



**HAL**  
open science

# Strong $c$ -concavity and stability in optimal transport

Anatole Gallouët, Quentin Mérigot, Boris Thibert

► **To cite this version:**

Anatole Gallouët, Quentin Mérigot, Boris Thibert. Strong  $c$ -concavity and stability in optimal transport. 2022. hal-03735775

**HAL Id: hal-03735775**

**<https://hal.science/hal-03735775>**

Preprint submitted on 21 Jul 2022

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# STRONG C-CONCAVITY AND STABILITY IN OPTIMAL TRANSPORT

ANATOLE GALLOUËT, QUENTIN MÉRIGOT, AND BORIS THIBERT

ABSTRACT. The stability of solutions to optimal transport problems under variation of the measures is fundamental from a mathematical viewpoint: it is closely related to the convergence of numerical approaches to solve optimal transport problems and justifies many of the applications of optimal transport. In this article, we introduce the notion of strong  $c$ -concavity, and we show that it plays an important role for proving stability results in optimal transport for general cost functions  $c$ . We then introduce a differential criterion for proving that a function is strongly  $c$ -concave, under an hypothesis on the cost introduced originally by Ma-Trudinger-Wang for establishing regularity of optimal transport maps. Finally, we provide two examples where this stability result can be applied, for cost functions taking value  $+\infty$  on the sphere: the reflector problem and the Gaussian curvature measure prescription problem.

## CONTENTS

1. Introduction	1
2. Stability under strong $c$ -concavity	5
3. Sufficient condition for strong $c$ -concavity	11
4. Stability of optimal transport map for MTW cost	18
5. Stability for the reflector cost on the sphere	19
6. Prescription of Gauss curvature measure	23
References	26

## 1. INTRODUCTION

The theory of optimal transport has had an important impact in applied mathematics, with applications in inverse problems, in variational modeling of evolution PDEs [25, 24], and in machine learning [23] to name but a few. Numerical applications of this theory have been made possible thanks to the tremendous progress of optimal transport solvers in the last decade [23, 21, 3].

The stability of solutions to optimal transport problems under variation of the data is fundamental from a mathematical viewpoint, making optimal transport a “well-posed” problem in the terminology of Hadamard. The question of *quantitative stability* is also of prime importance. The first and most obvious reason is that it is strongly related to the convergence of many

numerical approaches to solve optimal transport problems — both in statistical and in numerical analysis contexts — and explicitly or implicitly it justifies most of the applications of optimal transport. Quantitative stability is at the heart of several other applications, including the understanding of geometric embeddings of spaces of probability measures to Hilbert spaces used in statistics [10], the convergence analysis of numerical methods for evolution equations using optimal transport as a building block [6], the estimation of transport maps in high dimension [15] or the construction of precise asymptotics for random matching problems [2].

The stability of optimal transport plans can be established in a very general setting [26], under variations of the source and target measures, and even under variations of the cost. However, the question of *quantitative* stability has only been addressed rather recently, and most of the existing results deal with the cost function  $c(x, y) = \|x - y\|^2$  [14, 4, 10, 17], or with the squared geodesic distance on a Riemannian manifold [2]. The aim of this article is to establish stability results for more general cost functions, namely those that satisfy the *strong Twist* and *Ma-Trudinger-Wang* conditions on manifolds. We also identify *strong c-concavity* of the Kantorovitch potential as a central notion to get stability results.

**Optimal transport.** Let  $M, N$  be two Polish spaces, let  $\mu \in \mathcal{P}(M)$ ,  $\nu \in \mathcal{P}(N)$  be two probability measures on  $M$  and  $N$  and let  $c : M \times N \rightarrow \mathbb{R} \cup \{+\infty\}$  be a lower semi-continuous cost function which is bounded below. A *transport map* between  $\mu$  and  $\nu$  is a map  $T : M \rightarrow N$  such that the image measure  $T_{\#}\mu$  equals  $\nu$ . Monge’s optimal transport problem between  $\mu$  and  $\nu$  for the cost  $c$  amounts to finding a map  $T : M \rightarrow N$  that minimizes

$$\inf_{T_{\#}\mu=\nu} \int_M c(x, T(x)) d\mu(x). \quad (\text{MP})$$

Such a map, if it exists, is called an *optimal transport map* between  $\mu$  and  $\nu$ . Existence and uniqueness of such an optimal transport map is obtained for instance when the transport cost is quadratic, i.e.  $c(x, y) = \|x - y\|^2$ , when  $M, N$  are compact subsets of  $\mathbb{R}^d$ , and when  $\mu$  is absolutely continuous with respect to the Lebesgue measure [7]. Existence and uniqueness also hold for more general cost functions satisfying a so-called “twist” hypothesis [12].

Kantorovich’s relaxation consists in minimizing the same quantity, but among *transport plans*  $\Gamma(\mu, \nu)$ :

$$\min_{\gamma \in \Gamma(\mu, \nu)} \int_{M \times N} c(x, y) d\gamma(x, y). \quad (\text{KP})$$

We recall that a transport plan between  $\mu$  and  $\nu$  is a probability measure  $\gamma \in \mathcal{P}(M \times N)$  with marginals  $\mu$  and  $\nu$ . Under mild assumptions (e.g.  $c$  is lower-semicontinuous and bounded below), a minimizer to (KP) always exists – but uniqueness may fail. A minimizer to (KP) is called an optimal transport plan.

**Existing stability results.** The problem of stability of optimal transport maps can be expressed as a continuity property of the map  $(\mu, \nu) \mapsto T_{\mu \rightarrow \nu}$ , where  $T_{\mu \rightarrow \nu}$  is the optimal transport map between a source probability measure  $\mu$  and a target measure  $\nu$ . In order to have a common space in which

to consider the optimal transport map  $T_{\mu \rightarrow \nu}$ , we will mainly consider the problem of the stability of the map  $T_\nu := T_{\mu \rightarrow \nu}$  for a fixed  $\mu$ . As first noted by Li and Nchetto [17], the arguments implying quantitative stability of  $\nu \mapsto T_{\mu \rightarrow \nu}$  sometimes also imply general stability results, where both the source and target measures can change.

To the best of our knowledge, the first quantitative stability result in optimal transport is of “local” nature, in the sense that it only holds near a configuration  $(\mu, \nu)$ , and is established under strong assumptions on the data. It is due to Ambrosio and reported in an article of Gigli [14]. It can be phrased as follows.

**Theorem 1** (Ambrosio-Gigli). *Assume that  $M$  and  $N$  are compact subsets of  $\mathbb{R}^d$ , that  $\mu \in \mathcal{P}(M)$  is absolutely continuous, and that for some  $\nu_0 \in \mathcal{P}(N)$  the optimal transport map  $T_{\mu \rightarrow \nu_0}$  for the quadratic cost  $c(x, y) = \|x - y\|^2$  is Lipschitz. Then*

$$\forall \nu_1 \in \mathcal{P}(N), \|T_{\mu \rightarrow \nu_0} - T_{\mu \rightarrow \nu_1}\|_{L^2(\mu)}^2 \leq \text{diam}(M) \text{Lip}(T_{\mu \rightarrow \nu_0}) W_1(\nu_0, \nu_1). \quad (1.1)$$

In the above statement,  $\text{Lip}(T)$  is the Lipschitz constant of the map  $T$  and  $W_1(\nu_0, \nu_1)$  is the Wasserstein distance between  $\nu_0$  and  $\nu_1$  with respect to the Euclidean distance on  $N$ .

By Brenier theorem [7], we know that  $T_\nu = \nabla \varphi_\nu$ , where  $\varphi_\nu$  is convex. A convex analysis result shows that the Lipschitz regularity of  $T_\nu$  is equivalent to the strong convexity of the convex conjugate  $\psi_\nu = \varphi_\nu^*$ . Using these remarks, the proof of the stability estimate (1.1) can then be obtained in a few lines, see e.g. [10, Theorem 2.2]. Li and Nchetto [17] prove under the same hypothesis that if  $\gamma \in \mathcal{P}(M \times N)$  is the transport plan between  $\mu$  and  $\nu$  induced by the optimal map  $T_{\mu \rightarrow \nu}$ , and  $\tilde{\gamma}$  is *any* optimal transport plan between  $\tilde{\mu}$  and  $\tilde{\nu}$ , i.e. any solution to (KP) then

$$W_2(\gamma, \tilde{\gamma})^2 \leq C(W_2(\mu, \tilde{\mu}) + W_2(\nu, \tilde{\nu})),$$

where  $C$  is a constant that depends on  $\text{Lip}(T_{\mu \rightarrow \nu})$ , the diameters of  $M$  and  $N$ . The Wasserstein distance  $W_2$  in the left-hand side is with respect to a product metric on  $M \times N$ .

We mention that the “Euclidean” stability result (1.1) can be extended to optimal transport problems on a compact Riemannian manifold with the squared geodesic distance [2]. We also mention the more “global” stability results of [4, 10], which do not make regularity assumptions on  $T_{\mu \rightarrow \nu}$ , but come with worse continuity estimates. For instance, the main theorem of [10] shows that if  $\mu \in \mathcal{P}(\mathbb{R}^d)$  is a probability density on a compact convex subset of  $\mathbb{R}^d$ , which is bounded from above and below by a positive constant, then for any compact subset  $Y \subseteq \mathbb{R}^d$ , the map  $\nu \mapsto T_{\mu \rightarrow \nu}$  is  $\frac{1}{6}$ -Hölder from  $(\mathcal{P}(Y), W_1)$  to  $L^2(\mu, \mathbb{R}^d)$ , to be compared to the  $\frac{1}{2}$  exponent in (1.1).

**Strong  $c$ -concavity of the potential.** A key ingredient in the stability results for the quadratic cost [14, 2] is the strong convexity of the Kantorovich potentials  $\psi$  associated to the optimal transport maps. In order to get stability results for general cost functions  $c$ , we introduce below the notion of *strong  $c$ -concavity*.

We denote by  $d_N : N \times N \rightarrow \mathbb{R}_+$  the distance on  $N$ . The  $p$ -Wasserstein distance on  $\mathcal{P}(N)$  between two probability measures is defined with respect to the distance by

$$W_p^p(\nu_0, \nu_1) = \inf_{\gamma \in \Gamma(\nu_0, \nu_1)} \int_{N \times N} d_N(y, z)^p d\gamma(y, z),$$

**Definition 2** (Transport map induced by a potential). *Let  $T : M \rightarrow N$  be a measurable map, and  $\psi : N \rightarrow \mathbb{R}$ . We say that  $T$  is induced by  $\psi$ , or that  $\psi$  is a potential associated to  $T$  if*

$$\forall x \in M, \quad T(x) \in \operatorname{argmin}_{y \in N} c(x, y) - \psi(y)$$

Thanks to Kantorovich duality [25], we know that if a transport map  $T$  from  $\mu$  to  $\nu$  is induced by a potential  $\psi$  then  $T$  is a solution to the Monge problem (MP). Such a potential  $\psi$  can be constructed by solving the dual problem

$$\sup_{\psi : N \rightarrow \mathbb{R}} \int_M \psi^c d\mu + \int_N \psi d\nu \quad (\text{DP})$$

where  $\psi^c : M \rightarrow \mathbb{R}$  is the  $c$ -transform of  $\psi$ , defined by

$$\psi^c(x) = \inf_{y \in N} c(x, y) - \psi(y)$$

so that  $\psi^c(x) + \psi(y) \leq c(x, y)$ . The dual problem (DP) has a maximizer, for instance, if the cost  $c$  is continuous on the compact  $M \times N$ , but existence also holds with weaker hypothesis on  $c$ , see [26] for instance. When such a maximizer exists, and still by Kantorovich theory, we can assume that a map  $T$  solution of (MP) is induced by a  $c$ -concave potential  $\psi$ . We recall the notion of  $c$ -concavity, and we refer to [26].

**Definition 3** ( $c$ -concavity and  $c$ -conjugate). *We say that  $\psi : N \rightarrow \mathbb{R} \cup \{-\infty\}$  is  $c$ -concave if for any  $y \in N$  there exists  $x \in M$  such that*

$$\forall z \in N, \quad c(x, z) - \psi(z) \geq c(x, y) - \psi(y)$$

*An equivalent definition is that there exists a function  $\varphi : M \rightarrow \mathbb{R} \cup \{\pm\infty\}$  such that for any  $y \in N$*

$$\psi(y) = \inf_{x \in M} c(x, y) - \varphi(x).$$

*We denote the right-hand side of the above equation by  $\varphi^c(x)$ , and we call it the  $c$ -conjugate of  $\varphi$ . One can define similarly the notion of  $c$ -concave function on  $M$ .*

The  $c$ -superdifferential of  $\psi$  at a point  $y \in N$  is defined by

$$\partial^c \psi(y) = \{x \in M \mid \forall z \in N, \psi(z) - c(x, z) \leq \psi(y) - c(x, y)\} \quad (1.2)$$

Note that  $\psi$  is  $c$ -concave iff for any  $y \in N$  its  $c$ -superdifferential  $\partial^c \psi(y)$  is non-empty. We can now introduce the notion of strong  $c$ -concavity.

