\sin r})I',$$

where $r = |p|$, I is the identity matrix of order $n-1$, and I' is the identity matrix on $\mathrm{vect}(e_2, .., e_{n-1})$. So, for a given $v \in \mathbb{R}^{n-1}$,

$$M(x,p) \cdot v \cdot v = |v|^2 - (1 - \frac{r \cos r}{\sin r})|\Pi v|^2,$$

where Πv is the projection of v on $p^\perp$. This can be written, using intrinsic notations

$$\big(\mathcal{H}(c(\cdot, y))|_{y=\exp_x(z)} v, v\big)_g = |v|_g^2 - (1 - \frac{|z|_g \cos|z|_g}{\sin|z|_g})|\Pi v|_g^2,$$

where $(\cdot,\cdot)_g$ denotes the Riemannian scalar product on $\mathfrak{T}_x(\mathbb{S}^{n-1})$ and $|p|_g^2 = (p,p)_g$.

Let $\alpha, \beta \in \mathbb{R}^{n-1}$, $t \in \mathbb{R}$, $u_t = \alpha + t\beta$ and $r = |u_t|$. We assume that $r < \pi$, and $\Pi_t v$ is the projection of v on $u_t^\perp$. Assumption **As** is then equivalent to show

$$\frac{d^2}{dt^2}\left[t \to M(x, \alpha + \beta t) \cdot v \cdot v\right] \leq -C_0 |\beta|^2 |v|^2,$$

for all $\alpha, \beta \in \mathfrak{T}_x\mathbb{S}^{n-1}$, $v \perp \beta$, which is equivalent to

$$\frac{d^2}{dt^2}\left[t \to (1 - \frac{r \cos r}{\sin r})|\Pi_t v|^2\right] \geq C_0 |\beta|^2 |v|^2,$$

for all $\alpha, \beta \in \mathbb{R}^{n-1}$, $v \perp \beta$. Without loss of generality, we assume by changing t in $t - t_0$ that $\alpha \perp \beta$. We have

$$\Pi_t v = v - |u_t|^{-2}(v \cdot u_t)u_t,$$

hence, using $v \perp \beta$,

$$\begin{aligned} |\Pi_t v|^2 &= |v|^2 - 2(v \cdot u_t)^2 |u_t|^{-2} + (v \cdot u_t)^2 |u_t|^{-2} \\ &= |v|^2 - (v \cdot \alpha)^2 |u_t|^{-2}. \end{aligned}$$

Now, assuming $|v| = 1$ and $(v, \alpha)_g = c$, we have to evaluate $\frac{d^2}{dt^2}F$, where $F = G \circ r$, and

$$\begin{aligned} &G(r) = (1 - \frac{r \cos r}{\sin r})(1 - \frac{c^2}{r^2}), \\ &r(t) = |\alpha + t\beta|. \end{aligned}$$

A computation shows that

$$\begin{aligned} \frac{d^2}{dt^2}F &= \frac{\alpha^2\beta^2}{r^3}\left\{\frac{1}{\sin^2 r}(r - \sin r \cos r)(1 - \frac{c^2}{r^2})\right. \\ &\qquad \left. + \frac{1}{\sin r}(\sin r - r\cos r)\frac{2c^2}{r^3}\right\} \\ &+ \frac{t^2\beta^4}{r^2}\left\{\frac{2}{\sin^3 r}(\sin r - r \cos r)(1 - \frac{c^2}{r^2})\right. \\ &\qquad + \frac{2}{\sin^2 r}(r - \sin r \cos r)\frac{2c^2}{r^3} \\ &\qquad \left. - \frac{1}{\sin r}(\sin r - r \cos r)\frac{6c^2}{r^4}\right\}. \end{aligned}$$

Notice that we have for some $C > 0$, and for all $r \in [0, \pi]$,

$$\sin r - r\cos r \geq Cr^3, \tag{32}$$
$$r - \sin r \cos r \geq Cr^3. \tag{33}$$

