NECESSARY AND SUFFICIENT CONDITIONS FOR CONTINUITY OF OPTIMAL TRANSPORT MAPS ON RIEMANNIAN MANIFOLDS

A. FIGALLI, L. RIFFORD, AND C. VILLANI

Abstract. In this paper we continue the investigation of the regularity of optimal transport maps on Riemannian manifolds, in relation with the geometric conditions of Ma–Trudinger–Wang and the geometry of the cut locus. We derive some sufficient and some necessary conditions to ensure that the optimal transport map is always continuous. In dimension two, we can sharpen our result into a necessary and sufficient condition. We also provide some sufficient conditions for regularity, and review existing results.

1. Introduction

Throughout this paper, M will stand for a smooth compact connected n-dimensional Riemannian manifold ($n \geq 2$) with its metric tensor g, its geodesic distance d, and its volume vol. Reminders and basic notation from Riemannian geometry (exponential map, cut and focal loci, injectivity domain, etc.) are gathered in Appendix A.[1]

Let μ, ν be two probability measures on M and let $c(x,y) = d(x,y)^2/2$. The Monge problem with measures μ, ν and cost c consists in finding a map $T : M \to M$ which minimizes the cost functional $\int_M c(x, T(x))\, d\mu(x)$ under the constraint $T_\# \mu = \nu$ (ν is the image measure of μ by T).

If μ does not give mass to countably $(n-1)$-rectifiable sets, then according to McCann [29] (see also [9, 10]) this minimizing problem has a solution T, unique up to modification on a μ-negligible set; moreover T takes the form $T(x) = \exp_x(\nabla_x \psi)$, where $\psi : M \to \mathbb{R}$ is a c-convex function [31, Definition 5.2]. In the sequel, μ will be absolutely continuous with respect to the volume measure, and its density will be bounded from above and below; so McCann's theorem applies and T is uniquely determined up to modification on a set of zero volume.

[1] In this paper as in our previous works [27, 14, 15, 16, 17] we use the word "focal" as synonymous of "conjugate".

The regularity of T is in general a more subtle problem which has received much attention in recent years [31, Chapter 12]. A first question is whether the optimal transport map can be expected to be continuous.

Definition 1.1. We say that M satisfies the *transport continuity property* (abbreviated (**TCP**))[2] if, whenever μ and ν are absolutely continuous measures with respect to the volume measure, with densities bounded away from zero and infinity, the optimal transport map T with measures μ, ν and cost c is continuous, up to modification on a set of zero volume.

The aim of this paper is to give necessary and sufficient conditions for (**TCP**). This problem involves two geometric conditions: the condition of **convexity of injectivity domains**, and the **Ma–Trudinger–Wang condition**. These conditions were first introduced and studied by Ma, Trudinger and Wang [28] and Loeper [26] outside the Riemannian world; the natural Riemannian adaptation of these concepts was made precise in Loeper and Villani [27] and further developed by Figalli and Rifford [14]. Both conditions come in the form of more or less stringent variants.

• We say that M satisfies **(CI)** (resp. **(CI$^+$)**) if for any $x \in M$ the injectivity domain $\mathrm{I}(x) \subset T_xM$ is convex (resp. strictly convex).

• For any $x \in M$, $v \in \mathrm{I}(x)$ and $(\xi, \eta) \in T_xM \times T_xM$, let $y = \exp_x v$, we define the Ma–Trudinger–Wang tensor at (x, y), evaluated on (ξ, η), by the formula

$$\mathfrak{S}_{(x,y)}(\xi,\eta) = -\frac{3}{2}\,\frac{d^2}{ds^2}\bigg|_{s=0}\frac{d^2}{dt^2}\bigg|_{t=0} c\Big(\exp_x(t\xi),\,\exp_x(v+s\eta)\Big).$$

Then we say that M satisfies **(MTW)** if

$$(1.1)\quad \forall\,(x,v) \in TM \text{ with } v \in \mathrm{I}(x),\ \forall\,(\xi,\eta) \in T_xM \times T_xM,$$
$$\Big[\langle \xi, \eta\rangle_x = 0 \Longrightarrow \quad \mathfrak{S}_{(x,y)}(\xi,\eta) \geq 0\Big].$$

If the last inequality in (1.1) is strict unless $\xi = 0$ or $\eta = 0$, then M is said to satisfy the *strict* Ma–Trudinger–Wang condition **(MTW$^+$)**.

In the case of nonfocal Riemannian manifolds, that is, when the injectivity domains do not intersect the focal tangent cut locus, Loeper and Villani [27] showed that **(MTW$^+$)** implies **(CI$^+$)**, (**TCP**), and then further regularity properties.

The much more tricky focal case was first attacked in [14] and pursued in a series of works by the authors [15, 16, 17]. In presence of focalization, the robustness of the results becomes an issue, since then the stability of conditions **(MTW$^+$)** and

[2]This definition differs slightly from that in [14, Definition 1.1].

(CI$^+$) is not guaranteed under perturbations of the metric. In [15] it was shown that perturbations of the sphere satisfy **(MTW$^+$)** and **(CI$^+$)**. In [16] we showed that in dimension $n = 2$ these two conditions are stable under perturbation, provided that the nonfocal domains are strictly convex near the tangent cut locus (see [16] for a precise formulation). In [17] we showed that the same condition implies the semiconvexity of injectivity domains, without assumption on the Ma–Trudinger–Wang tensor. After this set of studies which were so to speak purely geometric, the goal of the current paper is to connect these conditions to the regularity problem, also in the possible presence of focalization.

Our first main result is as follows:

Theorem 1.2. *With the notation above,*

(i) *If M satisfies* (**TCP**), *then* **(CI)** *and* **(MTW)** *hold.*
(ii) *If M satisfies* **(CI$^+$)** *and* **(MTW$^+$)**, *then* (**TCP**) *holds.*

Notice the gap between the necessary (weak) and sufficient (strong) conditions. In dimension $n = 2$, we can fill the gap and show that the weak conditions are necessary and sufficient:

Theorem 1.3. *If M is a (compact connected Riemannian) surface, then M satisfies* (**TCP**) *if and only if* **(CI)** *and* **(MTW)** *hold.*

The proofs of Theorems 1.2 and 1.3, they are given in Sections 2 to 5. The improvement from Theorem 1.2 to Theorem 1.3 is based on the following two-dimensional results (combined with a delicate geometric reasoning):

1) In $\mathbb{R}^2$, continuity of optimal maps between densities bounded away from zero and infinity is known to be true under **(MTW)** [13].

2) If **(MTW)** holds, then, for any $x \in M$, the curvature of the tangent focal locus at any point of the tangent cut locus has to be nonnegative. (Although not explicitly stated in this way, this fact is an immediate consequence of the proof of [16, Proposition 4.1(ii)].)

Since property **(CI)** is closed under C^2-convergence of the metric, as an immediate consequence of Theorem 1.3 and Lemma 3.1, we deduce that the set of two-dimensional manifolds satisfying (**TCP**) is closed in C^2-topology (compare with the stability results in [32]).

In [14, 15, 27] it is shown how a Ma–Trudinger–Wang condition, combined with additional geometric assumptions, implies the convexity of injectivity domains. So one could make the more daring conjecture that (**TCP**) and **(MTW)** are in fact equivalent. For the moment this seems out of technical reach, and we do not know if one can dispend with the explicit assumptions of convexity of injectivity domains. Such conditions may be impossible to check in practice because of the complexity of the cut locus; so it may be useful to know that they can be replaced by stronger conditions on the shape of the nonfocal domain. This will be described in Section 6.

From the present study and previous works, we can write a list of simple or specific Riemannian manifolds known to satisfy (**TCP**):

– Flat tori in any dimension [4].
– Round spheres in any dimension [26]
– Small C^4 deformations of round spheres in any dimension [14, 15].
– Riemannian submersions of round spheres [23].
– Products of round spheres [12].
– Quotients of all the above examples by a discrete group of isometry [5].
– Compact quotients of products of spheres and Euclidean spaces with nonfocal cut locus (Proposition 6.4).
– Compact Riemannian surfaces satisfying **(CI)** and **(MTW)** (Theorem 1.3).

Now as far as the proofs are concerned, many ingredients are recycled from previous works [14, 15, 16, 22, 27], however they are improved and combined with new technical results, such as Lemma B.2 about the structure of the cut locus. Some notation and technical results are postponed to the appendices.

We close this introduction by noting that the body of work described above connects the classical nonsmooth topic of the regularity and shape of cut locus [3, 21, 24] with the not less classical smooth topic of the regularity of solutions of fully nonlinear partial differential equations [20]. See [33] for a short description of this connection.

2. Necessary conditions for (TCP)

The following assumption was introduced in [28] and further studied in [26, 31].

Assumption (C): *For any c-convex function* ψ *and any* $x \in M$*, the c-subdifferential* $\partial_c\psi(x)$ *is pathwise connected.*

Background on c-subdifferentials can be found in [31, Chapter 5]. It is a general fact that (**TCP**) implies **Assumption (C)**; this is true even for much more general cost functions c [31, Theorem 12.7]. Moreover, **Assumption (C)** and **(CI)** together imply **(MTW)** (see [31, Theorem 12.42]). So Theorem 1.2(i) is a straightforward consequence of the following result:

Proposition 2.1. *For the cost function $c = d^2/2$,* **Assumption (C)** *implies* **(CI)**.

To establish Proposition 2.1 we shall use a technical auxiliary result, whose proof is postponed to the end of the section:

Lemma 2.2. *Assume $\mathrm{I}(\bar{x})$ is not convex for some $\bar{x} \in M$. Then there are $v_{-1}, v_1 \in \mathrm{I}(\bar{x})$ such that:*

- *$v_0 = \frac{v_{-1}+v_1}{2}$ does not belong to $\overline{\mathrm{I}(\bar{x})}$;*
- *the mapping $v \in [v_{-1}, v_1] \cap \overline{\mathrm{I}(\bar{x})} \mapsto \exp_{\bar{x}}(v) \in M$ is injective.*

Proof of Proposition 2.1. Assume by contradiction that $\mathrm{I}(\bar{x})$ (or equivalently $\overline{\mathrm{I}(\bar{x})}$) is not convex for some $\bar{x} \in M$. Let v_{-1}, v_0, v_1 be as in Lemma 2.2. Set $y_i = \exp_{\bar{x}}(v_i)$ for $i = -1, 0, 1$ and let $\psi : M \to \mathbb{R}$ be the c-convex function defined by

$$\psi(x) = \max\Big\{c(\bar{x}, y_{-1}) - c(x, y_{-1}),\, c(\bar{x}, y_1) - c(x, y_1)\Big\} \qquad \forall\, x \in M.$$

The set of subgradients of ψ at $\bar{x}$ is given by the segment

$$\nabla^- \psi(\bar{x}) = [v_{-1}, v_1] \subset T_{\bar{x}}M,$$

which by [31, Theorem 10.25] implies

$$-\nabla^+_{\bar{x}} c(\bar{x}, y) \subset [v_{-1}, v_1] \qquad \forall\, y \in \partial_c \psi(\bar{x}),$$

where $\nabla^+_{\bar{x}} c(\bar{x}, y)$ denotes the set of supergradients of the semiconcave function $x \mapsto c(x, y)$ at $\bar{x}$. (Recall that, on a Riemannian maninfold M, the squared distance function is a locally semiconcave function on $M \times M$, see for instance [9, Appendix B].) Since both v_{-1} and v_1 belong to the injectivity domain $\mathrm{I}(\bar{x})$, $x \mapsto c(x, y_i)$ is differentiable at $\bar{x}$ and $\nabla_{\bar{x}} c(\bar{x}, y_i) = -v_i$, for $i = -1, 1$. Let us define the multivalued function

$$F : y \in \partial_c \psi(\bar{x}) \longmapsto -\nabla^+_{\bar{x}} c(\bar{x}, y) \subset [v_{-1}, v_1]$$