**Definition 4** (strong  $c$ -concavity on  $D$ ). *We say that a  $c$ -concave function  $\psi$  is strongly  $c$ -concave on a set  $D \subseteq M \times N$  and with modulus  $\omega$  if for all  $x, y, z$  such that  $(x, y) \in D$ ,  $(x, z) \in D$  and  $x \in \partial^c \psi(y)$ :*

$$\psi(z) - c(x, z) \leq \psi(y) - c(x, y) - \omega(d_N(y, z)) \quad (1.3)$$

In the above definition, the modulus  $\omega : \mathbb{R}_+ \rightarrow \mathbb{R}_+$  is an increasing function that satisfies  $\omega(0) = 0$ . One can check that when  $c(x, y) = -\langle x|y \rangle$  and  $\omega(r) = Cr^2$  the notion of strong concavity and strong  $c$ -concavity are equivalent. Moreover if a function  $\psi : N \rightarrow \mathbb{R}$  is strongly  $c$ -concave, then for  $y \neq z$  in  $N$ ,  $\partial^c \psi(y) \cap \partial^c \psi(z) = \emptyset$ , or equivalently for  $x \in M$  there exists a unique minimizer of  $y \mapsto c(x, y) - \psi(y)$ . This implies that the transport map associated to  $\psi$  is uniquely defined by minimizing  $c(x, \cdot) - \psi$ :

$$\forall x \in M \quad T(x) = \operatorname{argmin}_{y \in N} c(x, y) - \psi(y)$$

**Contribution.** This paper is concerned with stability problems in optimal transport. We introduce the notion of *strong  $c$ -concavity*, which is central to get stability results.

- We provide two stability results in Section 2 that depend on an assumption of strong  $c$ -concavity. First, we extend the 1/2-Hölder stability result of Ambrosio stated in [14] to general cost function  $c$  (Theorem 5). Our result is local around transport maps associated to *strongly  $c$ -concave* potential. Second, we generalize a result of Li and Nchetto [17] that estimates the distance of a transport plan to an optimal transport map (the source and target measures being fixed) in terms of the suboptimality gap (Proposition 8). We then use this result to obtain quantitative stability of the transport plan with respect to both measures (Proposition 10), following the strategy of Li-Nchetto [17] for the quadratic cost.
- We provide in Section 3 the central result of this paper (Theorem 22), which is a differential criterion for a potential function  $\psi$  to be strongly  $c$ -concave. This result generalizes a sufficient condition for  $c$ -convexity proposed by Villani [26, Th. 12.46]. It requires that  $M, N$  are two smooth  $d$ -dimensional complete Riemannian manifolds. Similarly to Villani, we require a local condition on the derivatives of the potential  $\psi$  and a weak Ma-Trudinger-Wang condition [20]. In Section 4, we combine Theorem 22 to the stability results of Section 2 to get local stability results for optimal transport maps.
- The last two sections are dedicated to the applications of our stability results to two optimal transport problems on the sphere, with cost functions taking the value  $+\infty$ . In Section 5 we consider the reflector antenna problem, which is a non-imaging optics problem that can be written as optimal transport [27]. Section 6 is dedicated to the prescription of the Gaussian curvature measure of a convex body, originally introduced by Alexandrov [1] and rephrased as an optimal transport problem by Oliker [22].

## 2. STABILITY UNDER STRONG C-CONCAVITY

In this section we assume that  $M$  and  $N$  are Polish spaces. We provide stability results in the neighborhood of transport maps that are associated to strongly  $c$ -concave Kantorovitch potential. The stability result of Section 2.1 is with respect to variations of the target measure, whereas the result in Section 2.3 is with respect to variations of both the source and the target measures. This last result is a consequence of an error bound for a fixed

optimal transport problem given in Section 2.2. As a side note, we also remark in the last section that strong  $c$ -concavity implies Hölder regularity of transport maps.

**2.1. Stability with respect to the target measure.** The following theorem extends to general cost functions a theorem of Ambrosio [14], using a reformulation proposed in [10]. The hypothesis that the transport map  $T$  is Lipschitz (in the formulation of [10]) is replaced by the assumption that the transport map is induced by a strongly  $c$ -concave potential  $\psi$ , i.e.

$$\forall x \in M \quad T(x) \in \operatorname{argmin}_{y \in N} c(x, y) - \psi(y).$$

**Theorem 5.** *Let  $D \subseteq M \times N$  be a compact set and  $c : M \times N \rightarrow \mathbb{R} \cup \{+\infty\}$  be a cost function of class  $C^1$  on  $D$ . Let  $\mu \in \mathcal{P}(M)$  and  $\nu_0, \nu_1 \in \mathcal{P}(N)$ . We assume that there exists optimal transport maps  $T_i$  from  $\mu$  to  $\nu_i$  with associated potential  $\psi_i : N \rightarrow \mathbb{R}$  ( $i = 0, 1$ ) such that:*

- $\psi_0$  is Lipschitz on  $N$  and  $c$ -concave on  $D$ .
- $\psi_1$  is Lipschitz on  $N$  and strongly  $c$ -concave with modulus  $\omega$  on  $D$ .
- The maps  $T_i$  satisfies for any  $x \in M$ ,  $(x, T_i(x)) \in D$ .

Then,

$$\int_M \omega(d_N(T_0(x), T_1(x))) d\mu(x) \leq (\operatorname{Lip}(\psi_0) + \operatorname{Lip}(\psi_1)) W_1(\nu_0, \nu_1) \quad (2.4)$$

**Remark 6.** *The left hand side of inequality (2.4) measures the distance between transport maps  $T_0$  and  $T_1$ . To see this let us consider a simpler case where  $M$  and  $N$  are domains of  $\mathbb{R}^d$  and  $\omega(r) = r^2$  then we get*

$$\int_M \omega(d_N(T_0(x), T_1(x))) d\mu(x) = \|T_1 - T_0\|_{L^2(\mu)}^2$$

and in that case, Theorem 5 amounts to bounding the  $L^2$  norm of the distance between transport maps.

**Remark 7** (Discretization of the target measure). *Assume that we have two absolutely continuous measures  $\mu \in \mathcal{P}(M)$  and  $\nu \in \mathcal{P}(N)$  and an optimal transport map  $T$  from  $\mu$  to  $\nu$  satisfying all the hypothesis of Theorem 5. One can pick a family of points  $(y_i)_{1 \leq i \leq n}$  in the target space  $N$  and approximate the measure  $\nu$  by a discrete measure  $\nu_h$  of the form*

$$\nu_h = \sum_i \nu(V_i) \delta_{y_i}$$

where  $(V_i)_{1 \leq i \leq n}$  is a Voronoi tessellation of  $N$  around the points  $(y_i)_{1 \leq i \leq n}$  chosen in an appropriate way in the support of  $\nu$ . The parameter  $h$  is given by  $h = \max_{1 \leq i \leq n} \operatorname{diam}(V_i)$  so that  $W_1(\nu, \nu_h) \leq h$ . We can compute the optimal transport map  $T_h$  between  $\mu$  and  $\nu_h$  using semi-discrete methods such as [16]. Then, Theorem 5 implies

$$\int_M \omega(d_N(T(x), T_h(x))) d\mu(x) \leq Ch$$

where the constant  $C$  depends on the Lipschitz constants of the potentials, which can be controlled explicitly in many cases. If the modulus  $\omega(r)$  is quadratic, then the  $L^2(\mu)$  distance between  $T$  and  $T_h$  is controlled by  $h^{1/2}$ .



*Proof of Theorem 5.* We have

$$\langle \nu_1 - \nu_0 | \psi_1 - \psi_0 \rangle = \int_N \psi_1 d_N(\nu_1 - \nu_0) + \int_N \psi_0 d_N(\nu_0 - \nu_1)$$

Let  $A = \int_N \psi_1 d_N(\nu_1 - \nu_0)$  and  $B = \int_N \psi_0 d_N(\nu_0 - \nu_1)$ . Since  $T_{i\#\mu} = \nu_i$  we have

$$\begin{aligned} A &= \int_N \psi_1 d\nu_1 - \int_N \psi_1 d\nu_0 \\ &= \int_M \psi_1(T_1(x)) d\mu(x) - \int_M \psi_1(T_0(x)) d\mu(x) \end{aligned}$$

For  $x \in M$  we have  $x \in \partial^c \psi_i(T_i(x))$ . Then the strong  $c$ -concavity of  $\psi_1$  gives

$$\begin{aligned} A &= \int_M \psi_1(T_1(x)) - \psi_1(T_0(x)) d\mu(x) \\ &\geq \int_M c(x, T_1(x)) - c(x, T_0(x)) + \omega(d_N(T_0(x), T_1(x))) d\mu \end{aligned}$$

Now since  $\psi_0$  is also  $c$ -concave, we have

$$B \geq \int_M -c(x, T_1(x)) + c(x, T_0(x)) d\mu$$

Summing these two inequalities gives

$$\int_M \omega(d_N(T_0(x), T_1(x))) d\mu(x) \leq \int_N \psi_1 - \psi_0 d_N(\nu_1 - \nu_0)$$

Since  $\psi_0$  and  $\psi_1$  are Lipschitz, we have

$$\int_N \psi_1 - \psi_0 d_N(\nu_1 - \nu_0) \leq (\text{Lip}(\psi_0) + \text{Lip}(\psi_1)) W_1(\nu_0, \nu_1)$$

where the last inequality is given by Kantorovich-Rubinstein theorem.  $\square$

**2.2. Error bounds for optimal transport problems.** In this section, we generalize in Proposition 8 a stability result of Li and Nochetto [17] to general cost functions, using the notion of strong  $c$ -concavity. This result allows to bound in Corollary 9 the Wasserstein distance between the optimal transport map and any transport plan with the same marginals by the suboptimality gap of the transport plan.

**Proposition 8.** *Let  $\mu \in \mathcal{P}(M)$ ,  $\nu \in \mathcal{P}(N)$  and  $T : M \rightarrow N$  be an optimal transport map from  $\mu$  to  $\nu$ . We assume that  $T$  is induced by a strongly  $c$ -concave potential  $\psi : N \rightarrow \mathbb{R}$  with modulus  $\omega$  on a compact subset  $D$  of  $M \times N$  which contains the graph of  $T$ . Then any transport plan  $\gamma \in \Gamma(\mu, \nu)$  supported on  $D$  satisfies*

$$\int_{M \times N} \omega(d_N(T(x), y)) d\gamma(x, y) \leq \int_{M \times N} c(x, y) d\gamma(x, y) - \int_M c(x, T(x)) d\mu(x)$$

The left hand side of this equation is called the suboptimality gap of  $\gamma$ , and measures how worse the transport plan  $\gamma$  behaves compared to the optimal transport map  $T$ .



*Proof.* The strong  $c$ -concavity of  $\psi$  implies that for any  $x, y \in D$ ,

$$\psi(y) \leq \psi(T(x)) - c(x, T(x)) + c(x, y) - \omega(d_N(T(x), y)).$$

Moreover since  $T_{\#}\mu = \nu$ , we have

$$\int_N \psi(y) d\nu(y) = \int_M \psi(T(x)) d\mu(x)$$

which combined with the strong  $c$ -concavity of  $\psi$  gives

$$\begin{aligned} 0 &= \int_N \psi(y) d\nu(y) - \int_M \psi(T(x)) d\mu(x) \\ &= \int_D \psi(y) - \psi(T(x)) d\gamma(x, y) \\ &\leq \int_D c(x, y) - c(x, T(x)) - \omega(d_N(T(x), y)) d\gamma(x, y) \\ &= \int_D c(x, y) d\gamma(x, y) - \int_M c(x, T(x)) d\mu(x) - \int_D \omega(d_N(T(x), y)) d\gamma(x, y) \end{aligned}$$

Rearranging this inequality gives the desired conclusion.  $\square$

We can rephrase this proposition using the the 1-Wasserstein distance  $W_1$  in  $\mathcal{P}(M \times N)$  induced by the distance

$$d_{M \times N}((x, y), (x', y')) = d_M(x, x') + d_N(y, y').$$

**Corollary 9.** *Under the assumptions of Proposition 8, if the modulus of the Kantorovitch potential  $\psi$  is  $\omega(r) = Cr^2$ , one has*

$$W_1(\gamma, \gamma_T) \leq \frac{1}{\sqrt{C}} \left( \int_{M \times N} c(x, y) d\gamma(x, y) - \int_M c(x, T(x)) d\mu(x) \right)^{1/2}$$

where  $\gamma_T = (Id, T)_{\#}\mu$ .

*Proof.* Let  $S : M \times N \rightarrow (M \times N)^2$  defined by

$$S(x, y) = (S_1(x, y), S_2(x, y))$$

where  $S_1(x, y) = (x, T(x))$  and  $S_2(x, y) = (x, y)$ . Let  $\pi = S_{\#}\gamma \in \mathcal{P}((M \times N)^2)$ . One can check that  $\pi \in \Gamma(\gamma_T, \gamma)$ , which implies

$$\begin{aligned} W_1(\gamma_T, \gamma) &\leq \int_{(M \times N)^2} d_{M \times N}((x, y), (x', y')) d\pi(x, y, x', y') \\ &= \int_{M \times N} d_{M \times N}(S_1(x, y), S_2(x, y)) d\gamma(x, y) \\ &= \int_{M \times N} d_N(T(x), y) d\gamma(x, y). \end{aligned}$$

We use the Cauchy-Schwarz inequality in  $L^2(M \times N, \gamma)$  and Proposition 8 to get the desired result.  $\square$

**2.3. Stability with respect to both measures.** Here we apply Corollary 9 to show stability results of transport plans with respect to both the source and the target measures. Our result holds for general cost functions and is inspired by a result of Li and Nohetto [17] that holds in the quadratic case. We denote by  $d_M$  the distance on  $M$  and  $d_N$  the distance on  $N$ . We also choose for distance on the product space  $d_{M \times N}((x, y), (x', y')) = d_M(x, x') + d_N(y, y')$ . Throughout this section, we require the cost function  $c$  to be Lipschitz on the whole product space  $M \times N$ .