Hence, all terms are positive, except the last one. The sum of the last two lines has for $r \in [0, \pi]$ the sign of

$$\begin{aligned} & 2(r^2 - r\sin r\cos r) - 3(\sin^2 r - r\sin r\cos r) \\ = \; & \frac{1}{2}\left(r\sin 2r + 3\cos 2r + 4r^2 - 3\right). \end{aligned}$$

One can check (by two successive differentiations) that this quantity is non-negative for $r \in [0, \pi]$. Hence we have, using (32, 33),

$$\begin{aligned} \frac{d^2}{dt^2}F &\geq C\beta^2\left[(1 - \frac{c^2}{r^2})\frac{\alpha^2 + t^2\beta^2}{r^2} + \frac{\alpha^2 c^2}{r^4}\right] \\ &= C\beta^2\left[(1 - \frac{c^2}{r^2}) + \frac{\alpha^2 c^2}{r^4}\right]. \end{aligned}$$

Remember that $c = \alpha \cdot v$, with $|v| = 1$, hence $|\alpha| \geq |c|$ and $r \geq |c|$, to conclude that $\frac{d^2}{dt^2}F \geq C_0|\beta|^2$.

Hence in a normal system of coordinates at x, we see that **As** is satisfied at x for any y such that $d(x, y) < \pi$. Through propositions 5.1 and 5.7, we see that condition **As** is satisfied in any system of coordinates, with a constant C_0 that depends on the choice of the coordinates as in Proposition 6.2. This proves Proposition 6.2.

□

6.4 Conclusion of the proof

C^2 a priori estimates for smooth solutions Let ϕ be a C^4 smooth c-convex potential such that $G_{\phi\#}\mu_0 = \mu_1$, with μ_0, μ_1 having positive $C^{1,1}$ smooth densities. Consider a system of geodesic coordinates around x_0. As we have seen, using Proposition 6.1, in this system of coordinates, for $r_0 = \sigma/3$, we have

$$\{(x, G_\phi(x)), x \in B_{r_0}(0)\} \subset B_{r_0}(0) \times B_{\pi - 2r_0}(0) \subset S^2_{\sigma/3}.$$

In this system of coordinates, ϕ will satisfy on $B_{r_0}(0)$ the Monge-Ampère equation (13) where c is $d^2/2$ expressed in our local coordinates. Moreover, c satisfies **A0-As** on $B_{r_0}(0) \times B_{\pi - 2r_0/3}(0)$.

Then from [17, Theorem 4.1], we have a bound on $D^2\phi(x_0)$ that depends only on r_0 and on the bounds on μ_0, μ_1. Finally, note r_0 can be chosen once for all once σ is known.

Smooth solutions In [13] it was established that given a C^∞ smooth c-convex potential ϕ, a C^∞ smooth positive measure μ_0, and $\mu_1 = G_{\phi\#}\mu_0$, the operator

$$F : \phi \to F(\phi) = G_{\phi\#}\mu_0$$

was locally invertible in C^∞ around μ_1. Hence the existence, for a given pair of C^∞ smooth positive probability measures μ_0, μ_1 of a C^∞ smooth c-convex potential ϕ was granted provided one could obtain a-priori estimates for the second derivatives. Indeed, this follows from the concavity of the equation and and the well known continuity method (see [16]).

We conclude the following: for μ_0, μ_1 having $C^{1,1}$ smooth probability densities, there exists a (unique up to a constant) c-convex potential $\phi \in C^{3,\alpha}$ for every $\alpha \in [0, 1[$ such that $G_{\phi\#}\mu_0 = \mu_1$. If moreover μ_0, μ_1 are C^∞, $\phi \in C^\infty$. This concludes the third point of Theorem 3.11.

C^1 **regularity for weak solutions** Having proved the existence of smooth solutions for smooth positive densities, we have the first point of Theorem 3.11. Then, using again the system of geodesic coordinates, we have $G_\phi(B_{r_0}(0)) \subset B_{R_0}(0)$ for some $R_0 < \pi$. We can now apply Proposition 5.3 on $B_{r_0}(0) \times B_{R_0}(0)$ (see the remark after).

