ON THE CONVEXITY OF INJECTIVITY DOMAINS ON NONFOCAL MANIFOLDS

A. FIGALLI, TH. GALLOUËT, AND L. RIFFORD

Abstract. Given a smooth nonfocal compact Riemannian manifold, we show that the so-called Ma–Trudinger–Wang condition implies the convexity of injectivity domains. This improves a previous result by Loeper and Villani.

1. Introduction

Let (M, g) be a smooth compact Riemannian manifold of dimension $n \geq 2$. The *injectivity domain* at a point $x \in M$ is defined as

$$\mathrm{I}(x) := \Big\{ v \in T_xM \,|\, \exists\, t > 1 \text{ s.t. } d(x, \exp_x(tv)) = |tv|_x \Big\},$$

where $\exp_x$ denotes the exponential mapping at x, d the geodesic distance on $M \times M$, and $|v|_x = \sqrt{g_x(v,v)} = \sqrt{\langle v, v\rangle_x}$. We recall that $I(x)$ is an open star-shaped subset of T_xM, and by the Itoh-Tanaka Theorem [4, 16, 19] its boundary $\mathrm{TCL}(x)$ (which is called *tangent cut locus* at x) is Lipschitz. Its image by the exponential mapping is called the *cut locus* of x,

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

Recall that the geodesic distance from x, that is the function $y \mapsto d(x,y)$, is smooth outside $\mathrm{cut}(x)$, and more generally the distance function d is smooth outside the set

$$\mathrm{cut}(M) := \Big\{ (x,y) \in M \times M \,|\, y \in \mathrm{cut}(x) \Big\}.$$

For every $x \in M$, $v \in \mathrm{I}(x)$, and $(\xi, \eta) \in T_xM \times T_xM$, the *Ma–Trudinger–Wang tensor* (or MTW tensor for short) at (x,v) evaluated on (ξ,η) is defined by the formula

$$\mathfrak{S}_{(x,v)}(\xi,\eta) := -\frac{3}{2} \left.\frac{d^2}{ds^2}\right|_{s=0} \left.\frac{d^2}{dt^2}\right|_{t=0} \frac{d^2}{2}\Big(\exp_x(t\xi),\, \exp_x(v + s\eta)\Big). \tag{1.1}$$

(The MTW tensor was introduced for the first time in [22] in a slightly different way, see also [25].) Since $v \in I(x)$ we have that $\exp_x(v) \notin \mathrm{cut}(x)$, hence pair of points $(\exp_x(t\xi), \exp_x(v + s\eta))$ does not belong to $\mathrm{cut}(M)$ provided s, t are small enough, and

the right-hand side in (1.1) is well-defined. As observed by Loeper in [20], if ξ, η are two unit orthogonal vectors in T_xM, then

$$\mathfrak{S}_{(x,0)}(\xi,\eta) = \sigma_x(P)$$

is the sectional curvature of M at x along the plane P generated by ξ and η.

Definition 1.1. We say that (M,g) satisfies **(MTW)** if the following property is satisfied:

$$\forall\, x \in M,\, \forall\, v \in \mathrm{I}(x),\, \forall\, \xi,\eta \in T_xM, \quad \Big[\langle \xi,\eta\rangle_x = 0 \implies \mathfrak{S}_{(x,v)}(\xi,\eta) \geq 0\Big].$$

We say that (M,g) satisfies **(MTW(K,C))** if there exists $(K,C) \in \mathbb{R}\times\mathbb{R}\cup\{+\infty\}$:

$$\forall\, x \in M,\, \forall\, v \in \mathrm{I}(x),\, \forall\, \xi,\eta \in T_xM, \quad \mathfrak{S}_{(x,v)}(\xi,\eta) \geq -C\,|\langle \xi,\eta\rangle_x|\,|\xi|_x|\eta|_x + K|\xi|_x^2|\eta|_x^2.$$

The **(MTW)** property imposes hard constraints on the geometry of (M,g). First, by Loeper's observation above, if (M,g) satisfies **(MTW)** then it must have nonnegative sectional curvatures. Moreover, as shown by Loeper and Villani in [21], the **(MTW)** property has some effects on the geometry of injectivity domains. They proved that if (M,g) is nonfocal and satisfies a stronger form of the **(MTW)** condition, then all its injectivity domain must be uniformly convex. The aim of the present paper is to improve the result by Loeper and Villani by showing that the strong form of **(MTW)** condition can be dropped. Before to state our main result, let us briefly recall the link between **(MTW)** and the regularity of optimal transports with quadratic geodesic costs, which was the initial motivation for the introduction of the Ma–Trudinger–Wang tensor, see [25].