By **Assumption (C)**, there exists a continuous curve $t \in [-1, 1] \to y(t) \in M$, valued in $\partial_c \psi(x)$, such that $y(-1) = y_{-1}$ and $y(1) = y_1$. Observe that $F \circ y$ is convex-valued and upper semicontinuous, taking values $\{v_{-1}\}$ and $\{v_1\}$ at $t = -1$ and $t = 1$ respectively. Since v_0 does not belong to $\overline{\mathrm{I}(\bar{x})}$, there has to be some $t_* \in (-1, 1)$ such that $F(y(t_*))$ is not contained in $\overline{\mathrm{I}(\bar{x})}$. (To prove this rigorously, one can appeal to [1, Theorem 9.2.1]: $F \circ y$ admits a Lipschitz approximate section, that is, for every $\epsilon > 0$ we can find a Lipschitz function $f_\epsilon : [-1, 1] \to [v_{-1}, v_1]$ such that

$$\mathrm{Graph}(f_\epsilon) \subset \mathrm{Graph}\big(F \circ y\big) + \epsilon B = \Big\{(t, v);\, t \in [-1, 1], v \in \big(F \circ y\big)(t)\Big\} + \epsilon B,$$

in $[-1,1] \times [v_{-1}, v_1]$, where B denotes the open unit ball in $\mathbb{R}^2$; from this and a compactness argument the conclusion is easy.)

Thus there is $v_* \in F(y(t_*))$ such that

$$v_* \in -\nabla^+_{\bar{x}} c(\bar{x}, y(t_*)) \setminus \overline{\mathrm{I}(\bar{x})}.$$

But the set $-\nabla^+_{\bar{x}} c(\bar{x}, y(t_*))$ is the convex hull of the minimizing velocities joining $\bar{x}$ to $y(t_*)$. Hence, recalling that $-\nabla^+_{\bar{x}} c(\bar{x}, y(t_*)) \subset [v_{-1}, v_1]$, we can find two such minimizing velocities $v, v' \in [v_{-1}, v_1]$ such that v_* is a convex combination of v and v'. In particular $\exp_{\bar{x}} v = \exp_{\bar{x}} v'$, contradicting the injectivity condition in Lemma 2.2. This establishes **(CI)**. □

Proof of Lemma 2.2. It is sufficient to show that there are $v_{-1}, v_1 \in \mathrm{I}(\bar{x})$ such that $v_0 = \frac{v_{-1}+v_1}{2}$ does not belong to $\mathrm{I}(\bar{x})$ and such that the mapping

$$v \in [v_{-1}, v_1] \cap \overline{\mathrm{TCL}(\bar{x})} \mapsto |v|_{\bar{x}} \in \mathbb{R}$$

is injective. Indeed, if v, v' belong to $[v_{-1}, v_1]$ then: either at least one of them is in the injectivity domain and then necessarily $\exp_x v \neq \exp_x v'$; or both are in the tangent cut locus and then $d(x, \exp_x v) = |v| \neq |v'| = d(x, \exp_x v')$.

Denote by $\tau_C : U_{\bar{x}}M \to (0, \infty)$ the restriction of the cut time $t_C(\bar{x}, \cdot)$ (see Appendix A) to the unit sphere $U_{\bar{x}}M \subset T_{\bar{x}}M$; then $\mathrm{TCL}(\bar{x})$ is the image of the map $\widehat{w} \in U_{\bar{x}}M \mapsto w = \tau_C(\widehat{w})\widehat{w} \in T_{\bar{x}}M$. In particular, at each differentiability point $\widehat{w}$ of τ_C, there is a uniquely determined unit exterior normal $\xi(w)$ to $\mathrm{TCL}(\bar{x})$ at $w = \tau_C(\widehat{w})\widehat{w}$, and an associated closed affine halfspace

$$\mathcal{S}_{\xi(w)} = \left\{ w + h;\ \langle h, \xi(w) \rangle_{\bar{x}} \leq 0 \right\}.$$

Since $\overline{\mathrm{I}(\bar{x})}$ is Lipschitz [21] but not convex, there is w_0 in the boundary of $\mathrm{I}(\bar{x})$ such that $w_0/|w_0|$ is a differentiability point of τ_C, and such that $\overline{\mathrm{I}(\bar{x})}$ is not included in $\mathcal{S}_{\xi(w_0)}$. In other words there is $w_1 \in \overline{\mathrm{I}(\bar{x})}$ such that

$$\langle w_1 - w_0, \xi(w_0) \rangle_{\bar{x}} > 0,$$

and necessarily $(1-t)w_0 + tw_1 \notin \overline{\mathrm{I}(\bar{x})}$ for $t > 0$ small enough.

Then there is a maximal $\widehat{t} \in (0, 1]$ such that $w_t = (1-t)w_0 + tw_1 \notin \overline{\mathrm{I}(\bar{x})}$ for all $t \in (0, \widehat{t})$. Since $\langle w_1 - w_0, \xi(w_0) \rangle_{\bar{x}} > 0$ and $w_t \notin \overline{\mathrm{I}(\bar{x})}$ for small positive times, there exists $\epsilon_0 > 0$ small enough such that $(1+\epsilon)w_0 - \epsilon w_1$ belongs to $\mathrm{I}(\bar{x})$ and

$$\left[(1+\epsilon)w_0 - \epsilon w_1, w_{\widehat{t}}\right] \cap \mathrm{TCL}(\overline{x}) = \left\{w_0, w_{\widehat{t}}\right\}.$$

for all $\epsilon \leq \epsilon_0$. Hence, if we choose $\epsilon, \epsilon' \leq \epsilon_0$ such that $|(1+\epsilon)w_0 - \epsilon w_1| \neq |(1-\epsilon')w_{\widehat{t}}|$, then $v_{-1} = (1+\epsilon)w_0 - \epsilon w_1$ and $v_1 = (1-\epsilon')w_{\widehat{t}}$ satisfy the desired conclusion. □

3. About the regularity condition

Before proving Proposition 4.1, we shall establish a lemma of independent interest: Under assumptions **(CI)** and **(MTW)**, the cost $c = d^2/2$ is "regular" in the sense of [31, Definition 12.14]. This means that for every $\bar{x} \in M$ and $v_0, v_1 \in \mathrm{I}(\bar{x})$ it holds not only the convexity of $\mathrm{I}(\bar{x})$, but also, for all $x \in M$ and $t \in [0,1]$,

$$c(x,y_t) - c(\bar{x},y_t) \geq \min\Big(c(x,y_0) - c(\bar{x},y_0),\, c(x,y_1) - c(\bar{x},y_1)\Big), \tag{3.1}$$

where $y_t = \exp_{\bar{x}}(v_t)$, $v_t = (1-t)v_0 + tv_1$.

This result was stated in [31, Theorem 12.36], but the proof was incomplete; we present a complete proof below.

Lemma 3.1. **(CI)** *and* **(MTW)** *imply the regularity of* $d^2/2$.

Before proving this lemma, let us discuss some implications. If $\psi : M \to \mathbb{R}$ is a c-convex function, and y_0, y_1 both belong to $\partial_c\psi(\bar{x})$, then there are $v_0, v_1 \in \overline{\mathrm{I}(\bar{x})}$ such that $y_0 = \exp_{\bar{x}}(v_0)$, $y_1 = \exp_{\bar{x}}(v_1)$, and

$$\psi(\bar{x}) + c(\bar{x},y_i) = \min_{x\in M}\Big\{\psi(x) + c(x,y_i)\Big\} \qquad \forall\, i = 0,1.$$

The latter property can be written as

$$c(x,y_i) - c(\bar{x},y_i) \geq \psi(\bar{x}) - \psi(x) \qquad \forall\, x \in M,\ i = 0,1,$$

which, thanks to Lemma 3.1, implies

$$c(x,y_t) - c(\bar{x},y_t) \geq \min\Big(c(x,y_0) - c(\bar{x},y_0),\, c(x,y_1) - c(\bar{x},y_1)\Big) \geq \psi(\bar{x}) - \psi(x),$$

for every $x \in M$, where $y_t = \exp_x((1-t)v_0 + tv_1)$, $t \in [0,1]$. This shows that the path $t \mapsto y_t$ belongs to $\partial_c\psi(\bar{x})$. In particular, *the c-subdifferential of an arbitrary c-convex function is always pathwise connected.*

Proof of Lemma 3.1. By continuity of c, we may assume $v_0, v_1 \in \mathrm{I}(\bar{x})$. Then the reasoning is based on a maximum principle similar to arguments already used in [14, 22, 31] or [15, Section 6], however with a few modifications. For $t \in [0,1]$, we define

$$h(t) := -c(x,y_t) + c(\bar{x},y_t) = -c(x,y_t) + \frac{1}{2}|v_t|^2_{\bar{x}} \qquad \forall\, t \in [0,1].$$

(Observe that $c(\bar{x},y_t) = |v_t|^2_{\bar{x}}/2$, since $v_t \in I(\bar{x})$.) Our aim is to show that h can only achieve a maximum at $t=0$ or $t=1$. By Lemma B.2, generically the curve $(y_t)_{0\leq t\leq 1}$ intersects $\mathrm{cut}(x)$ only on a finite set of times $0 = t_0 < t_1 < \ldots < t_N < t_{N+1} = 1$, always intersects $\mathrm{cut}(x)$ transversally, and never intersects $\mathrm{fcut}(x) =$

$\exp_x(\mathrm{TFCL}(x))$. So, as in [18], by an approximation argument we may assume that (y_t) satisfies all these assumptions.

This implies that h is smooth on the intervals (t_j, t_{j+1}) for $j = 0, \ldots, N$ and is not differentiable at $t = t_j$ for $j = 1, \ldots, N$. Let us observe that the function $t \mapsto -c(x, y_t)$ is semiconvex, being the curve $t \mapsto y_t$ smooth and the function $c(x, \cdot)$ semiconcave. Hence, since $t \mapsto c(\bar{x}, y_t)$ is smooth, $h(t)$ is semiconvex as well. Thus, at every non-differentiability point $t = t_j$ we necessarily have $\dot{h}(t_j^+) > \dot{h}(t_j^-)$, and so $h(t)$ cannot achieve a local maximum in a neighborhood of $t_1, \ldots, t_N$. In particular, there exists $\eta > 0$ such that $h(t)$ cannot achieve its maximum in any interval of the form $[t_j - 2\eta, t_j + 2\eta]$ with $j \in \{1, \ldots, N\}$.