Note that $\nabla_x c(0, B_R(0)) = B_R(0)$, which is strictly convex. Hence, c being C^4 on $B_r(0) \times B_R(0)$, for r small enough, $\nabla_x c(x, B_R(0))$ is convex for all $x \in B_r(0)$, and $B_R(0)$ is c-convex with respect to $B_r(0)$. Hence we can now apply Proposition 5.5 on $B_r(0) \times B_R(0)$.

Then we can adapt with almost no modification Proposition 5.6 as follows:

Proposition 6.3 *Let* $\phi : \mathbb{S}^{n-1} \to R$ *be c-convex. Assume that* $G_{\phi\#}\mu_0 = \mu_1$. *Assume that* $\mu_1 \geq m\mathrm{dVol}$ *on* Ω'. *Then for all* $\omega \subset \mathbb{S}^{n-1}$, *we have*

$$\mu_0(\omega) \geq m\mathrm{Vol}(G_\phi(\omega)).$$

Hence, we can conclude in the same way the continuity of $G_\phi, \nabla\phi$. This achieves the proof of Theorem 3.11

□

7 Proof of Theorem 3.7

For this case it has already been checked (see [22] for example) that the cost function $c(x, y) = -\log|x - y|$ satisfies **As** for $x \neq y$. We will just prove that under some assumptions on the measures μ_0, μ_1 we can guarantee that $G_\phi(x)$ stays away from x. Then, the proof of the Theorem 3.7 follows the same lines as the proof of Theorem 3.11. We prove the

Proposition 7.1 *Let* $\mathbb{S}^{n-1}$ *be unit sphere of* $\mathbb{R}^n$. *Let* $c(x, y) = -\log|x - y|$. *Let* $T : \mathbb{S}^{n-1} \to \mathbb{S}^{n-1}$ *be a 2-monotone map, i.e. such that*

$$\forall x_1, x_2 \in \mathbb{S}^{n-1}, c(x_1, T(x_1)) + c(x_2, T(x_2)) \leq c(x_2, T(x_1)) + c(x_1, T(x_2)). \tag{34}$$

Let μ_0, μ_1 *be two probability measures on* $\mathbb{R}^n$. *Assume that* $\mu_1 \geq m\mathrm{dVol}$ *for* $m > 0$, *and that* μ_0 *satisfies (19). Then there exists* $\epsilon_0 > 0$ *depending only on* m, f *in (19) such that*

$$\forall x \in \mathbb{S}^{n-1}, d(x, T(x)) \geq \epsilon_0.$$

Proof. We follow the same lines as in [14]. From (34), we have

$$\log|x_1 - T(x_2)| \leq \log|x_1 - T(x_1)| + \log 2 - \log|x_2 - T(x_1)|.$$

Letting $M = -\log|x_1 - T(x_1)|$, the set $\{x : |x - T(x_1)| \geq 1\}$ is sent by T in the set $\{y : |y - x_1| \leq 2\exp(-M)\}$. Then the bounds on μ_0, μ_1 yield an upper bound on M as in the proof of Proposition 6.1.

□

8 Proof of Theorem 3.8

We consider condition **Aw** at $(x_0, y = x_0)$. We recall that

$$\mathfrak{S}_{\mathfrak{c}}(x_0, x_0)(\xi, \nu) = -D^2_{p_\nu p_\nu} D^2_{x_\xi x_\xi}[(x, p) \to c(x, \mathfrak{T}_{x_0}(p))].$$

for any ν, ξ in $T_{x_0}M$. Let us first take a normal system of coordinates at x_0, so that we will compute

$$Q = D^2_{tt} D^2_{ss}[(x, p) \to c(\mathfrak{T}_{x_0}(t\xi), \mathfrak{T}_{x_0}(s\nu))].$$