**Proposition 10** (Stability with respect to both measures). *Let  $\mu, \tilde{\mu} \in \mathcal{P}(M)$  and  $\nu, \tilde{\nu} \in \mathcal{P}(N)$ . Let  $c : M \times N \rightarrow \mathbb{R}$  be a cost function which is Lipschitz on  $M \times N$ . Let  $T : M \rightarrow N$  be an optimal transport map between  $\mu$  and  $\nu$ , and  $\tilde{\gamma}$  be an optimal transport plan between  $\tilde{\mu}$  and  $\tilde{\nu}$  for the cost  $c$ . We assume that  $T$  is induced by a strongly  $c$ -concave potential  $\psi : N \rightarrow \mathbb{R}$  with associated modulus  $\omega(r) = Cr^2$  on  $D = M \times N$ . Then we have*

$$W_1(\gamma_T, \tilde{\gamma}) \leq \varepsilon + \sqrt{\frac{2\text{Lip}(c)}{C}}\varepsilon, \quad \text{where } \varepsilon := W_1(\tilde{\mu}, \mu) + W_1(\nu, \tilde{\nu}).$$

The end of this section is devoted to the proof of this proposition. As in [17], we will use the gluing lemma [24, 26].

**Lemma 11** (gluing of measures). *Let  $(X_i, \mu_i)$  be probability spaces for  $i \in \{1, 2, 3\}$ , and  $\gamma_{12} \in \Gamma(\mu_1, \mu_2)$ ,  $\gamma_{23} \in \Gamma(\mu_2, \mu_3)$ . Then there exists  $\pi \in \mathcal{P}(X_1 \times X_2 \times X_3)$  such that  $\pi(\cdot, \cdot, X_3) = \gamma_{12}$  and  $\pi(X_1, \cdot, \cdot) = \gamma_{23}$ . Or equivalently*

$$p_{12\#}\pi = \gamma_{12} \quad p_{23\#}\pi = \gamma_{23}$$

where  $p_{ij}$  is the projection defined by  $p_{ij}(x_1, x_2, x_3) = (x_i, x_j)$ .

We also need the following (easy) lemma, showing that the transport cost

$$\mathcal{T}^c(\mu, \nu) := \min_{\gamma \in \Gamma(\mu, \nu)} \int c d\gamma$$

is Lipschitz with respect to perturbations of the measures when  $c$  is Lipschitz.

**Lemma 12.** *Let  $c : M \times N \rightarrow \mathbb{R}$  be a Lipschitz cost function. Let  $\mu, \tilde{\mu} \in \mathcal{P}(M)$  and  $\nu, \tilde{\nu} \in \mathcal{P}(N)$ . Then we have*

$$|\mathcal{T}^c(\mu, \nu) - \mathcal{T}^c(\tilde{\mu}, \tilde{\nu})| \leq \text{Lip}(c)(W_1(\mu, \tilde{\mu}) + W_1(\nu, \tilde{\nu})).$$

*Proof.* Kantorovich duality gives

$$\mathcal{T}^c(\mu, \nu) = \max_{\varphi \oplus \psi \leq c} \int_M \varphi d\mu + \int_N \psi d\nu.$$

Moreover, since the cost is Lipschitz, the maximum is attained in the dual problem; one can assume that the maximum is attained for two potentials  $\varphi, \psi$  satisfying  $\varphi = \psi^c$  and  $\varphi = \psi^c$ . In particular both  $\varphi$  and  $\psi$  are Lipschitz continuous with Lipschitz constant lower than  $\text{Lip}(c)$ . Kantorovich (weak) duality applied to the two measures  $\tilde{\mu}$  and  $\tilde{\nu}$  gives

$$\mathcal{T}^c(\tilde{\mu}, \tilde{\nu}) \geq \int_M \varphi d\tilde{\mu} + \int_N \psi d\tilde{\nu}.$$

We thus get

$$\mathcal{T}^c(\mu, \nu) - \mathcal{T}^c(\tilde{\mu}, \tilde{\nu}) \leq \int_M \varphi d(\mu - \tilde{\mu}) + \int_N \psi d(\nu - \tilde{\nu}) \leq \text{Lip}(c)(W_1(\mu, \nu) + W_1(\tilde{\mu}, \tilde{\nu}))$$

where the last inequality is given by Kantorovich-Rubinstein Theorem. By symmetry the same result holds when we exchange  $\mu, \nu$  and  $\tilde{\mu}, \tilde{\nu}$ .  $\square$

*Proof of Proposition 10.* Let  $\alpha \in \Gamma(\mu, \tilde{\mu})$  and  $\beta \in \Gamma(\tilde{\nu}, \nu)$  be optimal transport plans for the cost  $d_M$  and  $d_N$ . Let  $\pi \in \mathcal{P}(M^2 \times N^2)$  be a gluing of  $\alpha, \tilde{\gamma}$  and  $\beta$ , i.e.

$$p_{12\#}\pi = \alpha, \quad p_{23\#}\pi = \tilde{\gamma}, \quad p_{34\#}\pi = \beta$$

Defining  $\gamma = p_{14\#}\pi \in \Gamma(\mu, \nu)$ , we get

$$\begin{aligned} W_1(\gamma, \tilde{\gamma}) &\leq \int_{M^2 \times N^2} d_M(x, x') + d_N(y, y') d\pi(x, x', y, y') \\ &= \int_{M^2} d_M(x, x') d\alpha(x, x') + \int_{N^2} d_N(y, y') d\beta(y, y') \\ &= W_1(\tilde{\mu}, \mu) + W_1(\nu, \tilde{\nu}) \end{aligned} \quad (2.5)$$

We also have

$$\begin{aligned} &\int_{M \times N} c(x, y) d\gamma \\ &= \int_{M^2 \times N^2} c(x, y) d\pi(x, x', y', y) \\ &= \int_{M^2 \times N^2} c(x', y') + c(x, y) - c(x', y') d\pi(x, x', y', y) \\ &\leq \int_{M^2 \times N^2} c(x', y') + \text{Lip}(c)(d_M(x, x') + d_N(y, y')) d\pi(x, x', y', y) \\ &= \int_{M \times N} c(x', y') d\tilde{\gamma} + \text{Lip}(c) \left( \int_{M^2} d_M(x, x') d\alpha + \int_{N^2} d_N(y, y') d\beta \right) \\ &\leq \int_{M \times N} c(x, y) d\tilde{\gamma} + \text{Lip}(c)(W_1(\mu, \tilde{\mu}) + W_1(\nu, \tilde{\nu})) \end{aligned} \quad (2.6)$$

The transport plans  $\gamma_T = (Id, T)_\# \mu \in \Gamma(\mu, \nu)$  and  $\tilde{\gamma} \in \Gamma(\mu, \nu)$  are optimal, so that by Lemma 12,

$$\int_{M \times N} c(x, y) d\tilde{\gamma} \leq \int_{M \times N} c(x, y) d\gamma_T + \text{Lip}(c)(W_1(\mu, \tilde{\mu}) + W_1(\nu, \tilde{\nu}))$$

which combined with (2.6) gives

$$\int_{M \times N} c(x, y) d\gamma - \int_{M \times N} c(x, y) d\gamma_T \leq 2\text{Lip}(c)(W_1(\mu, \tilde{\mu}) + W_1(\nu, \tilde{\nu}))$$

Corollary 9 then implies that

$$W_1(\gamma, \gamma_T) \leq \left[ \frac{2\text{Lip}(c)}{C} (W_1(\tilde{\mu}, \mu) + W_1(\nu, \tilde{\nu})) \right]^{1/2}$$

Finally, using the triangle inequality along with (2.5) we get

$$\begin{aligned} W_1(\tilde{\gamma}, \gamma_T) &\leq W_1(\tilde{\gamma}, \gamma) + W_1(\gamma, \gamma_T) \\ &\leq W_1(\tilde{\mu}, \mu) + W_1(\nu, \tilde{\nu}) + \left( \frac{2\text{Lip}(c)}{C} (W_1(\tilde{\mu}, \mu) + W_1(\nu, \tilde{\nu})) \right)^{1/2} \\ &= \varepsilon + \sqrt{\frac{2\text{Lip}(c)}{C}} \varepsilon \end{aligned} \quad \square$$

**2.4. A remark on regularity.** The above results show that the notion of strong  $c$ -concavity is sufficient to get stability results. In fact, this notion can also lead to regularity of the associated transport maps, as expressed in the following lemma.

**Lemma 13** (Regularity under strong  $c$ -concavity). *Let us assume that the cost function  $c : M \times N \rightarrow \mathbb{R}$  is Lipschitz on  $M \times N$  and let  $T : M \rightarrow N$  be a transport map induced by a strongly  $c$ -concave potential  $\psi : N \rightarrow \mathbb{R}$ , with continuity modulus  $\omega(r) = Cr^2$  on  $M \times N$ . Then  $T$  is  $1/2$ -Hölder:*

$$d_N(T(x), T(x')) \leq \left( \frac{\text{Lip}(c)}{C} d_M(x, x') \right)^{1/2}$$

*Proof.* Let  $x \in M$ . Since  $T$  is induced by a strongly  $c$ -concave potential  $\psi$  we have  $T(x) = \operatorname{argmin}_{y \in N} c(x, y) - \psi(y)$ . The strong  $c$ -concavity of  $\psi$  implies that for every  $y \in N$

$$c(x, y) - \psi(y) \geq c(x, T(x)) - \psi(T(x)) + \omega(d_N(y, T(x)))$$

Now let  $x' \in M$ . By choosing  $y = T(x')$  the above inequality becomes

$$c(x, T(x')) - \psi(T(x')) \geq c(x, T(x)) - \psi(T(x)) + \omega(d_N(T(x'), T(x)))$$

This inequality still holds when we exchange  $x$  and  $x'$ , summing the two gives

$$2\omega(d_N(T(x), T(x'))) \leq c(x', T(x)) + c(x, T(x')) - c(x', T(x')) - c(x, T(x))$$

and since  $c$  Lipschitz we have

$$Cd_N(T(x), T(x'))^2 \leq \text{Lip}(c)d_M(x, x'). \quad \square$$

Thus, strong  $c$ -concavity of the potential entails some regularity of the transport map, generalizing what is well-known in the convex setting (i.e. if  $\psi$  is strongly convex, then  $\psi^*$  is  $\mathcal{C}^{1,1}$ ). The next section will show a partial converse statement, under strong assumptions on the cost function.

### 3. SUFFICIENT CONDITION FOR STRONG C-CONCAVITY

This section is all about the notion of strong  $c$ -concavity that we used through the previous section to deduce stability results of optimal transport maps. From now on, we assume that  $M$  and  $N$  are smooth complete Riemannian manifolds.

It is known that the notions of convexity and strong convexity can be easily characterized by conditions on the Hessian for smooth functions. The  $c$ -convexity is not that easy to study but for cost functions  $c$  that are regular enough in a certain sense, there is a differential criterion for  $c$ -convexity, given by Villani [26]. In this section we extend Villani's result for strong  $c$ -concavity, in other words we show that the strong  $c$ -concavity of a function can also be guaranteed by conditions on its derivatives. This result is presented in Corollary 23. To do so we need the cost function  $c : M \times N \rightarrow \mathbb{R} \cup \{+\infty\}$  to satisfy the Ma-Trudinger-Wang (MTW) condition, which is a well known condition in regularity theory of optimal transport.

**3.1. The Ma-Trudinger-Wang tensor.** We recall in this section the notion of MTW tensor [26]. Recall that we are working with two smooth complete Riemannian manifolds  $M$  and  $N$ , and a cost function  $c : M \times N \rightarrow \mathbb{R} \cup \{+\infty\}$ . We denote by  $\text{Dom}(\nabla_x c) \subseteq M \times N$  the set of differentiability of the cost  $c$  and  $\text{Dom}'(\nabla_x c(x, \cdot)) = \text{int}(\text{Dom}(\nabla_x c(x, \cdot)))$  the interior of the domain of definition of  $\nabla_x c(x, \cdot)$ , then

$$\text{Dom}'(\nabla_x c) = \{(x, y) \mid x \in \text{int}(M), y \in \text{Dom}'(\nabla_x c(x, \cdot))\} \quad (3.7)$$

**Definition 14** (Twisted cost). *The cost  $c$  satisfies the (Twist) condition if  $\nabla_x c(x, \cdot)$  is injective on its domain of definition, i.e. for any  $x, y, y'$  such that  $(x, y) \in \text{Dom}'(\nabla_x c)$  and  $(x, y') \in \text{Dom}'(\nabla_x c)$ :*

$$\nabla_x c(x, y) = \nabla_x c(x, y') \implies y = y'$$

**Definition 15** (STwist). *The cost satisfies the strong Twist condition (STwist) if  $c$  is  $\mathcal{C}^2$ ,  $\nabla_x c$  is one-to-one and  $D_{xy}^2 c$  is non singular on  $\text{Dom}'(\nabla_x c)$ .*

If the cost function satisfies (Twist), then for  $x \in \text{int}(M)$  the function  $-\nabla_x c(x, \cdot)$  is invertible on its image  $\mathfrak{J}_x \subseteq T_x M$ , i.e.

$$-\nabla_x c(x, \cdot) : \text{Dom}'(\nabla_x c(x, \cdot)) \subseteq N \rightarrow \mathfrak{J}_x \subseteq T_x M$$

is one-to-one.