Now let us show that h cannot have a maximum in any of the intervals $I_0 := (t_0, t_1 - \eta), I_N := (t_N + \eta, t_{N+1})$, and $I_j := (t_j + \eta, t_{j+1} - \eta)$ with $j \in \{1, \ldots, N-1\}$. Let $j \in \{0, \ldots, N\}$ be fixed. The function $y \mapsto c(x, y)$ is smooth in a neighborhood of the curve $(y_t)_{t \in I_j}$, so $q_t := -\nabla_y c(x, y_t)$ is well-defined for every $t \in I_j$. For every $t \in [0, 1]$, set $\bar{q}_t := -\nabla_y c(\bar{x}, y_t)$. Since v_0, v_1 belongs to $\mathrm{I}(\bar{x})$ which is an open convex set (by assumption **(CI)**), the tangent vector $\bar{q}_t$ always belong to $\mathrm{I}(y_t)$ and does not intersect $\mathrm{TFL}(y_t)$. Then, arguing as in [31, Proof of Theorem 12.36], there exists a constant $C > 0$ (depending on η) such that

$$\mathfrak{S}_{(y_t, sq_t + (1-s)\overline{q}_t)}(\dot{y}_t, q_t - \overline{q}_t) \geq -C\, |\langle \dot{y}_t, q_t - \overline{q}_t \rangle| \qquad \forall\, t \in I_j,\ \forall s \in [0,1].$$

Since (see [31, Proof of Theorem 12.36] or [14, Proof of Lemma 3.3])

$$\begin{cases} \dot{h}(t) &= \langle \dot{y}_t, q_t - \overline{q}_t \rangle \\ \ddot{h}(t) &= \dfrac{2}{3} \displaystyle\int_0^1 (1-s)\mathfrak{S}_{(y_t, sq_t + (1-s)\overline{q}_t)}(\dot{y}_t, q_t - \overline{q}_t)\, ds \end{cases}$$

we deduce

$$\ddot{h}(t) \geq -C|\dot{h}(t)| \qquad \forall\, t \in I_j. \tag{3.2}$$

Now, as in [31, Proof of Theorem 12.36] we consider the functions $h_\varepsilon(t) = h(t) + \varepsilon(t - 1/2)^k$, with k large enough (k will be chosen below). Assume by contradiction that h_ε attains a maximum at a time $t_0 \in I_j$ for some j. Then $\dot{h}_\varepsilon(t_0) = 0$ and $\ddot{h}_\varepsilon(t_0) \leq 0$, which implies

$$\dot{h}(t_0) = -\varepsilon k(t_0 - 1/2)^{k-1}, \qquad \ddot{h}(t_0) \leq -\varepsilon k(k-1)(t_0 - 1/2)^{k-2}.$$

This contradicts (3.2) for $k \geq 1 + C/2$. Moreover, since h_ε converges to h uniformly on $[0, 1]$ as $\varepsilon \to 0$, for ε sufficiently small the function h_ε cannot achieve its maximum on any interval of the form I_j. This implies that $h_\varepsilon(t) \le \max\{h_\varepsilon(0), h_\varepsilon(1)\}$ for ε small, and letting $\varepsilon \to 0$ we get (3.1) and the proof of Lemma 3.1 is complete. □

4. Sufficient conditions for (TCP)

Theorem 1.2(ii) follows from the following finer result.

Proposition 4.1. *Assume that M satisfies* **(CI$^+$)** *and* **(MTW$^+$)**. *Then the optimal map from μ to ν is continuous whenever μ and ν satisfy:*

(i) $\displaystyle\lim_{r\to 0^+} \frac{\mu(B_r(x))}{r^{n-1}} = 0$ *for any $x \in M$;*

(ii) $\nu \ge c_0 \,\mathrm{vol}$ *for some constant $c_0 > 0$.*

Proof of Proposition 4.1. The proof of Proposition 4.1 is based on an improvement of ingredients originating from [26] and further pushed in [14, Theorem 3.6] and [27, Theorem 5.1].

As we already mentioned, by [29] there exists a unique optimal transport map between μ and ν, given μ-almost everywhere by $T(x) = \exp_x\big(\nabla_x\psi\big)$, where ψ is a semiconvex function; moreover $\nabla_x\psi \in \overline{\mathrm{I}(x)}$ at all point of differentiability of ψ.

By assumption the sets $\mathrm{I}(x)$ are (strictly) convex for all x, therefore the subdifferential of ψ satisfies $\nabla^-\psi(x) \subset \overline{\mathrm{I}(x)}$ for all $x \in M$. We want to prove that ψ is C^1. To this aim, since ψ is semiconvex, by [2, Proposition 3.3.4(e)] it suffices to show that $\nabla^-\psi(x)$ is *everywhere* a singleton. We shall do this by contradiction, dividing the argument into several steps.

So assume that there is $\bar{x} \in M$ such that $v_0 \ne v_1 \in \nabla^-\psi(\bar{x})$. Let $y_0 = \exp_{\bar{x}}(v_0)$, $y_1 = \exp_{\bar{x}}(v_1)$. Then, by Lemma 3.1 and [31, Proposition 12.15(ii)], we get $y_i \in \partial_c\psi(\bar{x})$, i.e.

$$\psi(\bar{x}) + c(\bar{x}, y_i) = \min_{x\in M} \{\psi(x) + c(x, y_i)\}, \qquad i = 0, 1.$$

In particular

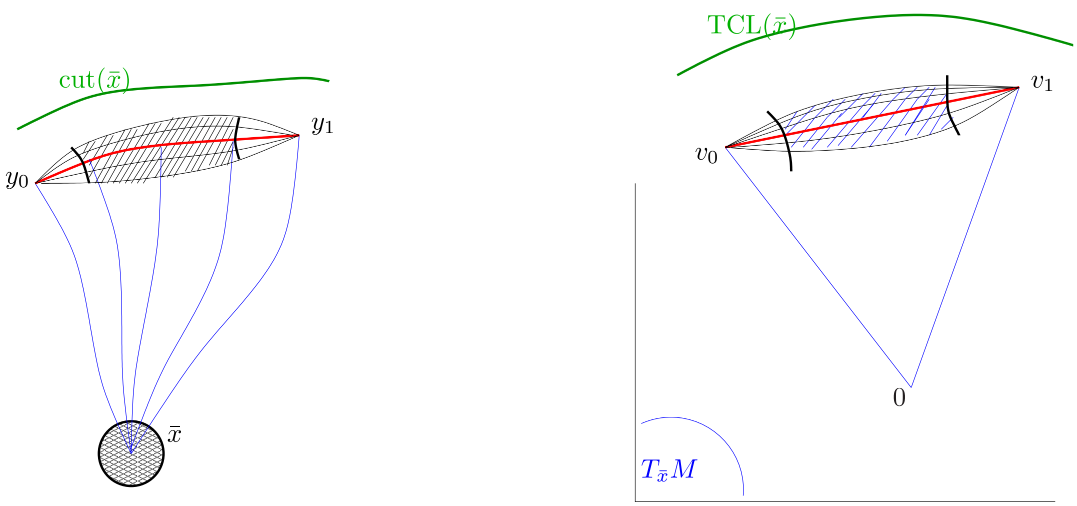
$$c(x, y_i) - c(\bar{x}, y_i) \ge \psi(\bar{x}) - \psi(x), \qquad \forall\, x \in M,\ i = 0, 1. \tag{4.1}$$

We shall show that this is impossible. For this we will perform a few estimates.

Step 1: For every $\delta > 0$ small, denote by $A_\delta \subset T_{\bar{x}}M$ the set swept by all the C^2 curves $t \in [0,1] \mapsto v_t \in T_{\bar{x}}M$ which satisfy $v_t = (1-t)v_0 + tv_1$ for any $t \in [0, 1/4] \cup [3/4, 1]$ and

$$|\ddot{v}_t|_{\bar{x}} \le \delta |\dot{y}_t|^2_{y_t} \quad \text{for } t \in [1/4, 3/4], \tag{4.2}$$

where $y_t = \exp_{\bar{x}}(v_t)$ (see figure).



Since the set $\overline{\mathrm{I}(\bar{x})}$ is strictly convex, the segment $I = [v_{1/4}, v_{3/4}]$ lies a positive distance away from $\mathrm{TCL}(\bar{x})$, and for every $y \in \exp_{\bar{x}}(I)$, the point $\bar{x}$ lies a positive distance away from $\mathrm{TCL}(y)$. Therefore, there is $\bar{\delta} > 0$ small enough such that $A_{\bar{\delta}} \subset I(\bar{x})$ and

$$\mathrm{conv}\Big(\exp_y^{-1}\big(\bar{B}_{\bar{\delta}}(\bar{x})\big)\Big) \subset \mathrm{I}(y) \qquad \forall\, y \in \exp_{\bar{x}}\big(A_{\bar{\delta}}\big). \tag{4.3}$$

By construction, any set A_δ (with $\delta \in (0, \bar{\delta})$, $\bar{\delta}$ small) contains a parallelepiped E_δ centered at $v_{1/2} = \frac{v_0+v_1}{2}$, with one side of length $\sim |v_0 - v_1|_{\bar{x}}$ and the other sides of length $\sim \delta |v_0 - v_1|^2_{\bar{x}}$, such that all points v in such a parallelepiped can be written as v_t for some $t \in [3/8, 5/8]$. Therefore, there is $c > 0$ such that $\mathcal{L}^n(E_\delta) \geq c\delta^{n-1}$, where $\mathcal{L}^n$ denotes the Lebesgue measure on $T_{\bar{x}}M$. Since E_δ lies a positive distance from $\mathrm{TCL}(\bar{x})$, we obtain

$$\mathrm{vol}\,(Y_\delta) \sim \mathcal{L}^n(E_\delta) \geq c\delta^{n-1}, \qquad \text{where} \qquad Y_\delta := \exp_{\bar{x}}(E_\delta). \tag{4.4}$$

On the other hand, by (4.3) the cost c is smooth on $\bar{B}_{\bar{\delta}}(\bar{x}) \times \exp_{\bar{x}}\big(A_{\bar{\delta}}\big)$ and moreover **(MTW$^+$)** holds. Hence, arguing as in [27, Lemma 2.3] we deduce that there are $K, C > 0$ such that the following property holds for any $y \in \exp_{\bar{x}}\big(A_{\bar{\delta}}\big)$ and any $x \in \mathcal{B} := \exp_y\Big(\mathrm{conv}\big(\exp_y^{-1}(\bar{B}_{\bar{\delta}})\big)\Big)$, where $\mathrm{conv}(S)$ denotes the convex envelope of a set S:

$$\forall\,(\xi,\eta) \in T_yM \times T_yM, \quad \mathfrak{S}_{(y,x)}(\xi,\eta) \geq K|\xi|_x^2|\eta|_y^2 - C\big|\langle \xi,\eta\rangle_y\big||\xi|_x|\eta|_y. \tag{4.5}$$

Step 2: Fix $f \in C_c^\infty([0,1])$ with $f \geq 0$ and $\{f > 0\} = (1/4, 3/4)$. We claim there exists $\lambda = \lambda(K, C, f) > 0$ such that for any $x \in \mathcal{B}$ and any C^2 curve $(v_t)_{0 \leq t \leq 1}$ drawn in $\mathrm{I}(\bar{x})$ satisfying

$$\begin{cases} |\ddot{v}_t|_{\bar{x}} = 0 \text{ for } t \in [0, 1/4] \cup [3/4, 1], \\ |\ddot{v}_t|_{\bar{x}} \leq \frac{K}{8} d(\bar{x}, x) |\dot{y}_t|^2_{y_t} \text{ for } t \in [1/4, 3/4], \\ v_t \in A_{\overline{\delta}} \text{ for any} t \in [1/4, 3/4], \end{cases} \tag{4.6}$$

where $y_t = \exp_x(v_t)$, for any $t \in [0,1]$ there holds

$$\begin{aligned} c(x, y_t) - c(\bar{x}, y_t) & \\ \geq \min\Big(c(x, y_0) - c(\bar{x}, y_0),\ c(x, y_1) - c(\bar{x}, y_1)\Big) + \lambda f(t) c(\bar{x}, x). \end{aligned} \tag{4.7}$$

To prove this we adapt the argument in Lemma 3.1, borrowing the strategy from [14, 27]: first of all, using Lemma B.2 (as we did in the proof of Lemma 3.1), up to slightly perturbing v_0 and v_1 we may assume that $v_0, v_1 \in \mathrm{I}(\bar{x})$, $(y_t)_{0 \leq t \leq 1}$ intersects $\mathrm{cut}(x)$ only at a finite set of times $0 = t_0 < t_1 < \ldots < t_N < t_{N+1} = 1$, and moreover $(y_t)_{0 \leq t \leq 1}$ never intersects $\mathrm{fcut}(x)$. Using the notation of Lemma 3.1, we consider the function $h : [0,1] \to \mathbb{R}$ given by

$$h(t) = -c(x, y_t) + c(\bar{x}, y_t) + \lambda d(\bar{x}, x)^2 f(t)/2 \qquad \forall\, t \in [0,1],$$

where $\lambda > 0$ is a positive constant to be chosen.