Let us write a finite difference version of this operator. We first introduce $y_- = \mathfrak{T}_{x_0}(-h\nu), y_+ = \mathfrak{T}_{x_0}(h\nu), x_- = \mathfrak{T}_{x_0}(-h\xi), x_+ = \mathfrak{T}_{x_0}(h\xi)$. We use the usual second order difference quotient, for example

$$D^2_{x_\xi, x_\xi} c(x, \mathfrak{T}_{x_0}(p)) = \frac{1}{h^2}(c(x_+, x_0) - 2c(x_0, x_0) + c(x_-, x_0)).$$

(Of course we have $c(x_0, x_0) = 0$.) We will have, as h goes to 0,

$$\lim_{h \to 0} \frac{1}{h^4} \left(\sum_{i,j=+,-} c(x_i, y_j) - 2 \sum_{i=+,-} (c(x_i, x_0) + c(y_j, x_0)) \right) = -Q.$$

Rearranging the terms, we find that the first line is equal to

$$\sum_{i,j=+,-} [c(x_i, y_j) - c(x_0, x_i) - c(x_0, y_j)].$$

Each of the terms inside brackets has a simple geometric interpretation: consider the triangle with vertices (x_0, x_i, y_j) whose sides are geodesics. This is a square angle triangle. If the metric is flat, by Pythagoras Theorem, the term inside the brackets is 0. In the general case, a standard computation shows that it is equal to $-2\kappa(x_0, \xi, \nu)h^4 + o(h^4)$ where $\kappa(x_0, \xi, \nu)$ is the sectional curvature at x_0 in the two-plane generated by ξ, ν. Hence, we get that $Q = 8\kappa(x, \xi, \nu)$.

Now to reach the more general formula of Theorem 3.8, we use the following expansion of the distance that Cédric Villani communicated us:

Lemma 8.1 *Let M be a smooth Riemannian manifold. Let γ_1, γ_2 be two unit speed geodesics that leave point $x_0 \in M$. Let θ be the angle between $\dot\gamma_1(0)$ and $\dot\gamma_2(0)$ (measured with respect to the metric), let κ be the sectional curvature of M at x_0 in the 2-plane generated by $\dot\gamma_1(0), \dot\gamma_2(0)$. Then we have*

$$d^2(\gamma_1(t), \gamma_2(t)) = 2(1-\cos(\theta))(1-\kappa(\cos^2(\theta/2))t^2 + O(t^4))^2.$$

Then, we obtain easily, following the same lines as in the case looked above that

$$\mathfrak{S}_\mathfrak{c}(x_0, x_0)(\xi, \nu) = 8\kappa(x_0, \xi, \nu)(|\xi|_g^2|\nu|_g^2 - (\xi, \nu)_g^2),$$

where $(\cdot,\cdot)_g, |\cdot|_g$ denote respectively the scalar product and the norm with respect to g. This proves the Theorem. □

8.1 Counterexample to regularity for a manifold with negative curvature

Consider the two dimensional surface $H = \{z = x^2 - y^2\} \subset \mathbb{R}^3$, with its canonical Riemannian metric. H has negative sectional curvature around 0. Then for r sufficiently small, $\Omega = H \cap B_r(0)$ is c-convex with respect to itself. Consider the function

$$\bar\phi(x) = \max\{-d^2/2(X, X_0), -d^2/2(X, X_1)\},$$

where $X_0 = (0, a, -a^2), X_1 = (0, -a, -a^2)$. Then, as shown by our proof of Theorem 3.2, for a small enough, no sequence of C^1 c-convex potentials can converge uniformly to $\bar\phi$ on Ω. Let μ_0 to be the Lebesgue measure of Ω, and $\mu_1 = \frac{1}{2}(\delta_{X_0} + \delta_{X_1})$. We have $G_{\bar\phi\#}\mu_0 = \mu_1$. Let $\mu_1^\epsilon \in C^\infty(\bar\Omega)$ be a positive mollification of μ_1 so that its total mass remains equal to 1, and that preserves the symmetries with respect to $x = 0$ and $y = 0$. Let ϕ_n be such that $G_{\phi_n\#}\mu_0 = \mu_n$. Then, for n large enough, ϕ_n is not differentiable at the origin. Indeed, for symmetry reasons, 0 belongs to the subdifferential of ϕ_n at 0, on the other hand, ϕ_n converges uniformly to $\bar\phi$, and we know from the fact that **Aw** is violated at 0 that $-\bar\phi - c(\cdot, 0)$ does not reach its global maximum on Ω at 0.