**Definition 16** (c-exponential). *When the cost  $c$  satisfies the (Twist) condition, we can define the c-exponential for  $x \in M$  by  $\text{c-exp}_x = (-\nabla_x c(x, \cdot))^{-1}$ , giving for  $p \in \mathfrak{J}_x$ :*

$$\begin{aligned} \text{c-exp}_x(p) : \mathfrak{J}_x \subseteq T_x M &\rightarrow \text{Dom}'(\nabla_x c(x, \cdot)) \subseteq N \\ p &\rightarrow (\nabla_x c(x, \cdot))^{-1}(-p) \end{aligned}$$

**Definition 17** (c-segment). *A c-segment is the image of a usual segment by the map  $\text{c-exp}_x$ . We denote  $(y_t)_{0 \leq t \leq 1} = [y_0, y_1]_x$  the c-segment between  $y_0$  and  $y_1$  with base  $x$  defined for  $p_0 = (-\nabla_x c)^{-1}(x, y_0)$  and  $p_1 = (-\nabla_x c)^{-1}(x, y_1)$  by*

$$y_t = \text{c-exp}_x((1-t)p_0 + tp_1)$$

**Definition 18** (c-convex set). *Let  $A \subseteq N$ .*

- *We say that  $A$  is c-convex with respect to  $x \in M$  if for any  $y_0, y_1 \in A$ , there is a c-segment  $[y_0, y_1]_x$  entirely contained in  $A$ .*
- *The set  $A$  is said to be c-convex with respect to a set  $B \subseteq M$  if  $A$  is c-convex with respect to any  $x \in B$ .*
- *A set  $D \subseteq M \times N$  is said to be totally c-convex if for any two points  $(x, y_0) \in D$  and  $(x, y_1) \in D$ , the c-segment  $(y_t)_{0 \leq t \leq 1} = [y_0, y_1]_x$  satisfies for any  $t$   $(x, y_t) \in D$ .*
- *We say that  $D \subseteq M \times N$  is symmetrically c-convex if both  $[x_0, x_1]_y \subseteq D$  and  $[y_0, y_1]_x \subseteq D$ .*

**Definition 19** (MTW tensor). *Assuming that  $c$  is of class  $\mathcal{C}^4$  on  $\text{Dom}'(\nabla_x c)$  and satisfies the (STwist) condition, the Ma-Trudinger-Wang tensor is defined for  $(x_0, y_0) \in \text{Dom}'(\nabla_x c)$  and  $(\zeta, \eta) \in T_x M \times T_y N$  by*

$$\mathfrak{S}_c(x_0, y_0)(\eta, \zeta) = -\frac{3}{2} \frac{\partial^2}{\partial q_\eta^2} \frac{\partial^2}{\partial y_\zeta^2} (c(\text{c-exp}_{y_0}(q), y)) \Big|_{y=y_0, q=-\nabla_y c(x_0, y_0)}$$

with  $\tilde{\eta} = -\nabla_{xy}^2 c(x_0, y_0)\eta \in T_x M$ .

In the above definition  $-\nabla_{xy}^2 c(x_0, y_0) : T_x M \times T_y N \rightarrow \mathbb{R}$  is a bilinear form which is non singular since (STwist) is satisfied. We then identify for  $\eta \in T_y N$  the linear form  $-\nabla_{xy}^2 c(x_0, y_0)\eta = \tilde{\eta} : T_x M \rightarrow \mathbb{R}$  with a vector of  $T_x M$  using the Riemannian structure.

**Definition 20** (weak MTW). *We say that the weak MTW condition (MTWw) is satisfied on a compact set  $D \subseteq M \times N$  if there exists a constant  $C > 0$  such that for any  $(x, y) \in D$  and  $(\zeta, \eta) \in T_x M \times T_y N$  we have*

$$\mathfrak{S}_c(x, y)(\eta, \zeta) \geq -C|\langle \zeta | \tilde{\eta} \rangle| \|\zeta\| \|\tilde{\eta}\| \quad (\text{MTWw})$$

*This condition was introduced by Ma, Trudinger and Wang [20] and is often referred to as (A3w).*

**3.2. Differential criterion for strong c-concavity.** The goal here is to generalize Villani's differential criterion [26] (detailed in the following theorem) for c-convexity to our definition of strong c-concavity. Our proof is highly inspired from Villani's one, in particular we study the same real valued function  $h : [0, 1] \rightarrow \mathbb{R}$  and show inequalities that are similar and also require positivity of the MTW tensor.

**Theorem 21** (Differential criterion for c-convexity, [26, Th. 12.46]). *Let  $D \subseteq M \times N$  be a closed symmetrically c-convex set and  $c \in \mathcal{C}^4(D, \mathbb{R})$  such that  $c$  and  $\check{c}$  satisfy (STwist) on  $D$ . Assume that the weak MTW condition (MTWw) is satisfied on  $D$ . Let  $\mathcal{X} = \text{proj}_M(D)$  and  $\psi \in \mathcal{C}^2(\mathcal{X}, \mathbb{R})$ . If for any  $x \in \mathcal{X}$  there exists  $y \in N$  such that  $(x, y) \in D$  and*

$$\begin{cases} \nabla \psi(x) + \nabla_x c(x, y) = 0 \\ D^2 \psi(x) + D_{xx}^2 c(x, y) \geq 0 \end{cases}$$

*Then  $\psi$  is c-convex on  $D$ .*

This theorem is given for a potential function  $\psi$  on  $\mathcal{X} \subseteq M$  and gives a c-convexity result while we consider  $\psi : N \rightarrow \mathbb{R}$  and work on c-concavity, but this is really just a matter of convention. Also Villani needs the Hessian  $D^2 \psi(x) + D_{xx}^2 c(x, y)$  to be positive semi-definite to obtain c-convexity, while we are naturally going to need the Hessian  $D_{yy}^2 c(x, y) - D^2 \psi(y)$  to have eigenvalues bounded from below by a positive constant to obtain strong c-concavity. A noticeable difference of c-convexity with respect to convexity is that it cannot be expressed locally, as we require the MTW tensor to be positive on the whole set  $D$  which is a global condition.

**Theorem 22** (Differential criterion for strong c-concavity). *We consider  $D \subseteq \text{Dom}'(\nabla_x c) \cap \text{Dom}'(\nabla_y c)$  a symmetrically c-convex compact set and denote  $\mathcal{X} = \text{proj}_M(D)$ ,  $\mathcal{Y} = \text{proj}_N(D)$ . We assume that  $c \in \mathcal{C}^4(D, \mathbb{R})$ , that  $c$  and  $\check{c}$  satisfy (STwist) on  $D$  where  $\check{c}(x, y) = \check{c}(y, x)$ . We also assume that the weak MTW condition is satisfied on  $D$ . Let  $\psi \in \mathcal{C}^2(\mathcal{Y}, \mathbb{R})$  be a c-concave function on  $D$  and such that there exists  $\lambda > 0$  satisfying for any  $x \in \partial^c \psi(y)$*

$$D^2 \psi(y) - D_{yy}^2 c(x, y) \geq \lambda Id$$

*Then  $\psi$  is strongly c-concave on  $D$  with modulus  $\omega(d_N(y, z)) = C d_N(y, z)^2$ , where  $C > 0$  is a constant depending on  $c$ ,  $\mathcal{X}$  and  $\mathcal{Y}$ . This means that we*

have

$$\psi(z) - c(x, z) \geq \psi(y) - c(x, y) + Cd_N(y, z)^2$$

for the points  $x \in \mathcal{X}$ ,  $y, z \in \mathcal{Y}$  such that  $x \in \partial^c \psi(y)$ ,  $(x, y) \in D$  and  $(x, z) \in D$ .

**Corollary 23** (Strong c-concavity). *We make the same hypothesis on  $c$  and  $D$ , and just assume  $\psi \in C^2(\mathcal{Y}, \mathbb{R})$ . Let  $T : \mathcal{X} \rightarrow \mathcal{Y}$  the map defined by  $T(x) = \operatorname{argmin}_y c(x, y) - \psi(y)$  be of class  $C^1$  and satisfying for any  $x \in \mathcal{X}$ ,  $(x, T(x)) \in D$ . Then the function  $\psi$  is strongly c-concave on the set  $D$  with modulus  $\omega(d_N(y, z)) = Cd_N(y, z)^2$ .*

**Remark 24** (Restriction of c-concavity to  $D$ ). *Theorem 22 actually gives the strong c-concavity of the potential  $\psi$  on a set  $D$  where the cost function is smooth enough. This can be an issue if we want to find a transport map  $T : M \rightarrow N$  such that  $T(x) = \operatorname{argmin}_{z \in N} c(x, z) - \psi(z)$ , because we cannot be sure that the argmin will be obtained at a point  $y$  such that  $(x, y) \in D$ . This issue has to be treated independently for each application.*

**3.3. Proof of Theorem 22.** In this subsection we consider that all the hypothesis of Theorem 22 are satisfied. For any  $x \in \mathcal{X}$ , we denote  $\mathcal{Y}^x = \{y \in N \mid (x, y) \in D\}$ . Let  $\bar{y} \in \mathcal{Y}$  and  $\bar{x} \in \partial^c \psi(\bar{y})$  such that  $(\bar{x}, \bar{y}) \in D$ . Note that  $\bar{x}$  always exists by hypothesis. Let us fix  $y \in \mathcal{Y}^{\bar{x}}$ . We want to show that there exists a constant  $C > 0$  independant of  $\bar{x}, \bar{y}$  and  $y$  such that

$$c(\bar{x}, y) - \psi(y) \geq c(\bar{x}, \bar{y}) - \psi(\bar{y}) + Cd_N(y, \bar{y})^2 \quad (3.8)$$

We put  $(y_t)_{0 \leq t \leq 1} = [\bar{y}, y]_{\bar{x}}$  the  $c$ -segment between  $\bar{y}$  and  $y$  with base  $\bar{x}$ . Remark that the  $c$ -convexity of  $D$  implies that for any  $t$  in  $[0, 1]$ ,  $(\bar{x}, y_t) \in D$ . We define the function  $h$  by

$$h(t) := c(\bar{x}, y_t) - \psi(y_t)$$

such that Equation (3.8) writes

$$h(1) \geq h(0) + Cd_N(y, \bar{y})^2 \quad (3.9)$$

The end of this section is devoted to the proof of Equation (3.9).

**Notation.** We first introduce some notations. Note that  $A_{\bar{x}} := \nabla_{xy}^2 c(\bar{x}, y_t) : T_{\bar{x}}M \times T_{y_t}N \rightarrow \mathbb{R}$  is a bilinear form which is assumed to be nonsingular. For any  $X \in T_{\bar{x}}M$  and  $Y \in T_{y_t}N$ , we can write  $\nabla_{xy}^2 c(\bar{x}, y_t)(X, Y) = \langle A_{\bar{x}}X | Y \rangle = \langle {}^t A_{\bar{x}}Y | X \rangle$  where in some local coordinates  $A_{\bar{x}}$  is an invertible matrix and  $X$  and  $Y$  are column matrices. Then  $\nabla_{xy}^2 c(\bar{x}, y_t)(X, \cdot)$  is a linear form on  $T_{y_t}N$  which is identified to the vector  $A_{\bar{x}}X \in T_{y_t}N$ . Similarly  ${}^t A_{\bar{x}}Y \in T_{\bar{x}}M$ . We take the same notation for  $A_{x_s} = \nabla_{xy}^2 c(x_s, y_t)$ .

**Lemma 25.**

$$h'(t) = \langle \zeta | \hat{\eta} \rangle$$

and

$$h''(t) = \left( D_{yy}^2 c(x^t, y_t) - D^2 \psi(y_t) \right) (\hat{\eta}, \hat{\eta}) + \frac{2}{3} \int_0^1 \mathfrak{S}_c \left( c\text{-exp}_{y_t}(\bar{y}_t + s\zeta), y_t \right) (\bar{\zeta}, \hat{\eta})(1-s) ds,$$



where  $x^t \in \partial^c \psi(y_t)$ ,

$$\begin{aligned} \eta &= \nabla_x c(\bar{x}, \bar{y}) - \nabla_x c(\bar{x}, y_t) \in T_{\bar{x}}M & \hat{\eta} &= -{}^t A_{\bar{x}}^{-1} \eta \in T_{y_t}N \\ \zeta &= \nabla_y c(\bar{x}, y_t) - \nabla \psi(y_t) \in T_{y_t}N & \hat{\zeta} &= -A_{\bar{x}}^{-1} \zeta \in T_{\bar{x}}M \\ \bar{q}_t &:= -\nabla_y c(\bar{x}, y_t) \in T_{y_t}M & \bar{\zeta} &= -A_{x_s}^{-1} \zeta \in T_{x_s}N \end{aligned}$$

*Proof of Lemma 25.* Since  $D$  is symmetrically  $c$ -convex and  $(\bar{x}, \bar{y}) \in D$ ,  $(\bar{x}, y) \in D$ , we can differentiate  $h$  as follows

$$h'(t) = \langle \nabla_y c(\bar{x}, y_t) - \nabla \psi(y_t) | \dot{y}_t \rangle$$

We also have by differentiating  $-\nabla_x c(\bar{x}, y_t) = \bar{p} + t\eta$ :

$$\eta = -\nabla_{xy}^2 c(\bar{x}, y_t) \dot{y}_t = -{}^t A_{\bar{x}} \dot{y}_t$$

So that  $\hat{\eta} = -{}^t A_{\bar{x}}^{-1} \eta = \dot{y}_t$  and thus

$$h'(t) = \langle \zeta | \hat{\eta} \rangle.$$

Differentiating  $h'$  gives

$$h''(t) = \left( \nabla_{yy}^2 c(\bar{x}, y_t) - \nabla^2 \psi(y_t) \right) (\dot{y}_t, \dot{y}_t) + \langle \zeta | \ddot{y}_t \rangle.$$

By differentiating  $-\eta = \nabla_{xy}^2 c(\bar{x}, y_t) \dot{y}_t$ , one gets

$$\nabla_{xyy} c(\bar{x}, y_t) (\dot{y}_t, \dot{y}_t) + \langle \nabla_{xy} c(\bar{x}, y_t) | \ddot{y}_t \rangle = 0$$

so that

$$\ddot{y}_t = -{}^t A_{\bar{x}}^{-1} \nabla_{xyy} c(\bar{x}, y_t) (\hat{\eta}, \hat{\eta})$$

and

$$\langle \zeta | \ddot{y}_t \rangle = \langle \zeta | -{}^t A_{\bar{x}}^{-1} \nabla_{xyy} c(\bar{x}, y_t) (\hat{\eta}, \hat{\eta}) \rangle = \langle -A_{\bar{x}}^{-1} \zeta | \nabla_{xyy} c(\bar{x}, y_t) (\hat{\eta}, \hat{\eta}) \rangle.$$