On the one hand, since $f(t) = 0$ and v_t is a segment for $t \in [0, 1/4] \cup [3/4, 1]$, arguing as in the proof of Lemma 3.1 we conclude that $h|_{[0,1/4] \cup [3/4,1]}$ achieves its maximum at $0, 1/4, 3/4$ or 1.

On the other hand, if $t \in [1/4, 3/4] \cap (t_j, t_{j+1})$ for some $j = 0, \ldots, N$, we can argue as in the proof of [27, Theorem 3.1] (see also [14, Lemma 3.3]) to check that the identity $\dot{h}(t) = 0$ gives $\ddot{h}(t) > 0$ for $\lambda = \lambda(K, C, f)$ sufficiently small. This implies that the function h cannot have any maximum on any interval $(t_j, t_{j+1}) \cap [1/4, 3/4]$. Moreover, since $\dot{h}(t_j^+) > \dot{h}(t_j^-)$ for $j = 1, \ldots, N$, h cannot achieve its maximum at any of the points t_j, $j = 1, \ldots, N$.

Hence h necessarily achieves its maximum at 0 or 1, and (4.7) follows.

Step 3: As a consequence of Step 2, for any $\delta \in (0, \bar{\delta})$, any $y \in Y_\delta$, and any $x \in B_{\bar{\delta}}(\bar{x}) \setminus B_{8\delta/K}(\bar{x})$, there holds

$$c(x, y) - c(\bar{x}, y) \geq \min\Big(c(x, y_0) - c(\bar{x}, y_0),\ c(x, y_1) - c(\bar{x}, y_1)\Big) + 2\lambda m_f d(\bar{x}, x)^2,$$

where $m_f = \inf\{f(t);\, t \in [3/8, 5/8]\} > 0$ and $\lambda = \lambda(K, C, f) > 0$. Combining this inequality with (4.1), we conclude that for any $\delta \in (0, \bar{\delta})$,

$$\forall\, y \in Y_\delta,\, \forall\, x \in B_{\bar{\delta}}(\bar{x}) \setminus B_{8\delta/K}(\bar{x}), \qquad y \notin \partial_c\psi(x). \tag{4.8}$$

We claim that taking $\delta \in (0, \bar{\delta})$ small enough, we may assume that the above property holds for any $x \in M \setminus B_{8\delta/K}(\bar{x})$. Indeed, if not, there exists a sequence $\{\delta_k\} \downarrow 0$, together with sequences $\{x_k\}$ in $M \setminus B_{\bar{\delta}}(\bar{x})$ and $\{y_k\} \in Y_{\delta_k}$, such that $y_k \in \partial_c\psi(x_k)$ for any k. By compactness, we deduce the existence of $x \in M \setminus B_{\bar{\delta}}(\bar{x})$ and $y_t \in \exp_{\bar{x}}([v_0, v_1])$, with $t \in [3/8, 5/8]$, such that $y_t \in \partial_c\psi(x)$. This implies that the c-concave potential $\psi^c : M \to \mathbb{R}$ defined as

$$\psi^c(y) := \min_{z \in M} \Big(\psi(z) + c(z, y)\Big) \tag{4.9}$$

satisfies $\bar{x}, x \in \partial^c\psi^c(y_t)$. Moreover, (4.8) gives that $\partial_c\psi^c(y_t) \cap B_{\bar{\delta}}(\bar{x}) \setminus B_{8\delta/K}(\bar{x}) = \emptyset$. However, from the discussion after Lemma 3.1 we know that the set $\partial_c\psi^c(y_t)$ is pathwise connected, absurd.

Step 4: Let us put the previous steps together. We have proved that for $\delta \in (0, \bar{\delta})$ small, all the mass brought into Y_δ by the optimal map comes from $B_{8\delta/K}(\bar{x})$, and so

$$\mu\big(B_{8\delta/K}(\bar{x})\big) \geq \nu\big(Y_\delta\big).$$

Thus, as $\mu\big(B_{8\delta/K}(\bar{x})\big) \leq o(1)\delta^{n-1}$ and $\nu(Y_\delta) \geq c_0 \mathrm{vol}\,(Y_\delta) \geq c'\delta^{n-1}$ (by assumption (ii) and (4.4)), we obtain a contradiction as $\delta \to 0$. □

5. The case of surfaces

Let us prove Theorem 1.3. Of course, in view of Theorem 1.2(i), we only need to prove the "if" part.

As in the proof of Proposition 4.1, let ψ be a c-convex function such that $T(x) = \exp_x\big(\nabla_x\psi\big)$. We want to prove that the subdifferential of ψ is a singleton everywhere.

We begin by observing that, thanks to Lemma 3.1 and [31, Proposition 12.15], we have

$$\nabla^-\psi(x) = \exp_x^{-1}\big(\partial_c\psi(x)\big) \qquad \forall\, x \in M. \tag{5.1}$$

First of all, we claim that the subdifferential of ψ at every point is at most one-dimensional. Indeed, if $\nabla^-\psi(x)$ is a two-dimensional convex C set for some $x \in M$, then by (5.1) $\exp_x(C) = \partial_c\psi(x)$ is a set with positive volume. But then, considering the optimal transport problem from ν to μ, the set $\partial_c\psi(x)$ is sent (by $\partial^c\psi^c$, with ψ^c as in (4.9)) onto the point x, which implies $\mu(\{x\}) \geq \nu(\partial_c\psi(x)) > 0$, impossible.

Now, assume by contradiction that ψ is not differentiable at some point x_0. Then there exist $v_{-1} \neq v_1 \in \overline{I(x_0)}$ such that $\nabla^- \psi(x_0) \subset \overline{I(x_0)}$ is equal to the segment $[v_{-1}, v_1] = \{v_t\}_{-1 \leq t \leq 1}$, $v_t = \frac{1+t}{2} v_1 + \frac{1-t}{2} v_{-1}$. We define $y_t = \exp_{x_0}(v_t)$.

We claim that that the following holds:

(A) $[v_{\varepsilon-1}, v_{1-\varepsilon}] \subset \mathrm{I}(x_0)$ for all $\varepsilon > 0$ (i.e. $v_t \notin \mathrm{TCL}(x_0)$ for all $t \in (-1, 1)$).

Since the proof of the above result is pretty involved, we postpone it to the end of this subsection.

Now the strategy is the following: by (A) we know that the cost function $c = d^2/2$ is smooth in a neighborhood of $\{x_0\} \times \{y_t;\, t \in [-3/4, 3/4]\}$ and satisfies all the assumptions of [13, Lemma 3.1] (observe that, even if that result is stated on domains of $\mathbb{R}^n$, everything is local so it also holds on manifolds), and we deduce that propagation of singularities holds. More precisely, [13, Lemma 3.1] gives the existence of a smooth injective curve $\gamma_{x_0} \ni x_0$, contained inside the set

$$\Gamma_{-3/4,3/4} = \left\{ d(\cdot, y_{-3/4})^2 - d(\cdot, y_{3/4})^2 = d(x_0, y_{-3/4})^2 - d(x_0, y_{3/4})^2 \right\},$$

such that

$$2\psi - 2\psi(x_0) = -d(\cdot, y_0)^2 + d(x_0, y_0)^2 \quad \text{on } \gamma_{x_0}. \tag{5.2}$$

(Recall that $c = d^2/2$.) Moreover, restricting γ_{x_0} if necessary, we can assume that $\gamma_{x_0} \cap \mathrm{cut}(y_0) = \emptyset$.

We now observe that, since $y_t \in \exp_{x_0}([v_{-1}, v_1]) = \partial_c \psi(x_0)$, we have

$$2\psi - 2\psi(x_0) \geq -d(\cdot, y_t)^2 + d(x_0, y_t)^2 \qquad \forall\, t \in [-1, 1]. \tag{5.3}$$

Moreover, by Lemma 3.1,

$$-d(\cdot, y_0)^2 + d(x_0, y_0)^2 \leq \max\Big(-d(\cdot, y_t)^2 + d(x_0, y_t)^2, -d(\cdot, y_\tau)^2 + d(x_0, y_\tau)^2 \Big) \tag{5.4}$$

for all $-1 \leq t \leq 0 \leq \tau \leq 1$. Hence, combining (5.2), (5.3), and (5.4), we deduce that

$$2\psi - 2\psi(x_0) = -d(\cdot, y_t)^2 + d(x_0, y_t)^2 \quad \text{on } \gamma_{x_0} \qquad \forall\, t \in [-1, 1] \tag{5.5}$$

and

$$\gamma_{x_0} \subset \bigcap_{-1 < t < \tau < 1} \Gamma_{t,\tau},$$

where

$$\Gamma_{t,\tau} = \left\{ d(\cdot, y_t)^2 - d(\cdot, y_\tau)^2 = d(x_0, y_t)^2 - d(x_0, y_\tau)^2 \right\}. \tag{5.6}$$

By (5.5) we deduce that that $y_t \in \partial_c \psi(x)$ for any $x \in \gamma_{x_0}$. Moreover, if we parameterize γ_{x_0} as $s \mapsto x_s$, by differentiating with respect to s the identity

$$d(x_s, y_t)^2 - d(x_s, y_0)^2 = d(x_0, y_t)^2 - d(x_0, y_0)^2$$

we obtain, for all s, t ,

$$\dot{x}_s \cdot \left[p - \nabla_x d(x_s, y_0)^2\right] = 0 \qquad \forall\, p \in \nabla_x^+ d(x_s, y_t)^2$$

(recall that $\gamma_{x_0} \cap \mathrm{cut}(y_0) = \emptyset$). This implies that, for any fixed s, there exists a segment which contains all elements in the superdifferential of $d(\cdot, y_t)^2$ at x_s for all $t \in [-1, 1]$. Hence, by the convexity of $I(x_s)$, we can apply (A) with x_0 replaced by x_s: we can repeat the argument above starting from any point x_s, and by a topological argument as in [13, Proof of Lemma 3.1] (showing that the maximal time interval on which we can extend the curve is both open and closed) we immediately get that γ_{x_0} is a simple curve which either is closed or has infinite length.