Let μ,ν be two probability measures on M and $c : M\times M \to \mathbb{R}$ be the quadratic geodesic cost defined by

$$c(x,y) := \frac{d(x,y)^2}{2} \qquad \forall\,(x,y) \in M\times M.$$

The Monge problem from μ to ν and cost c consists in finding a measurable 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 is absolutely continuous, then according to McCann [23] this minimizing problem has a solution T, unique up to modification on a μ-negligible set. A natural question is whether the optimal transport map can be expected to be continuous. To this purpose, we introduce the following definition.

Definition 1.2. We say that (M,g) satisfies the *transport continuity property* (abbreviated $\mathcal{TCP}$) 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 from μ to ν with cost c is continuous, up to modification on a set of zero volume.

The following results give necessary and sufficient conditions for $\mathcal{TCP}$ in terms of the **(MTW)** property and convexity properties of injectivity domains, see [11]. Theirs proofs are based on previous works by many authors, see [5, 7, 8, 17, 18, 20, 21, 22, 25].

Theorem 1.3. *Assume that (M, g) satisfies the $\mathcal{TCP}$ condition. Then (M, g) satisfies* **(MTW)** *and all its injectivity domains are convex.*

Theorem 1.4. *Assume that M has dimension 2. Then the $\mathcal{TCP}$ condition holds if and only if (M, g) satisfies* **(MTW)** *and all its injectivity domains are convex.*

Let us now state our main result. The nonfocal domain at some $x \in M$ is defined as

$$\mathrm{NF}(x) := \Big\{ v \in T_xM \,|\, d_{tv} \exp_x \text{ is not singular for any } t \in [0,1] \Big\}.$$

It is an open star-shaped subset of T_xM whose boundary $\mathrm{TFL}(x)$ is called the tangent focal domain at x. The set $\overline{\mathrm{NF}}(x) = \mathrm{NF}(x) \cup \mathrm{TFL}(x)$ can be shown to be locally semiconvex (see [4] and Appendix A), and the following inclusion always holds:

$$\mathrm{I}(x) \subset \mathrm{NF}(x) \qquad \forall\, x \in M,$$

see for instance [14, Corollary 3.77] or [25, Problem 8.8].

Definition 1.5. We say that (M, g) is nonfocal provided

$$\mathrm{TCL}(x) \subset \mathrm{NF}(x) \qquad \forall\, x \in M.$$

In [21], Loeper and Villani proved that if (M, g) is nonfocal and satisfies the following strict form of the **(MTW)** condition,

$$\mathfrak{S}_{(x,v)}(\xi, \eta) \geq K |\xi|_x^2 |\eta|_x^2 \qquad \forall\, x \in M,\, \forall\, v \in \mathrm{I}(x),\, \forall\, \xi, \eta \in T_xM$$

for some $K > 0$, then all its injectivity domain are uniformly convex. Our main result shows that the **(MTW)** condition alone is sufficient for the convexity of injectivity domains.

Theorem 1.6. *Let (M, g) be a nonfocal Riemannian manifold satisfying* **(MTW)**. *Then all injectivity domains of M are convex.*

Our proof is based on techniques relying on the extended Ma–Trudinger–Wang tensor, which were introduced by the first and third author in [8], together with bootstrap arguments. In fact, Theorem 1.6 provides a partial answer to a conjecture formulated by Villani in [26].

Villani's Conjecture. Let (M, g) be a smooth compact Riemannian manifold satisfying **(MTW)**. Then all its injectivity domains are convex.

We will address the above conjecture in the case of analytic surfaces in a forthcoming paper [6]. In fact, we take opportunity of the present paper to present a slight improvement (Theorem 4.1) of Theorem 1.6 that will be useful in [6].