We therefore have

$$h''(t) = \left( \nabla_{yy}^2 c(\bar{x}, y_t) - \nabla^2 \psi(y_t) \right) (\hat{\eta}, \hat{\eta}) + \langle \hat{\zeta} | \nabla_{xyy} c(\bar{x}, y_t) (\hat{\eta}, \hat{\eta}) \rangle$$

We define  $\Phi(x) := \left( \nabla_{yy}^2 c(x, y_t) - \nabla^2 \psi(y_t) \right) (\hat{\eta}, \hat{\eta})$ . Then we have for  $X \in T_x M$

$$D\Phi(x).X = \langle X | \nabla_{xyy}^3 c(x, y_t) (\hat{\eta}, \hat{\eta}) \rangle,$$

so that

$$h''(t) = \Phi(\bar{x}) + D\Phi(\bar{x}) \hat{\zeta}$$

We put  $\tilde{\Phi}(q) = \tilde{\Phi}(-\nabla_y c(x, y_t)) = \Phi(x)$ , so that for  $X \in T_x M$

$$D\Phi(\bar{x})X = D\tilde{\Phi}(\bar{q}_t)(-A_{\bar{x}}X)$$

For  $x = \bar{x}$  and  $\eta \in T_{y_t}N$

$$D\Phi(\bar{x})\hat{\eta} = D\tilde{\Phi}(\bar{q}_t)(-A_{\bar{x}}\hat{\eta}) = D\tilde{\Phi}(\bar{q}_t)\eta$$

We put  $q_t := -\nabla_y c(x^t, y_t)$  with  $x^t \in \partial^c \psi(y_t)$  and recall  $\bar{q}_t = -\nabla_y c(\bar{x}, y_t)$ . We get  $\nabla \psi(y_t) = \nabla_y c(x^t, y_t) = -q_t$  and therefore get  $\zeta = q_t - \bar{q}_t$ . Using the  $c$ -convexity of  $D$  to differentiate  $c$  at  $(x_t, c\text{-exp}_{x_t}(\bar{p}_t + s\zeta))$ , we get

$$h''(t) = \tilde{\Phi}(\bar{q}_t) + D\tilde{\Phi}(\bar{q}_t)(q_t - \bar{q}_t) = \tilde{\Phi}(q_t) - \int_0^1 D_q^2 \tilde{\Phi}(\bar{q}_t + s\zeta)(\zeta, \zeta)(1-s) ds$$

We have

$$\tilde{\Phi}(q_t) = \Phi(x^t) = \left( \nabla_{yy}^2 c(x^t, y_t) - \nabla^2 \psi(y_t) \right) (\hat{\eta}, \hat{\eta})$$

Using the change of variable  $q = -\nabla_y c(x, y_t) \in T_{y_t}M$  (or equivalently  $x = \text{c-exp}_{y_t}(q)$ ), we get

$$\tilde{\Phi}(q) = \left( \nabla_{yy}^2 c(\text{c-exp}_{y_t}(q), y_t) - \nabla^2 \psi(y_t) \right) (\hat{\eta}, \hat{\eta})$$

Since  $\nabla^2 \psi(y_t)$  does not depend on  $q$ , one gets by the definition of the MTW tensor, for any  $\zeta \in T_q(T_{y_t}N) = T_{y_t}N$ :

$$D_q^2 \tilde{\Phi}(q)(\zeta, \zeta) = \frac{\partial^2}{\partial q_\zeta^2} \frac{\partial^2}{\partial y_{\hat{\eta}}^2} (c(\text{c-exp}_{y_t}(q), y_t)) = -\frac{2}{3} \mathfrak{S}_c(\text{c-exp}_{y_t}(q), y_t) (\bar{\zeta}, \hat{\eta})$$

where we put  $\bar{\zeta} := -A_{x_s}^{-1} \zeta$  so as to have  $\tilde{\zeta} = \zeta$ .  $\square$

**Lemma 26.** *Let  $y \in C^1([0, 1], \mathbb{R})$  satisfying for  $C > 0$ ,*

$$\begin{cases} y'(t) \geq -C|y(t)| \\ y(0) = 0 \end{cases}$$

*then  $y(t) \geq 0$  for any  $t \in [0, 1]$ .*

*Proof.* First remark that there exists  $g \in C^0([0, 1], \mathbb{R}_+)$  such that  $y$  is solution of

$$\begin{cases} y'(t) = -C|y(t)| + g(t) \\ y(0) = 0 \end{cases}$$

Assume that  $y \leq 0$  on  $[0, 1]$  then  $|y| = -y$  and  $y$  satisfies

$$\begin{cases} y'(t) = Cy(t) + g(t) \\ y(0) = 0 \end{cases}$$

yet the unique solution to this system is  $t \mapsto \int_0^t g(s)e^{C(t-s)} ds \geq 0$  which gives  $y = 0$ . Now assume that there exists  $t_0$  such that  $y(t_0) := y_0 > 0$ , then the system may be rewritten

$$\begin{cases} y'(t) = -C|y(t)| + g(t) \\ y(t_0) = y_0 \end{cases}$$

and has for unique solution  $t \mapsto y_0 e^{C(t_0-t)} + \int_{t_0}^t g(s)e^{C(s-t)} ds \geq 0$  on  $[t_0, 1]$ . To conclude let us consider  $t_* = \inf\{t, y(t) > 0\}$ , then we have  $y(t) \leq 0$  on  $[0, t_*]$  which implies  $y = 0$  on  $[0, t_*]$  as we have seen previously.  $\square$

**Proposition 27.** *Under hypothesis of Theorem 22,*

$$h''(t) \geq -Ch'(t) + \lambda \|\hat{\eta}\|^2$$

*Proof.* We have

$$|h'(t)| = |\langle \zeta | \hat{\eta} \rangle|.$$

We also have

$$h''(t) = \left( D_{yy}^2 c(x^t, y_t) - D^2 \psi(y_t) \right) (\hat{\eta}, \hat{\eta}) + \frac{2}{3} \int_0^1 \mathfrak{S}_c(x_s, y_t) (\bar{\zeta}, \hat{\eta}) (1-s) ds,$$

where  $x_s = \text{c-exp}_{y_t}(\bar{q}_t + s\zeta)$ . By hypothesis we have

$$\left( D_{yy}^2 c(x^t, y_t) - D^2 \psi(y_t) \right) (\hat{\eta}, \hat{\eta}) \geq \lambda \|\hat{\eta}\|^2$$

and (MTWw) gives

$$\mathfrak{S}_c(x_s, y_t)(\bar{\zeta}, \hat{\eta}) \geq -C|\langle \nabla_{xy}^2 c(x_s, y_t) \hat{\eta} | \bar{\zeta} \rangle| \|\hat{\eta}\| \|\bar{\zeta}\|$$

The norms  $\|\hat{\eta}\|$  and  $\|\bar{\zeta}\|$  can be integrated in the constant by compactness, so we get

$$\mathfrak{S}_c(x_s, y_t)(\bar{\zeta}, \hat{\eta}) \geq -C|\langle \nabla_{xy}^2 c(x_s, y_t) \hat{\eta} | \bar{\zeta} \rangle| = -C|\langle {}^t A_{x_s} \hat{\eta} | \bar{\zeta} \rangle|$$

Recall that  $\bar{\zeta} = -A_{x_s}^{-1} \zeta$ . Therefore we get

$$|\langle {}^t A_{x_s} \hat{\eta} | \bar{\zeta} \rangle| = |\langle {}^t A_{x_s} \hat{\eta} | A_{x_s}^{-1} \zeta \rangle| = |\langle \zeta | \hat{\eta} \rangle| = |h'(t)|,$$

We thus have  $h''(t) \geq -C|h'(t)| + \lambda \|\hat{\eta}\|^2$ . Note that  $\zeta|_{t=0} = 0$  so  $h'(0) = 0$ . Then we can apply Lemma 26 to  $h'$ , which gives  $h'(t) \geq 0$ , so we can drop the absolute value and we obtain  $h''(t) \geq -Ch'(t) + \lambda \|\hat{\eta}\|^2$ .  $\square$

*Proof of Theorem 22.* By compactness we have

$$C_1 := \inf_{(x,y) \in D, u \in T_x M, \|u\|=1} \|\nabla_{xy}^2 c(x, y)^{-1} u\|^2 > 0$$

and

$$C_2 := \inf_{x \in \mathcal{X}, y, z \in \mathcal{Y}^x} \frac{\|\nabla_x c(x, y) - \nabla_x c(x, z)\|^2}{d_N(y, z)^2} > 0$$

such that  $\|\hat{\eta}\|^2 \geq C_1 C_2 d_N(y, \bar{y})^2$ . By Proposition 27, we get

$$h''(t) \geq -Ch'(t) + \lambda C_1 C_2 d_N(y, \bar{y})^2$$

Using Grönwall's Lemma we then have that  $h'(t) \geq g(t)$  with  $g$  solution of

$$\begin{cases} g'(t) = -Cg(t) + \lambda C_1 C_2 d_N(y, \bar{y})^2 \\ g(0) = 0 \end{cases}$$

which immediatly gives  $g(t) = \left(\frac{\lambda C_1 C_2}{C} d_N(y, \bar{y})^2\right) (1 - e^{-Ct})$ , so finally we have for  $t \in [0, 1]$ ,  $h'(t) \geq \left(\frac{\lambda C_1 C_2}{C} d_N(y, \bar{y})^2\right) (1 - e^{-Ct})$ , and the by integrating for  $t \in [0, 1]$  we conclude that

$$\int_0^1 h'(t) dt \geq \lambda C_1 C_2 e^{-C} d_N(y, \bar{y})^2$$

which is exactly what we wanted in Equation (3.9).  $\square$

*Proof of Corollary 23.* We want to show that under the hypothesis of Corollary 23, we have

$$\forall y \in \mathcal{Y} \forall x \in \partial^c \psi(y), D_{yy}^2 c(x, y) - D^2 \psi(y) \geq \lambda Id,$$

We recall that  $T : \mathcal{X} \rightarrow \mathcal{Y}$  is of class  $\mathcal{C}^1$ . Let  $x \in \mathcal{X}$ , we first assume that  $T(x) \in \text{int}(\mathcal{Y})$ . Since  $T(x)$  minimizes  $c(x, \cdot) - \psi(\cdot)$  we have

$$\nabla_y c(x, T(x)) - \nabla \psi(T(x)) = 0 \quad (3.10)$$

and

$$D_{yy}^2 c(x, T(x)) - D^2 \psi(T(x)) \geq 0 \quad (3.11)$$

By differentiating (3.10) with respect to  $x$ , we get

$$(D_{yy}^2 c(x, T(x)) - D^2 \psi(T(x))) \circ DT(x) = -D_{xy}^2 c(x, T(x)). \quad (3.12)$$

By (STwist) assumption,  $D_{xy}^2 c(x, T(x))$  is nonsingular, which implies that  $D_{yy}^2 c(x, T(x)) - D^2 \psi(T(x))$  is also nonsingular. Since we also know that it is positive semi-definite from (3.11) we get that

$$D_{yy}^2 c(x, T(x)) - D^2 \psi(T(x)) > 0.$$

We now need to extend this inequality for any  $T(x) \in \partial \mathcal{Y}$ , including the boundary. By continuity, since  $\psi$  is  $\mathcal{C}^2$  on  $\mathcal{Y}$ ,  $c$  is  $\mathcal{C}^2$  on  $D$  and  $T$  is  $\mathcal{C}^1$  on  $\mathcal{X}$ , Equations (3.11) and (3.12) still hold when  $T(x) \in \partial \mathcal{Y}$ . Moreover (STwist) being satisfied on  $D$ , we have  $D_{yy}^2 c(x, T(x)) - D^2 \psi(T(x)) > 0$  for any  $x \in \mathcal{X}$ . By compactness of  $\mathcal{X}$ , there exists  $\lambda > 0$  such that

$$\forall x \in \mathcal{X} \quad D_{yy}^2 c(x, T(x)) - D^2 \psi(T(x)) \geq \lambda Id.$$

We conclude using that  $T(x) = y$  is equivalent to  $x \in \partial^c \psi(y)$ .  $\square$

#### 4. STABILITY OF OPTIMAL TRANSPORT MAP FOR MTW COST

In this section, we show that the stability results of Section 2 can be applied to optimal transport maps. We consider two compact Riemannian manifolds  $M$  and  $N$  in  $\mathbb{R}^d$  and still denote by  $d_N$  the distance on  $N$ .

**Theorem 28** (Stability in optimal transport). *Let  $\mu \in \mathcal{P}(M)$  and  $\nu \in \mathcal{P}(N)$  be two probability measures. Let  $c : M \times N \rightarrow \mathbb{R}$  be a cost function of class  $\mathcal{C}^4$  that satisfies (STwist) and (MTWw) hypothesis. Let  $T : M \rightarrow N$  be an optimal transport map between  $\mu$  and  $\nu$  of class  $\mathcal{C}^1$  for the cost  $c$  and assume that its associated Kantorovich potential  $\psi : M \rightarrow \mathbb{R}$  is of class  $\mathcal{C}^2$ .*

- Let  $\tilde{\nu} \in \mathcal{P}(N)$  be any probability measure, and  $S : M \rightarrow N$  be an optimal transport map between  $\mu$  and  $\tilde{\nu}$ . Then we have

$$\|d_N(T, S)\|_{L^2(\mu)}^2 \leq CW_1(\nu, \tilde{\nu})$$

where  $W_1$  denotes the 1-Wasserstein distance and  $C$  is a constant depending on the cost  $c$ ,  $M$  and  $N$ .