9 Appendix

Proof of Proposition 3.4. We prove only the last point, the other points being elementary. Consider on $\mathbb{R}^n$ a measure locally equal to $\mu_0 = \mathcal{L}^{n-1} \otimes \mu$, where $\mathcal{L}^{n-1}$ is the $n-1$-dimensional Lebesgue measure, and μ is a probability measure on $[0,1]$ equal to the derivative of the Devil's staircase. Then, $\mu_0 \notin L^1$. On the other hand, for all $[a,b] \subset [0,1]$, $\mu([a,b]) \leq |b-a|^\alpha$, for some $\alpha \in (0,1)$. Then, for $x = (x_1, .., x_n)$,

$$\mu_0(B_r(x)) \leq Cr^{n-1}\mu[x_n - r, x_n + r]) \leq Cr^{n-1+\alpha} = Cr^{n(1-1/p)}$$

for some $p > n$. Hence $\mu_0 \notin L^1_{loc}$ and μ_0 satisfies (18) for some $p > n$.

Proof of Proposition 4.1 We know (see [24, chapter 2]) that there exists π a probability measure on $\mathbb{R}^n \times \mathbb{R}^n$, with marginals μ and ν, and such that

$$\int_{\mathbb{R}^n} \phi(x)d\mu(x) + \psi(x)d\nu(x) = -\int c(x,y)d\pi(x,y),$$

and moreover, there exists $\bar{\phi}$ a c-convex potential such that

$$supp(\pi) \subset \{(x, G_{\bar{\phi}}(x)), x \in \mathbb{R}^n\}.$$

Let us decompose π as $\pi = \mu \otimes \gamma_x$, where for $d\mu$ almost all $x \in \mathbb{R}^n$, γ_x is a probability measure on $\mathbb{R}^n$ and γ_x is supported in $G_{\bar{\phi}}(x)$. Hence we have

$$\int_{\mathbb{R}^n} d\mu(x) \left[\int_{\mathbb{R}^n} d\gamma_x(y)(\phi(x) + \psi(y) - c(x,y)) \right] = 0.$$

This implies that for $d\mu$ a.e. x, for $d\gamma_x$ a.e. y, we have $y \in G_\phi(x)$. Since for $d\mu$ a.e. x, we have $y \in G_{\bar{\phi}}(x)$ $d\gamma_x$ a.s., we deduce that for $d\mu$ a.e. x, (and hence for Lebesgue a.e. x, since $\mu > 0$ a.e.), we have $G_{\bar{\phi}}(x) \cap G_\phi(x) \neq \emptyset$. This implies that $\nabla\phi = \nabla\bar{\phi}$ Lebesgue a.e., and that $\phi - \bar{\phi}$ is constant. This shows that ϕ is uniquely defined up to a constant. Now the pair ψ^{c*}, ψ can only improve the infimum (10) compared to (ϕ, ψ), hence it is also optimal. Hence ψ^{c*} is also uniquely defined up to a constant. If ψ is c*-convex, then $\psi^{c*c} = \psi$, and ψ is thus uniquely defined.

□

References

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

[2] L. A. Caffarelli. Interior $W^{2,p}$ estimates for solutions of the Monge-Ampère equation. *Ann. of Math. (2)*, 131(1):135–150, 1990.

[3] L. A. Caffarelli. A localization property of viscosity solutions to the Monge-Ampère equation and their strict convexity. *Ann. of Math. (2)*, 131(1):129–134, 1990.