The paper is structured as follows: In Section 2 we provide some preliminary results about injectivity and nonfocal domains. Then, Section 3 contains the proof of Theorem 1.6. Section 4 is devoted to the proof of Theorem 4.1 whose core of the proof follows the strategy developed in Section 3 together with additional technicalities, and in Section 5 we show how to recover Loeper-Villani's result with our techniques. Finally, in the appendices we collect some useful results on semiconvex functions and tangent cut loci.

2. Preliminary results

Let M be a smooth compact Riemannian manifold, and denote by $UM \subset TM$ the unit tangent bundle. Let us introduce some definitions and notation.

The distance function to the cut locus at some $x \in M$, $t_{cut} : UM \to (0, \infty)$, is defined as

$$\begin{aligned} t_{cut}(x,v) &:= \sup\Big\{ t \geq 0 \,|\, tv \in \mathrm{I}(x) \Big\} \\ &= \max\Big\{ t \geq 0 \,|\, d(x, \exp_x(tv)) = t \Big\}. \end{aligned}$$

Then, for every $x \in M$, there holds

$$\mathrm{I}(x) = \Big\{ tv \,|\, 0 \leq t < t_{cut}(x,v),\ v \in U_x M \Big\}, \qquad \mathrm{TCL}(x) = \Big\{ t_{cut}(x,v)v \,|\, v \in U_x M \Big\}.$$

For every $x \in M$, we denote by ρ_x the radial distance on T_xM, that is

$$\rho_x(v,w) := \begin{cases} |v|_x + |w|_x & \text{if} \quad g_x(v,w) \neq |v|_x |w|_x \\ |v - w|_x & \text{if} \quad g_x(v,w) = |v|_x |w|_x. \end{cases}$$

Then the radial distance to $\mathrm{I}(x)$ satisfies for any $v \in T_xM$,

$$\begin{aligned} \rho_x\big(v, \mathrm{I}(x)\big) &:= \inf\Big\{ \rho_x(v,w) \,|\, w \in \mathrm{I}(x) \Big\} \\ &= \begin{cases} \left| v - t_{cut}\left(x, \frac{v}{|v|_x}\right) \frac{v}{|v|_x} \right|_x & \text{if } v \notin \mathrm{I}(x), \\ 0 & \text{otherwise.} \end{cases} \end{aligned}$$

For every $v \in \mathrm{TCL}(x)$ we set

$$\delta(v) := \max\Big\{ |v - w|_x \,|\, w \in \mathrm{TCL}(x) \ \text{ s.t.} \exp_x v = \exp_x w \Big\},$$

for every compact set $V(x) \subset T_xM$

$$\delta(V(x)) := \min\Big\{ \delta(v) \,|\, v \in V(x) \cap \mathrm{TCL}(x) \Big\},$$

and finally for every compact set $V \subset TM$ we let

$$\delta(V) := \min\Big\{\delta\big(V(x)\big) \,|\, x \in M\Big\},$$

where for each $x \in M$, $V(x)$ denotes the fiber of V over x (which might be empty, in which case $\delta(V(x)) = +\infty$). Notice that nonfocal compact Riemannian manifolds satisfy $\delta(TM) > 0$. However, Riemannian manifolds satisfying $\delta(TM) > 0$ are not necessarily nonfocal, as the property $\delta(TM) > 0$ only rules out purely focal velocities.

Lemma 2.1. *Let V be a compact subset of TM with $\delta(V) > 0$ such that each $V(x) \neq \emptyset$ is starshaped with respect to the origin. Then, there exists $K > 0$ such, that for every $(x, v) \in V$,*

$$\rho_x\big(v, \mathrm{I}(x)\big) \leq K\left(|v|_x^2 - d\big(x, \exp_x(v)\big)^2\right).$$

In particular assume that (M, g) is nonfocal. Then, there exists $K > 0$ such, that for every $x \in M$ and every $v \in T_xM$,

$$\rho_x\big(v, \mathrm{I}(x)\big) \leq K\left(|v|_x^2 - d\big(x, \exp_x(v)\big)^2\right).$$