- Let  $\tilde{\mu} \in \mathcal{P}(M)$ ,  $\tilde{\nu} \in \mathcal{P}(N)$  and  $\tilde{\gamma}$  be an optimal transport plan between  $\tilde{\mu}$  and  $\tilde{\nu}$ . Then we have

$$W_1(\tilde{\gamma}, \gamma_T) \leq C (W_1(\tilde{\mu}, \mu) + W_1(\nu, \tilde{\nu}))^{1/2}$$

where  $\gamma_T = (Id, T)_{\#} \mu$  and  $C$  is a constant depending on the cost  $c$ ,  $M$  and  $N$ .

*Proof.* Since  $M$  and  $N$  are compact we have strong duality with a cost  $c$  that is Lipschitz on  $M \times N$  so  $S$  is induced by a Lipschitz potential. Since  $T \in \mathcal{C}^1$ ,  $\psi \in \mathcal{C}^2$  and  $c \in \mathcal{C}^4$  satisfies (STwist) and (MTWw), we can then apply Corollary 23 to  $\psi$ , which gives that it is strongly  $c$ -concave on  $N$ , with modulus  $\omega(d_N(y, z)) = Cd_N(y, z)^2$ . Then the first result is given by Theorem 5 and the second is given by Proposition 10.  $\square$

For simplicity, the above theorem is stated in a restrictive way as it requires  $c$  to be smooth on the whole product space  $M \times N$ . It may happen that the regularity conditions such as (STwist) and (MTWw) are not satisfied on the whole product space  $M \times N$ , but only on a subset  $D \subseteq M \times N$ . In this case we can still obtain stability with respect to the target measure if we can show that optimal transports plans are supported on this subset  $D$ . This is treated independently on examples of Sections 5 and 6.

## 5. STABILITY FOR THE REFLECTOR COST ON THE SPHERE

In this section, we apply a stability result of Section 2 to the reflector antenna problem. It is known that this problem amounts to solving an optimal transport problem on the unit sphere  $M = N = \mathcal{S}^{d-1}$  for the cost function  $c(x, y) = -\ln(1 - \langle x|y \rangle)$  [27], extended by  $+\infty$  on the diagonal  $\{x = y\}$ . One of the key element in the proof is to show that optimal transport maps are supported on compact sets that avoid the diagonal

$$D_\varepsilon = \{(x, y) \in M^2 \mid d_M(x, y) \geq \varepsilon\} \quad (5.13)$$

where  $d_M$  is the geodesic distance on  $M$ . We first need the following definition.

**Definition 29.** *Given a probability measures  $\mu \in \mathcal{P}(M)$ , we put*

$$M_\mu(r) = \sup_{x \in M} \mu(B(x, r)).$$

**Theorem 30.** *Let  $c(x, y) = -\ln(1 - \langle x|y \rangle)$  be the reflector cost on the sphere  $M = \mathcal{S}^{d-1}$ . Let  $\mu, \nu_0, \nu_1 \in \mathcal{P}(M)$  be such that  $\mu$  and  $\nu_0$  are absolutely continuous with respect to the Lebesgue measure with strictly positive  $\mathcal{C}^{1,1}$  densities. Let  $T_i$  be optimal transport maps between  $\mu$  and  $\nu_i$ . Then for all  $\beta > 0$ , there exists a constant  $C > 0$  depending on  $\mu, \nu_0$  and  $\beta$  such that*

$$\forall \nu_1 \in \mathcal{P}(N) \text{ s.t. } M_{\nu_1}(\beta) < 1/8, \quad \|d_M(T_0, T_1)\|_{L^2(\mu)}^2 \leq C W_1(\nu_0, \nu_1)$$

where  $d_M$  is the geodesic distance on  $M$ .

The main difficulty to prove the previous theorem is to show that the optimal transport plan is supported on the compact set  $D_\varepsilon$  for some  $\varepsilon$ . This is done in the following subsection in a more general setting.

**5.1. Support of the optimal transport plan.** In this subsection, we show that optimal transport plans are supported on compact sets of the form  $D_\varepsilon$ . Since our result holds in a slightly more general context than the sphere, we consider that  $M$  can be any smooth complete Riemannian manifold. Let  $c : M \times M \rightarrow \mathbb{R}$  be any cost bounded from below that satisfies  $c(x, y) = h(d_M(x, y))$  where  $h : \mathbb{R}_+ \rightarrow \mathbb{R}$  is a continuous decreasing function such that  $h(0) = +\infty$  and  $h(t) < +\infty$  for  $t > 0$ .

**Theorem 31.** *Let  $\mu, \nu \in \mathcal{P}(M)$  and  $\beta > 0$  such that both  $M_\mu(\beta) < 1/8$  and  $M_\nu(\beta) < 1/8$ , then there exists a constant  $\varepsilon > 0$  such that any optimal transport plan  $\gamma \in \Gamma(\mu, \nu)$  is concentrated on  $D_\varepsilon$ .*

Similar results have already been obtained in different settings [13, 8, 19], but none of them can be applied to discrete measures and therefore does not imply our result. W. Gangbo and V.Oliker [13] work with Borel measures that vanish on  $(d-1)$ -rectifiable sets. G. Buttazzo et al. [8] consider multi-marginal optimal transport problems for constant measures. G.Loeper [19] considers two measures  $\mu$  and  $\nu$  such that  $\mu \geq m dVol$  with  $m > 0$  and  $\nu$  satisfies for any  $\varepsilon \geq 0$  and  $x \in M$ ,  $\nu(B(x, \varepsilon)) \leq f(\varepsilon)\varepsilon^{n(1-1/n)}$  for some function  $f : \mathbb{R}_+ \rightarrow \mathbb{R}_+$  satisfying  $\lim_{t \rightarrow 0} f(t) = 0$ . These hypothesis imply that neither  $\mu$  nor  $\nu$  can be discrete.

Our proof is an adaptation of their proofs in a different context. Lemma 36 is inspired by [13] while Lemma 35 and the overall strategy of the proof come

from [8]. The main difference is that here we work on any measure satisfying  $M(\beta) < 1/8$ , including discrete measures, which is useful for semi-discrete optimal transport.

**Remark 32.** *Our proof requires  $M(\beta) < 1/8$  but we believe that the theoretical bound is  $M(\beta) \leq 1/2$ , which is enough to guarantee that there exists a transport plan with finite global cost, as showed in the following lemma. It is easy to show that we cannot expect a greater bound. Take for example  $x \neq y$  in  $\mathcal{S}^{d-1}$ ,  $\varepsilon \in ]0, 1/2[$ ,  $\mu = 1/2(\delta_x + \delta_y)$  and  $\nu = (1/2 + \varepsilon)\delta_x + (1/2 - \varepsilon)\delta_y$ . Any transport plan between  $\mu$  and  $\nu$  will send a set of measure at least  $\varepsilon$  from  $x$  to itself for which the cost is infinite.*

The end of this section is mainly dedicated to the proof of Theorem 31, which is necessary to guarantee that the optimal transport plan is supported where the cost is regular enough and the MTW tensor is non-negative, and allow us to apply our strong c-concavity result in order to obtain the stability results of Theorem 30. We first show in the following lemma that there exists a transport plan with bounded total cost.

**Lemma 33.** *If  $M_\mu(\beta) \leq 1/2$  and  $M_\nu(\beta) \leq 1/2$  for some  $\beta > 0$ , then there exists  $\gamma \in \Gamma(\mu, \nu)$  s.t.*

$$\int c d\gamma \leq h(\beta/2).$$

The proof of Lemma 33 relies on the following result, which can be seen as a continuous formulation of Hall's marriage Lemma. A proof is given in [25, Theorem 1.27].

**Lemma 34** (Continuous Hall's marriage lemma). *Let  $M, N$  be Polish spaces, and let  $P$  be a closed subset of  $M \times N$ . Given  $\mu \in \mathcal{P}(M)$  and  $\nu \in \mathcal{P}(N)$ , the following propositions are equivalent:*

- (i)  $\exists \gamma \in \Gamma(\mu, \nu)$  such that  $\text{spt}(\gamma) \subseteq P$  ;
- (ii) for every Borel subset  $B \subseteq M$ ,

$$\nu(\{y \in N \mid \exists x \in B \text{ s.t. } (x, y) \in P\}) \geq \mu(B).$$

*Proof of Lemma 33.* We are going to apply the Continuous Hall's marriage lemma to the set  $P = \{(x, y) \in M \times N, d_M(x, y) \geq \beta/2\}$ . Let  $B$  be any Borel set of  $X$ . We first assume that the diameter of  $B$  is at most  $\beta$  so that  $B \subseteq B(x_0, \beta)$  for some  $x_0 \in B$ . Then,  $\mu(B) \leq \mu(B(x_0, \beta)) \leq 1/2$  using  $M_\mu(\beta) \leq 1/2$ . having also  $M_\nu(\beta) \leq 1/2$  we get

$$\begin{aligned} \nu(\{y \in N \mid \exists x \in B, d_M(x, y) \geq \beta/2\}) &\geq \nu(\{y \in N \mid d_M(x_0, y) \geq \beta/2\}) \\ &= 1 - \nu(B(x_0, \beta/2)) \\ &\geq 1/2 \\ &\geq \mu(B). \end{aligned}$$

Assume now that the diameter of  $B$  is greater than  $\beta$ . Then there exist  $x, x' \in B$  such that  $d_M(x, x') \geq \beta$  and the left hand side of the previous inequation is equal to 1 and the condition is obviously satisfied. We can therefore apply Lemma 34, which implies the existence of a transport plan  $\gamma$  between  $\mu$  and  $\nu$  such that for any pair  $(x, y) \in \text{spt}(\gamma)$  one has  $d_M(x, y) \geq \beta/2$ . Since  $h$  is

decreasing, we have  $c(x, y) \leq h(\beta/2)$  for every pair  $(x, y) \in \text{spt}(\gamma)$ , which implies the desired result.  $\square$

**Lemma 35.** *Let  $\gamma$  be an optimal transport map between  $\mu$  and  $\nu$  for the cost  $c$  and let  $\beta > 0$  such that  $M_\mu(\beta) < 1/8$  and  $M_\nu(\beta) < 1/8$ . Then, for any optimal transport plan  $\gamma \in \Gamma(\mu, \nu)$ , there exists pairs  $(x_0, y_0), (x'_0, y'_0) \in \text{spt}(\gamma)$  such that the four points  $x_0, y_0, x'_0, y'_0$  are at distance at least  $\min(\varepsilon, \beta)$  with  $\varepsilon := h^{-1}(4h(\beta/2))$  from each other.*

*Proof.* Since  $\gamma$  is an optimal transport plan, its cost is less than the cost of the transport plan constructed in Lemma 33. Since  $h$  is decreasing and by definition of  $\Delta_\varepsilon$ , we have for any  $\varepsilon > 0$ ,

$$h(\varepsilon)\gamma(\Delta_\varepsilon) \leq \int_{\Delta_\varepsilon} cd\gamma \leq h(\beta/2)$$

Note that we can consider  $h(\beta/2) > 0$ , choosing a smaller  $\beta$  if necessary. Then if  $\varepsilon = h^{-1}(4h(\beta/2))$  we get  $\gamma(\Delta_\varepsilon) \leq \frac{1}{4}$ , thus proving the existence of a pair  $(x_0, y_0) \in \text{spt}(\gamma) \setminus \Delta_\varepsilon$ .

Since  $M_\mu(\beta) < 1/8$ , one has

$$\gamma((\text{B}(x_0, \beta) \cup \text{B}(y_0, \beta)) \times \mathcal{S}^{d-1}) \leq \mu(\text{B}(x_0, \beta)) + \mu(\text{B}(y_0, \beta)) < \frac{1}{4},$$

Similarly,  $M_\nu(\beta) < 1/8$ , gives

$$\gamma(\mathcal{S}^{d-1} \times (\text{B}(x_0, \beta) \cup \text{B}(y_0, \beta))) \leq \nu(\text{B}(x_0, \beta)) + \nu(\text{B}(y_0, \beta)) < \frac{1}{4},$$

so that

$$\begin{aligned} & \gamma(\{(x, y) \in M^2 \mid d_M(x, x_0) > \beta, d_M(y, y_0) > \beta, d_M(y, -x_0) > \beta, \\ & \quad d_M(x, y_0) > \beta \text{ and } d_M(x, y) > \varepsilon\}) \\ &= \gamma\left(M^2 \setminus \left[ (\text{B}(x_0, \beta) \cup \text{B}(y_0, \beta)) \times \mathcal{S}^{d-1} \right. \right. \\ & \quad \left. \left. \cup \mathcal{S}^{d-1} \times (\text{B}(x_0, \beta) \cup \text{B}(y_0, \beta)) \cup \Delta_\varepsilon \right] \right) \\ & \geq 1 - \gamma((\text{B}(x_0, \beta) \cup \text{B}(y_0, \beta)) \times \mathcal{S}^{d-1}) \\ & \quad - \gamma(\mathcal{S}^{d-1} \times (\text{B}(x_0, \beta) \cup \text{B}(y_0, \beta))) - \gamma(\Delta_\varepsilon) > 1/4. \end{aligned}$$

This proves the existence of  $(x'_0, y'_0) \in \text{spt}(\gamma)$  such that  $d_M(x_0, x'_0) > \beta$  and  $d_M(y_0, y'_0) > \beta$  and  $d_M(x'_0, y'_0) \geq \varepsilon$  and allows us to conclude.  $\square$