To summarize, we finally have the following geometric picture: there exists a smooth closed curve $\gamma \subset M$, which is either closed or has infinite length, such that:

(A-a) For any $x \in \gamma$, $y_t \notin \mathrm{cut}(x)$ for all $t \in (-1, 1)$.
(A-b) $\gamma \subset \bigcap_{-1<t<\tau<1} \Gamma_{t,\tau}$, where $\Gamma_{t,\tau} = \{d(\cdot, y_t)^2 - d(\cdot, y_\tau)^2 = d(x_0, y_t)^2 - d(x_0, y_\tau)^2\}$.

Let us show that the compactness of M prevents this.

By differentiating with respect to t at $t = 0$ the identity

$$d(x, y_t)^2 - d(x, y_0)^2 = d(x_0, y_t)^2 - d(x_0, y_0)^2 \qquad \forall\, x \in \gamma,$$

we obtain (using (A-a))

$$\left[\nabla_y d(x, y_0)^2 - \nabla_y d(x_0, y_0)^2\right] \cdot \dot{y}_0 = 0 \qquad \forall\, x \in \gamma.$$

Observe that, since $y_0 \notin \mathrm{cut}(x)$, $\dot{y}_0 \neq 0$. Hence there exists a segment $\Sigma \subset \overline{I(y_0)}$ such that $\exp_{y_0}^{-1}(\gamma) \subset \Sigma$. This is impossible since $\gamma \subset M \setminus \mathrm{cut}(y_0)$, so $\exp_{y_0}^{-1}(\gamma)$ is either closed or it has infinite length.

This concludes the proof of the C^1 regularity of ψ. It remains to show the validity of property (A) above.

Proof of (A). To prove (A), we distinguish two cases:

(i) Either v_{-1} or v_1 does not belong to $\mathrm{TCL}(x_0)$.
(ii) Both v_{-1} and v_1 belong to $\mathrm{TCL}(x_0)$.

In case (i), by the convexity of $\mathrm{I}(x)$ we immediately get that either $[v_{-1}, v_{1-\varepsilon}] \subset \mathrm{I}(x)$ or $[v_{\varepsilon-1}, v_1] \subset \mathrm{I}(x)$ for any $\varepsilon > 0$, so (A) holds.

In case (ii), assume by contradiction that (A) is false. By the convexity of $\mathrm{I}(x_0)$ we have $[v_{-1}, v_1] \subset \mathrm{TCL}(x_0)$. Let $\bar{t} \in [-1, 1]$ be such that $|v_{\bar{t}}|_{x_0}$ is minimal on $[v_{-1}, v_1]$

(by uniform convexity of the norm, there exists a unique such point). We consider two cases:

(ii-a) $\exp_{x_0}^{-1}(y_{\bar{t}})$ is not a singleton.
(ii-b) $\exp_{x_0}^{-1}(y_{\bar{t}}) = v_{\bar{t}}$.

In case (ii-a), since $|v_{\bar{t}}|_{x_0}$ is the unique vector of minimal norm on $[v_{-1}, v_1]$, there exists $\bar{v} \in \mathrm{TCL}(x_0) \setminus [v_{-1}, v_1]$ such that $\exp_{x_0}(\bar{v}) = \exp_{x_0}(v_{\bar{t}}) = y_{\bar{t}}$ ($\bar{v}$ cannot belong to $[v_{-1}, v_1]$ since $|\bar{v}|_{x_0} = |v_{\bar{t}}|_{x_0}$). However this is impossible since (5.1) implies $\bar{v} \in \exp_{x_0}^{-1}\big(\partial_c \psi(x_0)\big) \subset \nabla^- \psi(x_0) = [v_{-1}, v_1]$.

In case (ii-b), without loss of generality we assume that the metric at x_0 concides with the identity matrix. Let us recall that by [16, Proposition A.6] the function $w \in U_{x_0}M \mapsto t_F(x_0, w)$ has vanishing derivative at all w such that $t_F(x, w)w \in \mathrm{TFCL}(x_0)$. Hence, since the derivative of $t \mapsto |v_t|_{x_0}$ is different from 0 at every $t \neq \bar{t}$ and $|v_t|_{x_0} \leq t_F(x_0, v_t)$ for every t, we deduce that $v_t \in \mathrm{TCL}(x_0) \setminus \mathrm{TFL}(x_0)$ for all $t \neq \bar{t}$. Let us choose any time $s \neq \bar{t}$. Since $v_s \in \mathrm{TCL}(x_0) \setminus \mathrm{TFL}(x_0)$, there exists a vector $v' \neq v_s$ such that $\exp_{x_0}(v') = \exp_{x_0}(v_s) \in \partial_c \psi(x_0)$. By (5.1) this implies $v' \in \nabla^- \psi(x_0) = [v_{-1}, v_1]$. Hence there exists a time $s' \neq s$ such that $v' = v_{s'}$. Moreover, since $|v_s|_{x_0} = |v_{s'}|_{x_0} > |v_{\bar{t}}|_{x_0}$, we get $|s - \bar{t}| = |s' - \bar{t}|$. Thus, by the arbitrariness of $s \in [-1, 1] \setminus \{\bar{t}\}$ we easily deduce that the only possibility is $\bar{t} = 0$, and so $y_t = y_{-t}$ for all $t \in [-1, 1]$.

By performing a change of coordinates in a neighborhood of the minimizing geodesic γ_0 going from x_0 to y_0, we can assume that $x_0 = (0, 0), y_0 = (1, 0)$, $v_0 = (1, 0)$, $[v_{-1}, v_1] = \big[(1, -1), (1, 1)\big]$, that the metric g at x_0 and y_0 is the identity matrix I_2, and that the geodesic starting from x_0 with initial velocity v_0 is given by $\gamma(t) = (t, 0)$. Now, to simplify the computation, we slightly change the definition of v_δ and y_δ for $\delta > 0$ small (this should not create confusion, since we will adopt the following notation in all the sequel of the proof): denote by v_δ the speed which belongs to the segment $[v_{-1}, v_1]$ and whose angle with the horizontal axis is δ, that is

$$v_\delta = \big(1, \tan\delta\big), \quad \bar{v}_\delta := \frac{v_\delta}{\big|v_\delta\big|} = \big(\cos\delta, \sin\delta\big), \quad t_\delta := \big|v_\delta\big| = \frac{1}{\cos\delta}.$$

Consider γ_δ the geodesic starting from x_0 with initial velocity v_δ, and set

$$y_\delta := \exp_{x_0}(v_\delta), \quad w_\delta := -\dot{\gamma}_\delta(1), \quad \text{and} \quad \bar{w}_\delta := \frac{w_\delta}{\big|w_\delta\big|}.$$

The geodesic flow sends $\big(x_0, v_\delta\big)$ to $\big(y_\delta, -w_\delta\big)$, and the linearization at $\delta = 0$ gives

$$\dot{y}_0 = 0 \quad \text{and} \quad -\dot{w}_0 = \big(0, \dot{f}_0(1)\big), \tag{5.7}$$

where f_0 denotes the solution (starting with $f_0(0) = 0, \dot{f}_0(0) = 1$) to the Jacobi equation associated with the geodesic starting from x_0 with initial velocity v_0. The curve $\delta \mapsto y_\delta$ is a smooth curve valued in a neighborhood of y_0. Moreover, since $y_\delta = y_{-\delta}$ for any small δ, y_δ has the form

$$y_\delta = y_0 + \frac{\delta^2}{2} Y + o(\delta^2)$$

for some vector Y. We now observe that, for every $\delta > 0$, the vector $\dot{y}_\delta$ satisfies (because the distance function to x_0 is semiconcave and y_δ is contained in the cut locus of x_0)

$$\langle \dot{y}_\delta, w_\delta \rangle = \langle \dot{y}_\delta, w_{-\delta} \rangle,$$

which can be written as

$$\left\langle \frac{\dot{y}_\delta}{|y_\delta|}, w_\delta - w_{-\delta} \right\rangle = 0.$$

Thanks to (5.7), we deduce that y_δ takes the form

$$y_\delta = y_0 + \frac{\delta^2}{2} (\lambda, 0) + o(\delta^2) \tag{5.8}$$

for some $\lambda \geq 0$.

We now need some notation. For every nonzero tangent vector v at x_0, we denote by $f_0(\cdot, v), f_1(\cdot, v)$ the solutions to the Jacobi equation

$$\ddot{f}(t) + k(t) f(t) = 0 \qquad \forall\, t \geq 0, \tag{5.9}$$

along the geodesic starting from x_0 with unit initial velocity $v/|v|$ which satisfies

$$f_0(0, v) = 0, \quad \dot{f}_0(0, v) = 1, \quad f_1(0, v) = 1, \quad \dot{f}_1(0, v) = 0 \tag{5.10}$$

(in (5.9), $k(t)$ denotes the Gauss curvature along such geodesic). Set now

$$\bar{v}_\delta^\perp := \big(-\sin \delta, \cos \delta\big).$$

Since

$$\dot{v}_\delta = \left(0, \frac{1}{\cos^2 \delta}\right) = \frac{\sin \delta}{\cos^2 \delta}\, \bar{v}_\delta + \frac{1}{\cos \delta}\, \bar{v}_\delta^\perp,$$

we have

$$\dot{y}_\delta = \frac{\sin \delta}{\cos^2 \delta} \big(-\bar{w}_\delta\big) + f_0\big(t_\delta, v_\delta\big) \frac{1}{\cos \delta} \big(-\bar{w}_\delta^\perp\big). \tag{5.11}$$

Then, since $\bar{w}_\delta = (-1, 0) + O(\delta)$, this means that

$$\dot{y}_\delta = \delta(1, 0) + O(\delta^2),$$

which implies that the constant λ appearing in (5.8) is equal to 1. Hence $\ddot{y}_0 = (1,0)$, and we get

$$y_\delta = y_0 + \frac{\delta^2}{2}(1,0) + O(\delta^4), \tag{5.12}$$

because $y_\delta = y_{-\delta}$.