[4] L. A. Caffarelli. Some regularity properties of solutions of Monge Ampère equation. *Comm. Pure Appl. Math.*, 44(8-9):965–969, 1991.

[5] L. A. Caffarelli. Boundary regularity of maps with convex potentials. *Comm. Pure Appl. Math.*, 45(9):1141–1151, 1992.

[6] L. A. Caffarelli. The regularity of mappings with a convex potential. *J. Amer. Math. Soc.*, 5(1):99–104, 1992.

[7] L. A. Caffarelli. Boundary regularity of maps with convex potentials. II. *Ann. of Math. (2)*, 144(3):453–496, 1996.

[8] L. A. Caffarelli and X. Cabré. *Fully nonlinear elliptic equations*, volume 43 of *American Mathematical Society Colloquium Publications*. American Mathematical Society, Providence, RI, 1995.

[9] L. A. Caffarelli, C. E. Gutiérrez, and Q. Huang. On the regularity of reflector antennas. *Ann. of Math. (2)*, to appear.

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

[11] D. Cordero-Erausquin, R. McCann, and M. Schmuckenschläger. A Riemannian interpolation inequality à la Borell, Brascamp and Lieb. *Inventiones math.*, 146:219–257, 2001.

[12] P. Delanoë. Classical solvability in dimension two of the second boundary-value problem associated with the Monge-Ampère operator. *Ann. Inst. H. Poincaré Anal. Non Linéaire*, 8(5):443–457, 1991.

[13] P. Delanoë. Gradient rearrangement for diffeomorphisms of a compact manifold. *Diff. Geom. Appl.*, 20:145–165, 2004.

[14] P. Delanoë and G. Loeper. C^1 a priori estimates for potentials of invertible gradient mappings on the sphere. *Calc. Var. Partial Differential Equations.* in press.

[15] Wilfrid Gangbo and Robert J. McCann. The geometry of optimal transportation. *Acta Math.*, 177(2):113–161, 1996.

[16] D. Gilbarg and N. S. Trudinger. *Elliptic partial differential equations of second order*, volume 224 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, second edition, 1983.

[17] X.-N. Ma, N. Trudinger, and X.-J. Wang. Regularity of potential functions of the optimal transport problem. *Arch. Ration. Mech. Anal.*, 177(2):151–183, 2005.

[18] R. J. McCann. A convexity principle for interacting gases. *Adv. Math.*, 128(1):153–179, 1997.

[19] R. J. McCann. Polar factorization of maps on Riemannian manifolds. *Geom. Funct. Anal.*, 11(3):589–608, 2001.

[20] F. Otto and C. Villani. Generalization of an inequality by Talagrand and links with the logarithmic Sobolev inequality. *J. Funct. Anal.*, 173(2):361–400, 2000.

[21] K.-T. Sturm. On the geometry of metric-measure spaces. *Acta Mathematica*, to appear.

[22] N. Trudinger and X.-J. Wang. On the second boundary value problem for monge-ampère type equations and optimal transportation. *preprint.*

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

[24] C. Villani. *Topics in optimal transportation*, volume 58 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2003.

[25] C Villani and J. Lott. Ricci curvature for metric-measure spaces via optimal transport. *Ann. of Maths. 2*, to appear.

[26] X.-J. Wang. Some counterexamples to the regularity of Monge-Ampère equations. *Proc. Amer. Math. Soc.*, 123(3):841–845, 1995.

[27] X.-J. Wang. On the design of a reflector antenna. *Inverse Problems*, 12(3):351–375, 1996.

[28] X.-J Wang. On the design of a reflector antenna. II. *Calc. Var. Partial Differential Equations*, 20(3):329–341, 2004.

Grégoire Loeper
Université Claude Bernard Lyon 1
43 boulevard du 11 novembre 1918
69622 Villeurbanne cedex France
loeper@math.univ-lyon1.fr