**Lemma 36.** *Assume that  $c$  is bounded from below by a constant  $c_{\min}$ . Let  $S \subseteq M \times M$  be a  $c$ -cyclically monotone set, which contains two pairs  $(x_0, y_0), (x'_0, y'_0)$  such that the pairwise distance between the points  $x_0, y_0, x'_0, y'_0$  is at least  $\varepsilon > 0$ . Then,*

$$\forall (x, y) \in S, \quad c(x, y) \leq C_\varepsilon := h(\varepsilon) + 2h(\varepsilon/2) + 2|c_{\min}|.$$

*Proof.* Using the  $c$ -cyclical monotonicity of  $S$  and  $c \geq c_{\min}$  one has

$$c(x, y) \leq c(x, y) + c(x_0, y_0) + c(x'_0, y'_0) - 2c_{\min} \leq F(x, y) + 2|c_{\min}|$$



where

$$\begin{cases} F(x, y) = \min(c(x, y_0) + R_1(y), c(x, y'_0) + R_2(y)) \\ R_1(y) = \min(c(x_0, y) + c(x'_0, y'_0), c(x_0, y'_0) + c(x'_0, y)) \\ R_2(y) = \min(c(x_0, y) + c(x'_0, y_0), c(x_0, y_0) + c(x'_0, y)). \end{cases}$$

By assumption, we have  $d_M(x_0, x'_0) \geq \varepsilon$ , thus  $\max(d_M(x_0, y), d_M(x_0, y) \geq \varepsilon/2$ . Then, since  $h$  is decreasing, one has  $\min(c(x_0, y), c(x'_0, y)) \leq h(\varepsilon/2)$ . We also have  $c(x'_0, y'_0) \leq h(\varepsilon)$  and  $c(x_0, y'_0) \leq h(\varepsilon)$ , which leaves us with

$$R_1(y) \leq h(\varepsilon) + \min(c(x_0, y), c(x'_0, y)) \leq h(\varepsilon) + h(\varepsilon/2),$$

and the same bound holds for  $R_2(y)$ . Using the same argument we get  $\min(c(x, y_0), c(x, y'_0)) \leq h(\varepsilon/2)$  and thus,

$$F(x, y) \leq h(\varepsilon) + h(\varepsilon/2) + \min(c(x, y_0), c(x, y'_0)) \leq h(\varepsilon) + 2h(\varepsilon/2). \quad \square$$

**Proof of Theorem 31.** Let  $\beta > 0$  such that  $M(\beta) > 1/8$ . Let  $\gamma$  be an optimal transport plan, and denote by  $S$  its support. By Lemma 33, the cost of this transport plan is finite. This implies that  $S$  is  $c$ -cyclically monotone. Recall that by assumption, the cost  $c$  is bounded from below. Therefore by Lemmas 35 and 36 one has

$$\forall (x, y) \in S, \quad c(x, y) \leq C_\varepsilon := h(\varepsilon) + 2h(\varepsilon/2) + 2|c_{\min}|.$$

where  $\varepsilon = \min(\beta, h^{-1}(4h(\beta/2)))$ . This directly implies that  $S \subseteq D_\delta$  with  $\delta = h^{-1}(C_\varepsilon)$ .

**5.2. Proof of Theorems 30.** Here, we come back to the sphere case, i.e.  $M = \mathcal{S}^{d-1}$ . We recall that the reflector cost is given on  $M^2$  by  $c(x, y) = -\ln(1 - \langle x|y \rangle)$ . Note that on the unit sphere,  $d_M(x, y) = \arccos(\langle x|y \rangle)$ , hence the reflector cost is of the form  $c(x, y) = h(d_M(x, y))$  with  $h(t) = -\ln(1 - \cos(t))$  and satisfies the assumptions of Theorem 31.

**Lemma 37.** *For  $\varepsilon < 2$ ,  $D_\varepsilon$  is symmetrically  $c$ -convex.*

*Proof.* A simple computation gives for  $x \in M$ , that  $\nabla_x c(x, \cdot) : M \setminus \{x\} \rightarrow T_x M$  is one to one and given by

$$\nabla_x c(x, y) = \frac{y - \langle x|y \rangle x}{1 - \langle x|y \rangle}$$

and the inverse of  $-\nabla_x c(x, \cdot)$  is

$$c\text{-exp}_x(p) = \left(1 - \frac{2}{1 + \|p\|^2}\right)x - \frac{2}{1 + \|p\|^2}p$$

Let  $(x, y_0)$  and  $(x, y_1)$  in  $D_\varepsilon$ , and define the  $c$ -segment  $(y_t) = [y_0, y_1]_x$ . For  $p_0 = \nabla_x c(x, y_0)$  and  $p_1 = \nabla_x c(x, y_1)$ , we put  $p_t = (1 - t)p_0 + tp_1$ , so that  $y_t = c\text{-exp}_x(p_t)$ . We want to show that  $(x, y_t) \in D_\varepsilon$ , hence we only have to show that  $d_M(x, y_t) \geq \varepsilon$ . We have

$$x - y_t = \frac{2}{1 + \|p_t\|^2}x + \frac{2}{1 + \|p_t\|^2}p_t.$$

Since  $x$  is orthogonal to  $p_t$  and  $\|x\| = 1$ , we get

$$d_M(x, y_t) = \arccos(\langle x|y_t \rangle) = \arccos\left(1 - \frac{2}{1 + \|p_t\|^2}\right).$$

So  $d_M(x, y_t) \geq \varepsilon$  is satisfied if  $1 - \frac{2}{1 + \|p_t\|^2} \leq \cos(\varepsilon)$ . Since  $\cos(\varepsilon) \geq 1 - \varepsilon^2/2$  it is sufficient to show that

$$\frac{2}{1 + \|p_t\|^2} \geq \varepsilon^2/2.$$

Since  $\|p_t\| \leq \max(\|p_0\|, \|p_1\|)$ , and by symmetry of  $p_0$  and  $p_1$  it is sufficient to show that  $\|p_0\|^2 \leq \frac{4}{\varepsilon^2} - 1$ . Again using that  $\|x\| = \|y_0\| = 1$ , we have

$$\|p_0\|^2 = \left\| \frac{y_0 - \langle x|y_0 \rangle x}{1 - \langle x|y_0 \rangle^2} \right\|^2 = \frac{1 - \langle x|y_0 \rangle^2}{(1 - \langle x|y_0 \rangle^2)^2} = \frac{1 + \langle x|y_0 \rangle}{1 - \langle x|y_0 \rangle}$$

Finally using the relation  $\langle x|y_0 \rangle = 1 - \|x - y_0\|^2/2$ , we get

$$\|p_0\|^2 = \frac{4}{\|x - y_0\|^2} - 1 \leq \frac{4}{\varepsilon^2} - 1$$

and in conclusion,  $D_\varepsilon$  is  $c$ -convex. Note that by symmetry it is obviously symmetrically  $c$ -convex.  $\square$

**End of proof of Theorem 30** Since  $\mu$  and  $\nu_0$  are absolutely continuous there exists  $\beta > 0$  such that  $M_\mu(\beta) < 1/8$ ,  $M_{\nu_0}(\beta) < 1/8$  and  $M_{\nu_1}(\beta) < 1/8$ . Therefore, by Theorem 31, there exists  $\varepsilon > 0$  such that for every  $x \in M$ ,  $(x, T_i(x)) \in D_\varepsilon$ . The set  $D_\varepsilon$  is a compact set and symmetrically  $c$ -convex by Lemma 37. Recall that the optimal transport map  $T_0$  between  $\mu$  and  $\nu_0$  is of the form  $T_0(x) = \operatorname{argmin}_{y \in N} c(x, y) - \psi_0(y)$ , where  $\psi_0 : N \rightarrow \mathbb{R}$  is a  $c$ -concave function. Since  $\mu$  and  $\nu_0$  have  $\mathcal{C}^{1,1}$  strictly positive densities, a result of Gregoire Loeper [19, Theorem 2.5] implies that  $\psi_0$  is of class  $\mathcal{C}^3$  and that  $T : x \mapsto c\text{-exp}_x(\nabla \psi^c(x))$  is of class  $\mathcal{C}^2$ . As seen in the proof of Theorem 31,  $\psi_1$  is  $c$ -concave for the truncated cost, which is Lipschitz, and is therefore also Lipschitz. Furthermore, it is known that the reflector cost satisfies MTW and (STwist) [19]. We can thus apply Corollary 23 which gives that  $\psi_0$  is strongly  $c$ -concave on  $D_\varepsilon$ . We then conclude by applying Theorem 5.

## 6. PRESCRIPTION OF GAUSS CURVATURE MEASURE

The problem of Gauss curvature measure prescription for a convex body has been introduced by A.D. Aleksandrov in 1950 [1] and has been shown to be equivalent to an optimal transport problem on the sphere [22, 5]. In this section we apply our stability result to this optimal transport problem.

To this purpose we define the Gauss curvature measure introduced in [1]. Let  $K \subseteq \mathbb{R}^d$  be a closed bounded convex body such that  $0 \in \operatorname{int}(K)$ . We denote by  $\rho_K : \mathcal{S}^{d-1} \rightarrow \mathbb{R}$  the radial parametrization of  $\partial K$  defined for any direction  $x$  in the sphere  $\mathcal{S}^{d-1}$  by  $\rho_K(x) = \sup\{r \in \mathbb{R} \mid rx \in K\}$ . This induces a homeomorphism  $\overrightarrow{\rho}_K$  from  $\mathcal{S}^{d-1}$  to  $\partial K$  defined by

$$\begin{aligned} \overrightarrow{\rho}_K : \mathcal{S}^{d-1} &\rightarrow \partial K \\ x &\mapsto \rho_K(x)x \end{aligned}$$

We call (multivalued) Gauss map, the map  $\mathcal{G}_K$  which maps a point  $x \in \partial K$  to the set of unit exterior normals to  $K$  at  $x$ , namely

$$G_K(x) = \{n \in \mathcal{S}^{d-1} \mid x \in \arg \max_K \langle n, \cdot \rangle\}.$$

Note that  $\mathcal{G}_K(x)$  is a set when  $K$  is not smooth at  $x$ . Through this section, we denote by  $\sigma$  the uniform probability measure on the sphere  $\mathcal{S}^{d-1}$ , i.e. the normalized  $(d-1)$ -dimensional Hausdorff measure.

**Definition 38** (Gauss curvature measure). *Let  $K$  be a bounded convex body containing  $0$  in its interior. The Gauss curvature measure of  $K$ , denoted  $\mu_K$ , is a probability measure over  $\mathcal{S}^{d-1}$  defined for any Borel subset  $A \subseteq \mathcal{S}^{d-1}$  by  $\mu_K(A) = \sigma(\mathcal{G}_K \circ \overrightarrow{\rho}_K(A))$ .*

The *Gauss curvature measure prescription problem* is the following inverse problem: given a measure  $\mu \in \mathcal{P}(\mathcal{S}^{d-1})$ , is it possible to find a convex body  $K$  such that  $\mu = \mu_K$ ? It is well-known that convexity of  $K$  implies that for every non-empty spherical convex subset  $\Theta \subsetneq \mathcal{S}^{d-1}$  — i.e. subsets  $\Theta$  that contains any minimizing geodesic between any pair of its points — we have

$$\mu_K(\Theta) < \sigma(\Theta_{\pi/2}) \quad (6.14)$$

with  $\Theta_{\pi/2} = \{x \in \mathcal{S}^{d-1} \mid d_M(x, \Theta) < \pi/2\}$ , and where  $d_M$  is the geodesic distance on the sphere. Aleksandrov's theorem states that Equation (6.14) is in fact a sufficient condition for  $\mu$  to be the Gauss curvature measure of a convex body.

**Theorem 39** (Aleksandrov). *Let  $\mu \in \mathcal{P}(\mathcal{S}^{d-1})$  be a probability measure satisfying condition (6.14), then there exists a unique (up to homotheties) convex body  $K \subseteq \mathbb{R}^d$  with  $0 \in \text{int}(K)$  such that  $\mu$  is the Gaussian curvature measure of  $K$ .*

**6.1. An optimal transport problem.** Following [22, 5] we briefly recall that this inverse problem can be recast as an optimal transport problem on the sphere for the cost  $c(x, n) = -\ln(\max(0, \langle x|n \rangle))$ , which takes value  $+\infty$  when  $\langle x|n \rangle \leq 0$ . Let  $\mu$  be any measure in  $\mathcal{P}(\mathcal{S}^{d-1})$  satisfying condition (6.14). Note that the very same cost plays an important role in the theory of unbalanced optimal transport [9, 18, 11].

In the following proposition, we use the notion of *support function* of a convex set  $K$ , defined by

$$h_K(n) = \sup_{x \in \mathcal{S}^{d-1}} \rho_K(x) \langle x|n \rangle.$$

**Proposition 40** ([22, 5]). *Let  $\sigma \in \mathcal{P}(\mathcal{S}^{d-1})$  be the uniform measure over the sphere, let  $K$  be a compact convex body containing zero in its interior, and let  $\mu = \mu_K$ . Then,*

- The map  $T_K : \mathcal{S}^{d-1} \rightarrow \mathcal{S}^{d-1}$  defined  $\sigma$ -a.e by

$$T_K(n) = (\mathcal{G}_K \circ \overrightarrow{\rho}_K)^{-1}(n)$$

*is the optimal transport map between  $\sigma$  and  $\mu$  for the cost  $c$ .*

- The functions  $\varphi_K = -\ln(h_K)$  and  $\psi_K = \ln(\rho_K)$  are maximizers of the Kantorovich dual problem. In particular we have

$$\int_{\mathcal{S}^{d-1}} c(T_K(n), n) d\sigma(n) = \int \varphi_K(n) d\sigma(n) + \int \psi_K(x) d\mu_K(x). \quad (6.15)$$

For the sake of completeness, we recall the proof of this proposition.