Define the curve $\delta \mapsto z_\delta$ by

$$z_\delta := \exp_{x_0}(u_\delta) \quad \text{with} \quad u_\delta := \tau_\delta \bar{v}_\delta = \big(\tau_\delta \cos\delta, \tau_\delta \sin\delta\big), \quad \tau_\delta := t_F(x_0, \bar{v}_\delta).$$

We now use a result from [16]: since **(MTW)** holds, the curvature of $\mathrm{TFL}(x_0)$ near any point of $\mathrm{TFCL}(x_0)$ has to be nonnegative, see [16, Proposition 4.1(ii)]. Since $[v_{-1}, v_1] \subset \overline{\mathrm{I}(x_0)} \subset \overline{\mathrm{NF}(x_0)}$ and $\mathrm{TFL}(x_0)$ is a smooth curve, the fact that the curvature $\mathrm{TFL}(x_0)$ at v_0 is nonnegative implies that $\mathrm{TFL}(x_0)$ separates at v_0 from $[v_{-1}, v_1]$ (and so also from $\mathrm{TCL}(x_0)$) at most at the forth order, so $\tau_\delta - t_\delta = O(\delta^4)$. Since the exponential map is smooth (so in particular Lipschitz), this also gives

$$\big|y_\delta - z_\delta\big| = O(\delta^4). \tag{5.13}$$

Denote by $\bar{a}_\delta$ the (unit) vector at time $t = \tau_\delta$ of the geodesic starting at x_0 with initial velocity $\bar{v}_\delta$. As for $\dot{y}_\delta$, we can express $\dot{z}_\delta$ in terms of $a_\delta, a_\delta^\perp$, and $f_0\big(\tau_\delta, v_\delta\big) = 0$. For that, we note that

$$\begin{aligned} \dot{u}_\delta &= \big(\dot{\tau}_\delta \cos\delta - \tau_\delta \sin\delta, \dot{\tau}_\delta \sin\delta + \tau_\delta \cos\delta\big) \\ &= \dot{\tau}_\delta\big(\cos\delta, \sin\delta\big) + \tau_\delta\big(-\sin\delta, \cos\delta\big) \\ &= \dot{\tau}_\delta \bar{v}_\delta + \tau_\delta \bar{v}_\delta^\perp, \end{aligned}$$

from which we deduce that

$$\dot{z}_\delta = \dot{\tau}_\delta\, \bar{a}_\delta + f_0\big(\tau_\delta, v_\delta\big)\tau_\delta \left(\bar{a}_\delta^\perp\right) = \dot{\tau}_\delta\, \bar{a}_\delta. \tag{5.14}$$

This gives

$$\ddot{z}_\delta = \ddot{\tau}_\delta\, \bar{a}_\delta + \dot{\tau}_\delta\, \dot{\bar{a}}_\delta$$

and

$$\dddot{z}_\delta = \dddot{\tau}_\delta\, \bar{a}_\delta + 2\ddot{\tau}_\delta\, \dot{\bar{a}}_\delta + \dot{\tau}_\delta\, \ddot{\bar{a}}_\delta.$$

Moreover, since $\tau_\delta - \frac{1}{\cos(\delta)} = \tau_\delta - t_\delta = O(\delta^4)$, we have

$$\dot{\tau}_0 = \dot{t}_0 = 0, \quad \ddot{\tau}_0 = \ddot{t}_0 = 1, \quad \dddot{\tau}_0 = \dddot{t}_0 = 0.$$

Hence we obtain

$$\dot{z}_\delta = 0, \quad \ddot{z}_\delta = \bar{a}_0 = -w_0, \quad \dddot{z}_\delta = 2\,\dot{\bar{a}}_0 (= -\dot{w}_0 \neq 0),$$

which yields

$$z_\delta = y_0 + \frac{\delta^2}{2}(1,0) + \frac{\delta^3}{3}\,\dot{\bar{a}}_0 + o(\delta^3), \quad \dot{\bar{a}}_0 \neq 0.$$

This contradicts (5.12) and (5.13), and concludes the proof of (A). □

6. Further sufficient conditions for regularity

In this section we present various sufficient conditions which, while not as sharp as the ones discussed previously, may be easier to check.

6.1. **Extended MTW conditions.** Extended MTW conditions may be more restrictive than the plain MTW conditions, but allow to replace the conditions on injectivity domains by conditions on nonfocal domains, which are easier to apprehend.

For every $x \in M$, let us denote by $\mathrm{NF}(x) \subset T_xM$ the *nonfocal domain* at x (see Appendix A). We shall say that (M,g) satisfies **(CNF)** (resp. **(CNF$^+$)**) if $\mathrm{NF}(x)$ is convex (resp. strictly convex) for all $x \in M$. As first suggested in [14], the MTW tensor may be extended by letting v vary in the whole nonfocal domain rather than in the injectivity domain. We refer for instance to [15, Subsection 2.3] for the construction of this generalized tensor $\overline{\mathfrak{S}}_{(x,v)}(\xi,\eta)$, where $v \in \mathrm{NF}(x)$. Then it is said that (M,g) satisfies the *extended Ma–Trudinger–Wang condition* **($\overline{\mathbf{MTW}}$)** if

$$(6.1) \quad \forall\,(x,v) \in TM \text{ with } v \in \mathrm{NF}(x),\ \forall\,(\xi,\eta) \in T_xM \times T_xM, \quad \Big[\langle \xi,\eta\rangle_x = 0 \implies \overline{\mathfrak{S}}_{(x,v)}(\xi,\eta) \geq 0\Big].$$

As before, if the last inequality in (6.1) is strict unless $\xi = 0$ or $\eta = 0$, then M is said to satisfy the *extended strict Ma–Trudinger–Wang condition* **($\overline{\mathbf{MTW}}^+$)**. Note that **($\overline{\mathbf{MTW}}$)** implies **(MTW)** and **($\overline{\mathbf{MTW}}^+$)** implies **(MTW$^+$)**.

Proposition 6.1. *Let (M,g) be a smooth compact connected Riemannian manifold of dimension $n \geq 2$. If M satisfies* **(CNF$^+$)** *and* **($\overline{\mathbf{MTW}}^+$)***, then it also satisfies* **(CI$^+$)***, and as a consequence* **(TCP)***.*

Proposition 6.2. *Let (M,g) be a smooth compact Riemannian surface. If M satisfies* **(CNF)** *and* **($\overline{\mathbf{MTW}}$)***, then it also satisfies* **(CI)***, and as a consequence* **(TCP)***.*

Proof of Proposition 6.1. By Theorem 1.2(ii), it suffices to show that (M,g) satisfies **(CI$^+$)**. We argue by contradiction. If **(CI$^+$)** fails, we can find $v_0 \neq v_1 \in \overline{\mathrm{I}(\bar{x})}$ such that $v_t = (1-t)v_0 + tv_1 \notin \mathrm{I}(\bar{x})$ for all $t \in (0,1)$. Then, since $\mathrm{NF}(\bar{x})$ is strictly convex, $v_t \in \mathrm{NF}(\bar{x})$ for all $t \in (0,1)$.

For any $t \in (0,1)$, set $y_t = \exp_{\bar{x}}(v_t)$ and $\bar{q}_t = -d_{v_t}\exp_{\bar{x}}(v_t)$ as in the proof of Lemma 3.1. Then there exists $q_t \in \overline{\mathrm{I}(y_t)}$ with $q_t \neq \bar{q}_t$ such that $\exp_{y_t}(q_t) = \exp_{y_t}(\bar{q}_t) = \bar{x}$. We now choose a sequence of points $\{x_k\} \to \bar{x}$ such that $y_t \notin \mathrm{cut}(x_k)$ and $-\nabla_y c(x_k, y_t) \to q_t$ for all $t \in [0,1]$ (see for instance [27] for such a construction). By repeating the proof of Lemma 3.1 with the smooth function $h_k(t) = -c(x_k, y_t) + |v_t|^2_{\bar{x}}/2$ over the time interval $[0,1]$ (see [14]), one can see that $\ddot{h}_k(t)$ is given by

$$\ddot{h}_k(t) = \frac{2}{3}\int_0^1 (1-s)\overline{\mathfrak{S}}_{(y_t,(1-s)\bar{q}_t - s\nabla_y c(x_k,y_t))}\big(\dot{y}_t, \bar{q}_t - [-\nabla_y c(x_k,y_t)]\big)\,ds,$$

which by $(\overline{\mathbf{MTW}}^+)$ is strictly positive whenever $\dot{h}_k(t) = \langle \dot{y}_t, \bar{q}_t - [-\nabla_y c(x_k,y_t)]\rangle_{y_t} = 0$. As in the proof Lemma 3.1, these facts imply easily that, for any $t \in (0,1)$,

$$\begin{aligned} d(x_k,y_t)^2 - |v_t|^2_{\bar{x}} &= 2h_k(t) \geq 2\min\{h_k(0), h_k(1)\} + r(t) \\ &= \min\Big(d(x_k,y_0)^2 - d(\bar{x},y_0)^2,\, d(x_k,y_1)^2 - d(\bar{x},y_1)^2\Big) + r(t), \end{aligned}$$

where $r : [0,1] \mapsto [0,1]$ is a continuous function (independent of k) such that $r > 0$ on $[1/4, 3/4]$. Hence, choosing for instance $t = 1/2$ and letting $k \to \infty$ we get

$$\begin{aligned} 0 = d(\bar{x},y_{1/2})^2 - d(\bar{x},y_{1/2})^2 &\geq d(\bar{x},y_{1/2})^2 - |v_{1/2}|^2_{\bar{x}} \\ &> \min\Big(d(\bar{x},y_0)^2 - d(\bar{x},y_0)^2,\, d(\bar{x},y_1)^2 - d(\bar{x},y_1)^2\Big) = 0, \end{aligned}$$

a contradiction. □

To prove Proposition 6.2, it suffices to establish the following "extended version" of Lemma 3.1, which is established along the same lines as in [14].

Lemma 6.3. *Let (M,g) be a Riemannian manifold satisfying* **(CNF)-($\overline{\mathbf{MTW}}$)**. *Fix $\bar{x} \in M$, $v_0, v_1 \in \overline{\mathrm{I}(\bar{x})}$, and let $v_t = (1-t)v_0 + tv_1 \in T_xM$. For any $t \in [0,1]$, set $y_t = \exp_x(v_t)$. Then, for any $x \in M$, for any $t \in [0,1]$,*

$$c(x,y_t) - \frac{1}{2}|v_t|^2_{\bar{x}} \geq \min\Big(c(x,y_0) - c(\bar{x},y_0),\, c(x,y_1) - c(\bar{x},y_1)\Big).$$

By choosing $x = \bar{x}$ we deduce that $c(x,y_t) \geq \frac{1}{2}|v_t|^2_{\bar{x}},$, which implies that $v_t \in \overline{\mathrm{I}(\bar{x})}$, as desired.

6.2. **Removing the orthogonality condition in MTW.** We say that (M, g) satisfies **(MTW$^{\not\perp}$)** if (1.1) holds without any orthogonality assumption, that is,

$$\forall\,(x,v) \in TM \text{ with } v \in \mathrm{I}(x),\ \forall\,(\xi,\eta) \in T_xM \times T_xM, \quad \mathfrak{S}_{(x,\exp_x v)}(\xi,\eta) \geq 0.$$

If $T = \exp \nabla\psi$ is an optimal transport map, let us define the contact set of $y \in M$ as

$$S(y) = \{x \in M;\ y \in \partial_c\psi(x)\} = \partial^c\psi^c(y),$$

where ψ^c is defined in (4.9). From (5.1) in Section 5 it follows that if (M, g) satisfies **(CI)** and **(MTW)**, then the equality $\exp_y^{-1}(\partial^c\psi^c(y)) = \nabla^-\psi^c(y)$ holds for any $y \in M$. In particular, $\exp_y^{-1}(S(y)) \subset \overline{\mathrm{I}(y)}$ is always a convex set.

Let us recall that, according to [27], a manifold is said to have nonfocal cut locus if $\mathrm{TFCL}(x) = \emptyset$ for all $x \in M$. As a consequence of the results in [11] and [5, Appendix C, Theorem 3], we obtain the following proposition, which generalizes the regularity result on the flat torus $\mathbb{T}^n$ [4] and the real projective space.