*Proof.* Let  $(x, n) \in \mathcal{S}^{d-1} \times \mathcal{S}^{d-1}$  be such that  $c(x, n) < +\infty$ , i.e.  $\langle x|n \rangle > 0$ . Then,

$$h_K(n) = \max_{y \in K} \langle n|y \rangle \geq \langle n|\rho_K(x)x \rangle = \rho_K(x) \langle n|x \rangle, \quad (6.16)$$

with equality if and only if  $n \in \mathcal{G}_K(x)$ . Since all quantities are positive, taking the logarithm, we see that  $\varphi_K(n) + \psi_K(x) \leq c(x, n)$ , ensuring that  $(\varphi_K, \psi_K)$  are admissible for the dual Kantorovich problem.

Note that e.g. by [5]  $\sigma$ -a.e. direction  $n \in \mathcal{S}^{d-1}$  is normal to a unique point in  $\partial K$ . This implies that the map  $T_K = (\mathcal{G}_K \circ \overrightarrow{\rho_K})^{-1}$  is well defined  $\sigma$ -a.e. The equality case of (6.16) gives

$$\varphi_K(n) + \psi_K(T_K(n)) \leq c(x, T_K(n)).$$

Integrating this equality with respect to  $\sigma$  directly gives (6.15). In turn, Kantorovich duality implies that  $T_K$  is an optimal transport between  $\sigma$  and  $\mu$ , and that  $(\varphi_K, \psi_K)$  is a maximizer in the dual Kantorovich problem.  $\square$

**6.2. Stability of transport maps.** In this subsection we apply our stability result to the Gauss curvature measure prescription problem. We introduce the following notation:

$$\mathcal{K}(r, R) = \{K \subseteq \mathbb{R}^d \text{ convex, compact} \mid B(0, r) \subseteq K \subseteq B(0, R)\}.$$

**Proposition 41.** *Let  $K$  be a strictly convex and  $\mathcal{C}^2$  compact convex body containing 0 in its interior. Then, for any  $R > r > 0$ , there exists a constant  $C$  depending on  $K$ ,  $r$  and  $R$  such that*

$$\forall L \in \mathcal{K}(r, R), \quad \|d_M(T_K, T_L)\|_{L^2(\sigma)}^2 \leq C W_1(\mu_K, \mu_L).$$

Note that in addition to the strict convexity and smoothness of  $K$ , the constant  $C$  also depends on the anisotropy of  $K$  — i.e. the radii  $R_K \geq r_K > 0$  such that  $K \in \mathcal{K}(r_K, R_K)$ . The end of the section is devoted to the proof of Proposition 41. We need to check that the hypothesis of Corollary 23 are satisfied for the cost  $c(x, n) = -\ln(\max(0, \langle x|n \rangle))$ .

**Lemma 42.** *Given any  $R > r > 0$ , there exists  $\varepsilon > 0$  such that for any set  $K \in \mathcal{K}(r, R)$  and any  $c$ -optimal transport plan  $\gamma \in \Gamma(\sigma, \mu_K)$ , one has*

$$\text{spt}(\gamma) \subseteq D_\varepsilon,$$

where  $D_\varepsilon = \{(x, n) \in (\mathcal{S}^{d-1})^2 \mid d_M(x, n) \leq \pi/2 - \varepsilon\}$ .

*Proof.* By hypothesis,  $r \leq \rho_K(x) \leq R$  for all  $x \in \mathcal{S}^{d-1}$ , where  $\rho_K$  is the radial function of the convex  $K$ . Since

$$h_K(n) = \sup_{x \in \mathcal{S}^{d-1}} \rho_K(x) \langle x|n \rangle,$$

we also have  $r < h_K(n) < R$ . Hence the two Kantorovich potential  $\varphi_K(n) = -\ln(h_K(n))$  and  $\psi_K(x) = \ln(\rho_K(x))$  therefore satisfy

$$\varphi_K(n) + \psi_K(x) \leq -\ln(r) + \ln(R) = \ln(R/r),$$

By strong Kantorovich duality  $\varphi_K(n) + \psi_K(x) = c(x, n)$  on  $\text{spt}(\gamma)$ , which implies that  $c$  is bounded by  $\ln(R/r)$  on  $\text{spt}(\gamma)$ , i.e. for any  $(x, n) \in \text{spt}(\gamma)$ , one has

$$c(x, n) = -\ln(\max(0, \langle x|n \rangle)) \leq \ln(R/r),$$

implying that  $\langle x|n \rangle \geq r/R$  and  $d_M(x, n) = \arccos(\langle x|n \rangle) \leq \arccos(r/R)$ . Finally  $(x, n) \in D_\varepsilon$  with  $\varepsilon = \pi/2 - \arccos(r/R)$ .  $\square$

**Lemma 43.** *The set  $D_\varepsilon = \{(x, n) \in (\mathcal{S}^{d-1})^2 \mid d_M(x, n) \leq \pi/2 - \varepsilon\}$  is symmetrically c-convex for the cost  $c(x, n) = -\ln(\max(0, \langle x|n \rangle))$ .*

*Proof.* We have

$$\nabla_x c(x, n) = -\frac{n}{\langle x|n \rangle} + x$$

and by inverting  $-\nabla_x c(x, \cdot)$  we get

$$\text{c-exp}_x(p) = \frac{p + x}{\sqrt{1 + \|p\|^2}}$$

Let  $(x, y_0) \in D_\varepsilon$  and  $(x, y_1) \in D_\varepsilon$  then we have  $y_t = \text{c-exp}_x(p_t)$  where  $p_0 = -\nabla_x c(x, y_0)$  and  $p_1 = -\nabla_x c(x, y_1)$  and  $p_t = (1-t)p_0 + tp_1$ . By symmetry we can consider that  $\|p_t\| \leq \|p_0\|$ , which implies  $\frac{1}{\sqrt{1+\|p_t\|^2}} \geq \frac{1}{\sqrt{1+\|p_0\|^2}}$  and thus

$$\begin{aligned} d_M(x, y_t) &= \arccos(\langle x|y_t \rangle) = \arccos\left(\frac{1}{\sqrt{1 + \|p_t\|^2}}\right) \\ &\leq \arccos\left(\frac{1}{\sqrt{1 + \|p_0\|^2}}\right) = d_M(x, y_0) \leq \frac{\pi}{2} - \varepsilon \quad \square \end{aligned}$$

*End of proof of Proposition 41.* The map  $T_K$  (resp.  $T_L$ ) is the optimal transport map between the uniform measure  $\sigma$  on  $\mathcal{S}^{d-1}$  and  $\mu_K$  (resp.  $\mu_L$ ) for the cost  $c(x, n) = -\ln(\max(0, \langle x|n \rangle))$ . From Lemma 42, for any  $n \in \mathcal{S}^{d-1}$  we have  $(T_K(n), n) \in D_\varepsilon$  and  $(T_L(n), n) \in D_\varepsilon$ . Note that for  $(x, n) \in D_\varepsilon$ , one has  $\langle x|n \rangle > 0$  and therefore  $c(x, n) = -\ln(\langle x|n \rangle) = -\ln(\cos(d_M(x, n)))$ . It has been shown in [11] that this cost satisfies (STwist) and (MTWw) on  $D_\varepsilon$ . By Lemma 43 the set  $D_\varepsilon$  is a symmetrically c-convex compact set.

Finally it remains to show that  $\psi_K$  is of class  $\mathcal{C}^2$  and  $T_K$  is of class  $\mathcal{C}^1$ . Since  $\partial K$  is  $\mathcal{C}^2$ , its radial parametrization  $\rho_K$  is also  $\mathcal{C}^2$ , so  $\psi_K = \ln(\rho_K)$  of class  $\mathcal{C}^2$ . Furthermore  $\overrightarrow{\rho_K}(x) = \rho_K(x)x$  is a  $\mathcal{C}^1$  diffeomorphism. Since  $K$  is strictly convex and  $\partial K$  is of class  $\mathcal{C}^2$ , its associated Gauss map  $\mathcal{G}_K$  is a  $\mathcal{C}^1$  diffeomorphism. We thus have that  $T_K = (\mathcal{G}_K \circ \overrightarrow{\rho_K})^{-1}$  is of class  $\mathcal{C}^1$ . By Corollary 23, we know that  $\psi_K$  is strongly c-concave. We conclude by applying Theorem 5.  $\square$

## REFERENCES

- [1] AD Alexandrov. Convex polyhedra. springer monographs in mathematics. springer-verlag, berlin, 2005. 1950.
- [2] Luigi Ambrosio, Federico Glaudo, and Dario Trevisan. On the optimal map in the 2-dimensional random matching problem. *Discrete & Continuous Dynamical Systems*, 39(12):7291–7308, 2019.
- [3] Jean-David Benamou. Optimal transportation, modelling and numerical simulation. *Acta Numerica*, 30:249–325, 2021.
- [4] Robert J Berman. Convergence rates for discretized monge–ampère equations and quantitative stability of optimal transport. *Foundations of Computational Mathematics*, 21(4):1099–1140, 2021.

- [5] Jérôme Bertrand. Prescription of gauss curvature using optimal mass transport. *Geometriae Dedicata*, 183, 08 2016.
- [6] David P Bourne, Charlie P Egan, Beatrice Pelloni, and Mark Wilkinson. Semi-discrete optimal transport methods for the semi-geostrophic equations. *Calculus of Variations and Partial Differential Equations*, 61(1):1–34, 2022.
- [7] Yann Brenier. Polar factorization and monotone rearrangement of vector-valued functions. *Communications on pure and applied mathematics*, 44(4):375–417, 1991.
- [8] Giuseppe Buttazzo, Thierry Champion, and Luigi De Pascale. Continuity and estimates for multimarginal optimal transportation problems with singular costs. *Applied Mathematics & Optimization*, 78(1):185–200, 2018.
- [9] Lenaïc Chizat, Gabriel Peyré, Bernhard Schmitzer, and François-Xavier Vialard. An interpolating distance between optimal transport and fisher–rao metrics. *Foundations of Computational Mathematics*, 18(1):1–44, 2018.
- [10] Alex Delalande and Quentin Merigot. Quantitative stability of optimal transport maps under variations of the target measure. *arXiv preprint arXiv:2103.05934*, 2021.
- [11] Thomas Gallouët, Roberta Ghezzi, and François-Xavier Vialard. Regularity theory and geometry of unbalanced optimal transport, 2021.
- [12] Wilfrid Gangbo and Robert J McCann. Optimal maps in monge’s mass transport problem. *Comptes Rendus de l’Academie des Sciences-Serie I-Mathematique*, 321(12):1653, 1995.
- [13] Wilfrid Gangbo and Vladimir Oliker. Existence of optimal maps in the reflector-type problems. *ESAIM Control Optim. Calc. Var.*, 13, 03 2007.
- [14] Nicola Gigli. On hölder continuity-in-time of the optimal transport map towards measures along a curve. *Proceedings of the Edinburgh Mathematical Society*, 54(2):401–409, 2011.
- [15] Jan-Christian Hütter and Philippe Rigollet. Minimax estimation of smooth optimal transport maps. *The Annals of Statistics*, 49(2):1166–1194, 2021.
- [16] Jun Kitagawa, Quentin Mérigot, and Boris Thibert. Convergence of a newton algorithm for semi-discrete optimal transport. *Journal of the European Mathematical Society*, 21(9):2603–2651, 2019.
- [17] Wenbo Li and Ricardo H Nochetto. Quantitative stability and error estimates for optimal transport plans. *IMA Journal of Numerical Analysis*, 41(3):1941–1965, 2021.
- [18] Matthias Liero, Alexander Mielke, and Giuseppe Savaré. Optimal entropy-transport problems and a new hellinger–kantorovich distance between positive measures. *Inventiones mathematicae*, 211(3):969–1117, 2018.
- [19] Grégoire Loeper. Regularity of optimal maps on the sphere: The quadratic cost and the reflector antenna. *Archive for rational mechanics and analysis*, 199(1):269–289, 2011.
- [20] Xi-Nan Ma, Neil S Trudinger, and Xu-Jia Wang. Regularity of potential functions of the optimal transportation problem. *Archive for rational mechanics and analysis*, 177(2):151–183, 2005.
- [21] Quentin Merigot and Boris Thibert. Optimal transport: discretization and algorithms. In *Handbook of Numerical Analysis*, volume 22, pages 133–212. Elsevier, 2021.
- [22] Vladimir Oliker. Embedding  $S^n$  into  $R^{n+1}$  with given integral gauss curvature and optimal mass transport on  $S^n$ . *Advances in Mathematics*, 213(2):600–620, 2007.
- [23] Gabriel Peyré, Marco Cuturi, et al. Computational optimal transport: With applications to data science. *Foundations and Trends® in Machine Learning*, 11(5-6):355–607, 2019.
- [24] Filippo Santambrogio. Optimal transport for applied mathematicians. *Birkhäuser, NY*, 55(58-63):94, 2015.
- [25] C. Villani. *Topics in Optimal Transportation*. Graduate studies in mathematics. American Mathematical Society, 2003.
- [26] C. Villani. *Optimal Transport: Old and New*. Grundlehren der mathematischen Wissenschaften. Springer Berlin Heidelberg, 2008.
- [27] Xu-Jia Wang. On the design of a reflector antenna ii. *Calculus of Variations and Partial Differential Equations*, 20(3):329–341, 2004.

UNIVERSITÉ GRENOBLE ALPES, CNRS, GRENOBLE INP, LJK, 38000 GRENOBLE,  
FRANCE

UNIVERSITÉ PARIS-SACLAY, CNRS, LABORATOIRE DE MATHÉMATIQUES D'ORSAY,  
91405, ORSAY, FRANCE

UNIVERSITÉ GRENOBLE ALPES, CNRS, GRENOBLE INP, LJK, 38000 GRENOBLE,  
FRANCE