Proposition 6.4. *Let (M, g) be a compact quotient of $\mathbb{S}^{n_1}_{r_1} \times \ldots \times \mathbb{S}^{n_i}_{r_i} \times \mathbb{R}^n$ with nonfocal cut locus. Then* (**TCP**) *holds. Moreover, if $\mu = f\mathrm{vol}$ and $\nu = g\mathrm{vol}$ with $f, g > 0$ and of class C^∞, then the optimal transport is C^∞ too.*

Actually, as can easily be seen from the proof of the above result, by the results in [11] the following general statement (already present in the proof of [12, Corollary 5.2]) holds:

Theorem 6.5. *Assume that (M, g) satisfies* **(CI)** *and* **(MTW$^{\not\perp}$)***, and suppose that there exist two positive constants $\lambda, \Lambda > 0$ such that*

$$\lambda\mathrm{vol} \leq \mu \leq \Lambda\mathrm{vol}\,, \qquad \lambda\mathrm{vol} \leq \nu \leq \Lambda\mathrm{vol}\,.$$

Then, for any $y \in M$, either $S(y)$ is a singleton or all exposed points of $\exp_y^{-1}(S(y))$ *belong to* $\mathrm{TCL}(y)$.

We recall that a point z in the boundary of a compact convex set $Z \subset \mathbb{R}^n$ is said to be an exposed point of Z if there exists a hyperplane $\Pi \subset \mathbb{R}^n$ such that $Z \cap \Pi = \{z\}$.

Proof of Proposition 6.4. Let μ to ν be two probability measures such that

$$\lambda\mathrm{vol} \leq \mu \leq \Lambda\mathrm{vol}\,, \qquad \lambda\mathrm{vol} \leq \nu \leq \Lambda\mathrm{vol}$$

for two positive constants $\lambda, \Lambda > 0$. Let $T = \exp_x\big(\nabla_x\psi\big) : M \to M$ denote the transport map from μ to ν, and define the "c-Monge-Ampère" measure $|\partial^c\psi|$ as

$$|\partial^c\psi|(A) = \mathrm{vol}\left(\cup_{x\in A}\partial^c\psi(x)\right) \qquad \text{for all } A \subset M \text{ Borel.}$$

As shown for instance in [11, Lemma 3.1] (see also [28]), under our assumptions on μ and ν the following upper and lower bounds on $|\partial^c\psi|$ hold:

$$\frac{\lambda}{\Lambda}\,\mathrm{vol}\,(A) \le |\partial^c\psi|(A) \le \frac{\Lambda}{\lambda}\,\mathrm{vol}\,(A) \qquad \text{for all } A\subset M \text{ Borel.} \tag{6.2}$$

Consider now the cost function $\widetilde{c} = \widetilde{d}^2/2$ on $\widetilde{M} = \mathbb{S}^{n_1}_{r_1}\times\ldots\times\mathbb{S}^{n_i}_{r_i}\times\mathbb{R}^n$, with $\widetilde{d}$ the Riemannian distance on $\widetilde{M}$. By [23], the cost $\widetilde{c}$ satisfies $\mathbf{MTW}^{\not\perp}$. Moreover, if $\widetilde{\pi} : \widetilde{M} \to M$ denotes the quotient map, we can use $(\widetilde{\pi})^{-1}$ to lift ψ: we define $\widetilde{\psi} : \widetilde{M} \to \mathbb{R}$ by $\widetilde{\psi} = \psi\circ\widetilde{\pi}$. It is not difficult to check that $\widetilde{\psi}$ is $\widetilde{c}$-convex. Moreover, thanks to (6.2), we have

$$\frac{\lambda}{\Lambda}\,\widetilde{\mathrm{vol}}\,(A) \le |\partial^{\widetilde{c}}\widetilde{\psi}|(A) \le \frac{\Lambda}{\lambda}\,\widetilde{\mathrm{vol}}\,(A) \qquad \text{for all } A\subset \widetilde{M} \text{ Borel,}$$

where $\widetilde{\mathrm{vol}}$ is the volume measure on $\widetilde{M}$. Finally, since set of subgradients $\nabla^-\psi(x)$ at any point x belongs to $\overline{I(x)} \subset T_xM$ (see for instance [29]), by identifying the tangent spaces $T_{\widetilde{x}}M$ and $T_{\widetilde{\pi}(\widetilde{x})}M$ we obtain

$$\nabla^-\widetilde{\psi}(\widetilde{x}) = \nabla^-\psi\big(\widetilde{\pi}(\widetilde{x})\big) \subset \overline{I(\widetilde{\pi}(\widetilde{x}))}.$$

However, since $\widetilde{M}$ is a product of spheres and $\mathbb{R}^n$, thanks to the nonfocality assumption on M it is easily seen that

$$\mathrm{TCL}(\widetilde{\pi}(\widetilde{x})) \subset\subset \mathrm{I}(\widetilde{x})$$

(again we are identifying $T_{\widetilde{x}}M$ with $T_{\widetilde{\pi}(\widetilde{x})}M$). This implies that $\nabla^-\widetilde{\psi}(\widetilde{x})$ lies at a positive distance from $\mathrm{TCL}(\widetilde{x})$ for every $\widetilde{x}\in\widetilde{M}$. In particular, for every $\widetilde{y}\in\widetilde{M}$, the set

$$\widetilde{S}(\widetilde{y}) = \Big\{\widetilde{x}\in\widetilde{M};\ \widetilde{y}\in\partial^{\widetilde{c}}\widetilde{\psi}(\widetilde{x})\Big\}$$

cannot intersect $\mathrm{cut}(\widetilde{y})$.

Now, let us consider the change of coordinates $\widetilde{x}\mapsto\widetilde{q} = -\nabla_{\widetilde{x}}\widetilde{c}(\widetilde{x},\widetilde{y})$ which sends $\widetilde{M}\setminus\mathrm{cut}(\widetilde{y})$ onto $\mathrm{I}(\widetilde{y})$. Since $\widetilde{x}_0 = \exp_{\widetilde{y}}(\widetilde{q}_0) \notin \mathrm{cut}(\widetilde{y})$, the cost $\widetilde{c}$ is smooth in a neighborhood of $\{\widetilde{x}_0\}\times\{\widetilde{y}\}$. Moreover, the support of $|\partial^{\widetilde{c}}\widetilde{\phi}|$ is the whole manifold $\widetilde{M}$. So, we can apply [11, Theorem 8.1 and Remark 8.2] to obtain that $\widetilde{S}(\widetilde{y})$ has no exposed points inside the open set $\mathrm{I}(\widetilde{y})$. Since $\widetilde{S}(\widetilde{y})$ cannot intersect $\mathrm{cut}(\widetilde{y})$ either, the only possibility is that $\widetilde{S}(\widetilde{y})$ is a singleton for every $\widetilde{y}$, so $\widetilde{\psi}^{\widetilde{c}}$ (the $\widetilde{c}$-transform of $\widetilde{\psi}$) is C^1. This implies that also ψ^c (the c-transform of ψ) is C^1 (indeed, one can easily show that $\widetilde{\psi}^{\widetilde{c}} = \psi^c\circ\widetilde{\pi}$), so the transport map T is injective.

Now, we observe that the same argument can be repeated considering the transport problem from ν to μ (recall that $\exp_y\big(\nabla_y \psi^c\big)$ is the optimal map sending ν onto μ, see for instance [31]). So, we deduce that $\widetilde{\psi}$ is C^1 as well, which implies that ψ is C^1, and T is continuous as desired.

Let us finally remark that, as already observed in [12], the continuity and injectivity of T combined with the results in [25] implies higher regularity ($C^{1,\alpha}/C^\infty$) of optimal maps for more smooth (C^α/C^∞) densities. This concludes the proof. □

7. Open problems

By Theorem 1.2(i), any Riemannian manifold verifying (**TCP**) must satisfy **(CI)** and **(MTW)**. As shown by Theorem 1.3, the combination **(CI)**-**(MTW)** and (**TCP**) are equivalent on surfaces. We do not know if such a result holds in higher dimension.

In [6] the authors showed that small C^4 perturbations of $\mathbb{S}^2$ equipped with the round metric satisfy **(MTW$^{\not\perp}$)**. Thanks to [23, Theorem 1.2(3)], this implies that any product of them satisfy **(MTW)**$^{\not\perp}$. We dot not know if such Riemannian products satisfy (**TCP**).

Finally, we point out that to our knowledge there is no concrete example of a Riemannian manifold satisfying **(MTW)** but not **(MTW$^{\not\perp}$)**.

Appendix A. Some notation in Riemannian geometry

Given (M, g) a C^∞ compact connected Riemannian manifold of dimension $n \geq 2$, we denote by TM its tangent bundle, by UM its unit tangent bundle, and by $\exp : (x, v) \longmapsto \exp_x v$ the exponential mapping. We write $g(x) = g_x$, $g_x(v, w) = \langle v, w\rangle_x$, $g_x(v, v) = |v|_x$ and equip M with its geodesic distance d. We further define:

- $t_C(x, v)$: the *cut time* of (x, v):
$$t_C(x, v) = \max\Big\{t \geq 0;\ (\exp_x(sv))_{0\leq s\leq t} \text{ is a minimizing geodesic}\Big\}.$$
- $t_F(x, v)$: the *focalization time* of (x, v):
$$t_F(x, v) = \inf\Big\{t \geq 0;\ \det(d_{tv}\exp_x) = 0\Big\}.$$
- $\mathrm{TCL}(x)$: the *tangent cut locus* of x:
$$\mathrm{TCL}(x) = \big\{t_C(x, v)v;\ v \in T_xM \setminus \{0\}\big\}.$$

• $\mathrm{cut}(x)$: the *cut locus* of x:

$$\mathrm{cut}(x) = \exp_x(\mathrm{TCL}(x)).$$

• $\mathrm{TFL}(x)$: the *tangent focal locus* of x:

$$\mathrm{TFL}(x) = \big\{t_F(x,v)v;\ v \in T_xM \setminus \{0\}\big\}.$$

• $\mathrm{TFCL}(x)$: the *tangent focal cut locus* of x:

$$\mathrm{TFCL}(x) = \mathrm{TCL}(x) \cap \mathrm{TFL}(x).$$

• $\mathrm{fcut}(x)$: the *focal cut locus* of x:

$$\mathrm{fcut}(x) = \exp_x(\mathrm{TFCL}(x)).$$

• $\mathrm{I}(x)$: the *injectivity domain* of the exponential map at x:

$$\mathrm{I}(x) = \big\{tv;\ 0 \le t < t_C(x,v),\ v \in T_xM\big\}.$$

• $\mathrm{NF}(x)$: the *nonfocal domain* of the exponential map at x:

$$\mathrm{NF}(x) = \big\{tv;\ 0 \le t < t_F(x,v),\ v \in T_xM\big\}.$$

• $\exp^{-1}$: the *inverse of the exponential map*; by convention $\exp_x^{-1}(y)$ is the set of *minimizing* velocities v such that $\exp_x v = y$. In particular $\mathrm{TCL}(x) = \exp_x^{-1}(\mathrm{cut}(x))$, and $I(x) = \exp_x^{-1}(M \setminus \mathrm{cut}(x))$.

We notice that, for every $x \in M$, the function $t_C(x,\cdot) : U_xM \to \mathbb{R}$ is locally Lipschitz (see [3, 21, 24]) while the function $t_F(x,\cdot) : U_xM \to \mathbb{R}$ is locally semiconcave on its domain (see [3]). In particular, the regularity property of $t_C(x,\cdot)$ yields

$$\mathcal{H}^{n-1}\big(\mathrm{cut}(x)\big) < +\infty \qquad \forall\, x \in M. \tag{A.1}$$

Appendix B. On the size of the focal cut locus

Recall that, for every $x \in M$, the focal cut locus of a point x is defined as

$$\mathrm{fcut}(x) = \exp_x(\mathrm{TFCL}(x)).$$

The focal cut locus of x is always contained in its cut locus. However it is much smaller, as the following result (which we believe to be of independent interest) shows:

Proposition B.1. *For every $x \in M$ the set* $\mathrm{fcut}(x)$ *has Hausdorff dimension bounded by $n-2$. In particular we have*

$$\forall\, x \in M, \quad \mathcal{H}^{n-1}\big(\mathrm{fcut}(x)\big) = 0. \tag{B.1}$$

Proof. For every $k = 0, 1, \ldots, n$, denotes by Σ_x^k the set of $y \neq x \in M$ such that the convex set $\nabla_x^+ c(x, y)$ has dimension k. By [2, Corollary 4.1.13], since the function $y \mapsto c(x, y)$ is semiconcave, the set Σ_x^k is countably $(n-k)$ rectifiable for every $k = 2, \ldots, n$, which means in particular that all the sets $\Sigma_x^2, \ldots, \Sigma_x^n$ have Hausdorff dimension bounded by $n-2$. Thus, we only need to show that the set

$$J_x = \left(J_x \cap \Sigma_x^0\right) \cup \left(J_x \cap \Sigma_x^1\right)$$

has Hausdorff dimension $\leq n-2$. The fact that $J_x^0 = J_x \cap \Sigma_x^0$ has Hausdorff dimension $\leq n-2$ is a consequence of [30, Theorem 5.1]. Now, consider $\bar{y} \in J_x^1 = J_x \cap \Sigma_x^1$. Then there are exactly two minimizing geodesics $\gamma_1, \gamma_2 : [0,1] \to M$ joining x to $\bar{y}$. By upper semicontinuity of the set of minimizing geodesics joining x to y, for $i = 1, 2$ we can modify the metric g in a small neighborhood of $\gamma_i(1/2)$ into a new metric g_i in such a way that the following holds: there exists an open neighborhood $\mathcal{V}_i$ of $\bar{y}$ such that, for any $y \in J_x^1 \cap \mathcal{V}_i$, there is only one minimizing geodesic (with respect to g_i) joining x to y. In that way, we have

$$J_x^1 \cap \mathcal{V}_1 \cap \mathcal{V}_2 \subset \left(J_x^0\right)_1 \cup \left(J_x^0\right)_2,$$

where $\left(J_x^0\right)_i$ denotes the set $J_x^0 = J_x \cap \Sigma_x^0$ with respect to the metric g_i. Hence we conclude again by [30, Theorem 5.1]. □

As a corollary the following holds:

Lemma B.2. *Let $\bar{x} \in M$, $v_0, v_1 \in \mathrm{I}(\bar{x})$ and $x \in M$ be fixed. Up to slightly perturbing v_0 and v_1, we can assume that $v_0, v_1 \in \mathrm{I}(\bar{x})$, $(y_t)_{0 \leq t \leq 1}$ intersects $\mathrm{cut}(x)$ only at a finite set of times $0 < t_1 < \ldots < t_N < 1$, and moreover $(y_t)_{0 \leq t \leq 1}$ never intersects $\mathrm{fcut}(x) = \exp_x(\mathrm{TFCL}(x))$.*

Proof of Lemma B.2. The proof of this fact is a variant of an argument in [18]: fix $\sigma > 0$ small enough so that

$$w \perp v_1 - v_0,\ |w|_x \leq \sigma \implies \quad v_0 + w, v_1 + w \in I(\bar{x}),$$

and consider the cylinder C_σ in $T_x M$ given by $\{v_t + w\}$, with $t \in [0,1]$ and w as above. By convexity of $\mathrm{TFL}(\bar{x})$, for σ sufficiently small we have $C_\sigma \subset \mathrm{NF}(\bar{x})$. Let us now consider the sets

$$C_\sigma^c = C_\sigma \cap \exp_{\bar{x}}^{-1}\left(\exp_{\bar{x}}(C_\sigma) \cap \mathrm{cut}(x)\right),$$
$$C_\sigma^{cf} = C_\sigma \cap \exp_{\bar{x}}^{-1}\left(\exp_{\bar{x}}(C_\sigma) \cap \mathrm{fcut}(x)\right).$$

Since $C_\sigma \subset \mathrm{NF}(\bar{x})$, $\exp_{\bar{x}}^{-1}$ is locally Lipschitz on $\exp_{\bar{x}}(C_\sigma)$, and therefore (A.1) and (B.1) imply

$$\mathcal{H}^{n-1}(C_\sigma^c) < +\infty, \quad \mathcal{H}^{n-1}(C_\sigma^{cf}) = 0.$$

We now apply the co-area formula in the following form (see [7, p. 109] and [8, Sections 2.10.25 and 2.10.26]): let $f : v_t + w \longmapsto w$ (with the notation above), then

$$\mathcal{H}^{n-1}(A) \geq \int_{f(A)} \mathcal{H}^0[A \cap f^{-1}(w)]\, \mathcal{H}^{n-1}(dw)$$

for any $A \subset C_\sigma$ Borel. Since the right-hand side is exactly $\int \#\{t;\ v_t + w \in A\}\, \mathcal{H}^{n-1}(dw)$, we immediately deduce that particular there is a sequence $w_k \to 0$ such that each $(v_t + w_k)$ intersects C_σ^c finitely many often, and $(v_t + w_k)$ never intersects C_σ^{cf}.

We now also observe that, if $y \in \mathrm{cut}(x) \setminus \mathrm{fcut}(x)$, then $\mathrm{cut}(x)$ is given in a neighborhood of y by the intersection of a finite number of smooth hypersurfaces (see for instance [27]). Thus, up to slightly perturbing v_0 and v_1, we may further assume that at the points y_{t_j} the curve $t \mapsto y_t$ intersects $\mathrm{cut}(x)$ transversally. □

References

[1] J.-P. Aubin and H. Frankowska. *Set-valued analysis*. Systems & Control: Foundations & Applications, 2. Birkhäuser Boston Inc. Boston, MA, 1990.

[2] P. Cannarsa and C. Sinestrari. *Semiconcave functions, Hamilton-Jacobi equations, and optimal control*. Progress in Nonlinear Differential Equations and their Applications, 58. Birkhäuser Boston Inc., Boston, MA, 2004.

[3] M. Castelpietra and L. Rifford. Regularity properties of the distance function to conjugate and cut loci for viscosity solutions of Hamilton-Jacobi equations and applications in Riemannian geometry. *ESAIM Control Optim. Calc. Var.*, 16(3):695–718, 2010.

[4] D. Cordero-Erausquin. Sur le transport de mesures priodiques. *C.R. Acad. Sci. Paris Sér. I Math.*, 329:199–202, 1999.

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

[6] P. Delanoë and Y. Ge. Positivity of C-curvature on surfaces locally near to sphere and its application. Preprint 2010.

[7] L.C. Evans and R. Gariepy. *Measure theory and fine properties of functions*. CRC Press, Boca Raton, FL, 1992.

[8] H. Federer. *Geometric measure theory*. Die Grundlehren der mathematischen Wissenschaften, Band 153. Springer-Verlag New York Inc., New York, 1969.

[9] A. Fathi and A. Figalli. Optimal transportation on non-compact manifolds. *Israel J. Math.*, 175(1):1–59, 2010.

[10] A. Figalli and N. Gigli. Local semiconvexity of Kantorovich potentials on non-compact manifolds. *ESAIM Control Optim. Calc. Var.*, to appear.

[11] A. Figalli, Y. H. Kim and R. J. McCann. Hölder continuity and injectivity of optimal maps. Preprint, 2011.

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

[13] A. Figalli and G. Loeper. C^1 regularity of solutions of the Monge-Ampère equation for optimal transport in dimension two. *Calc. Var. Partial Differential Equations*, 35(4):537–550, 2009.

[14] A. Figalli and L. Rifford. Continuity of optimal transport maps and convexity of injectivity domains on small deformations of $\mathbb{S}^2$. *Comm. Pure Appl. Math.*, 62(12):1670–1706, 2009.

[15] A. Figalli, L. Rifford and C. Villani. Nearly round spheres look convex. *Amer. J. Math.*, to appear.

[16] A. Figalli, L. Rifford and C. Villani. On the Ma-Trudinger-Wang curvature on surfaces. *Calc. Var. Partial Differential Equations*, 39(3-4):307–332, 2010.

[17] A. Figalli, L. Rifford and C. Villani. Tangent cut loci on surfaces. *Differential Geom. Appl.*, to appear.

[18] A. Figalli and C. Villani. An approximation lemma about the cut locus, with applications in optimal transport theory. *Methods Appl. Anal.*, 15(2):149–154, 2008.

[19] A. Figalli and C. Villani. Optimal transport and curvature. Notes for a CIME lecture course in Cetraro, June 2008.

[20] D. Gilbarg and N. Trudinger. *Elliptic partial differential equations of second order.* Classics in Mathematics, Reprint of the 1998 edition. Springer-Verlag, Berlin, 2001.

[21] J. Itoh and M. Tanaka. The Lipschitz continuity of the distance function to the cut locus. *Trans. Amer. Math. Soc.*, 353(1):21–40, 2001.

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

[23] Y.-H. Kim and R. J. McCann. Towards the smoothness of optimal maps on Riemannian submersions and Riemannian products (of round spheres in particular). *J. Reine Angew. Math.*, to appear.

[24] Y. Li and L. Nirenberg. The distance function to the boundary, Finsler geometry, and the singular set of viscosity solutions of some Hamilton-Jacobi equations. *Comm. Pure Appl. Math.*, 58(1):85–146, 2005.

[25] J. Liu, N. Trudinger and X.-J. Wang. Interior $C^{2,\alpha}$-regularity for potential functions in optimal transportation. *Comm. Partial Differential Equations*, to appear.

[26] G. Loeper: On the regularity of solutions of optimal transportation problems. *Acta Math.*, 202(2):241-283, 2009.

[27] G. Loeper and C. Villani. Regularity of optimal transport in curved geometry: the nonfocal case. *Duke Math. J.*, 151(3):431–485, 2010.

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

[29] R. J. McCann. Polar factorization of maps in Riemannian manifolds. *Geom. Funct. Anal.*, 11:589–608, 2001.

[30] C. Pignotti. Rectifiability results for singular and conjugate points of optimal exit time problems. *J. Math. An. Appl.*, 270(2):681–708, 2002.

[31] C. Villani. Optimal transport, old and new. Grundlehren des mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], Vol. 338, Springer-Verlag, Berlin-New York, 2009.

[32] C. Villani. Stability of a 4th-order curvature condition arising in optimal transport theory. *J. Funct. Anal.*, 255(9):2683–2708, 2008.

[33] C. Villani. Regularity of optimal transport and cut locus: from nonsmooth analysis to geometry to smooth analysis. To appear in *Disc. Cont. Dyn. Syst.*, special issue in the honor of L. Nirenberg, Volume III.

Alessio Figalli

Department of Mathematics
The University of Texas at Austin
1 University Station, C1200
Austin TX 78712, USA
email: figalli@math.utexas.edu

Ludovic Rifford

Université de Nice–Sophia Antipolis
Labo. J.-A. Dieudonné, UMR 6621
Parc Valrose
06108 Nice Cedex 02, FRANCE
email: rifford@unice.fr

Cédric Villani

Université de Lyon & Institut Henri Poincaré
11 rue Pierre et Marie Curie
75230 Paris Cedex 05, FRANCE

e-mail: villani@ihp.jussieu.fr