Proof of Lemma 2.1. By compactness of M, the geodesic distance (and thus the quantity $d(x, \exp_x(v))$) is uniformly bounded. Then since the right-hand side in the inequalities is quadratic in $|v|_x$ while the left-hand size has linear growth, it is sufficient to show that there is $\delta > 0$ such that

$$|v|_x^2 - d\big(x, \exp_x(v)\big)^2 \leq \delta \quad \Longrightarrow \quad \rho_x\big(v, \mathrm{I}(x)\big) \leq K\left(|v|_x^2 - d\big(x, \exp_x(v)\big)^2\right),$$

for every (x, v) as required. First, for every $(x, v) \in V$ we set

$$\psi_x(v) := d_v \exp_x(v),$$

so that if $\gamma : [0, 1] \to M$ is a constant-speed minimizing geodesic path going from x to y, with initial velocity v_0 and final velocity v_1, the map ψ_x is defined by $v_0 \mapsto v_1$. Since $\delta(V) > 0$ there exists $\Delta > 0$ such that, for every $x \in M$ with $V(x) \neq \emptyset$ and every $v \in V(x) \cap \mathrm{TCL}(x)$, there is a geodesic path starting at x with initial velocity w (with $|w|_x = |v|_x$), and finishing at $y = \exp_v(x)$ with final velocity $\psi_x(w)$, satisfying

$$|v|_x^2 - \langle \psi_x(v), \psi_x(w)\rangle_y > \Delta, \tag{2.1}$$

see for instance [21, Proposition C.5(a)]. Let $v \in \mathrm{TCL}(x) \cap V(x)$ and $y := \exp_x(v)$ be fixed. As before, consider a minimizing geodesic path from x to y with initial velocity w satisfying (2.1). Since $d^2(x, \cdot)$ is locally semiconcave on M, $2\psi_x(w)$ is a supergradient for $d^2(x, \cdot)$ at y, and the distance from x to its cut locus is uniformly bounded from below (see [25, Definition 10.5 and Proposition 10.15]), it is easy to show the existence of a smooth

function $h : M \to \mathbb{R}$, whose C^2 norm does not depend on x and v, and such that

$$\begin{cases} d(x,y)^2 = h(y) = |v|_x^2, \\ \nabla h(y) = 2\psi_x(w) \\ d(x,z)^2 \le h(z), \ \forall\, z \in M, \end{cases}$$

see for instance [21, Proposition C.6]. This gives

$$|(1+\epsilon)v|_x^2 - d\big(x, \exp_x((1+\epsilon)v)\big)^2 \ge (1+\epsilon)^2|v|_x^2 - h\big(\exp_x((1+\epsilon)v)\big) \qquad \forall\, \epsilon.$$

Hence, if C_0 denotes a uniform bound for the C^2 norm of h independent of x and v, we get

$$|(1+\epsilon)v|_x^2 - d\big(x, \exp_x((1+\epsilon)v)\big)^2 \ge 2\epsilon\left(|v|_x^2 - \langle \psi_x(v), \psi_x(w)\rangle\right) - C_0\epsilon^2 \qquad \forall \epsilon.$$

Then, using (2.1), we deduce that

$$|(1+\epsilon)v|_x^2 - d\big(x, \exp_x((1+\epsilon)v)\big)^2 \ge \epsilon\Delta \qquad \forall\, \epsilon \in (-\epsilon_0, \epsilon_0),$$

where $\epsilon_0 := \Delta/C_0$. Since

$$\rho_x\big((1+\epsilon)v, \mathrm{I}(x)\big) = |(1+\epsilon)v - v|_x = \epsilon|v|_x,$$

we finally obtain

$$\rho_x\big((1+\epsilon)v, \mathrm{I}(x)\big) \le \frac{|v|_x}{\Delta}\Big(|(1+\epsilon)v|_x^2 - d\big(x, \exp_x((1+\epsilon)v)\big)^2\Big) \qquad \forall\, \epsilon \in (-\epsilon_0, \epsilon_0).$$

To conclude the proof it suffices to observe that, by a simple compactness argument together with the fact that each $V(x) \neq \emptyset$ is starshaped, one can easily check that there exists $\delta > 0$ such that any $w \in V(x) \setminus \mathrm{I}(x)$, with $|w|_x^2 - d\big(x, \exp_x(w)\big)^2 \le \delta$, has the form $(1+\epsilon)v$ for some $v \in \mathrm{TCL}(x) \cap V(x)$ and $\epsilon \in [0, \epsilon_0)$. □

Lemma 2.2. *There exists $K > 0$ such that for every $(x,v) \in TM$,*

$$K^{-1}\rho_x\big(v, \mathrm{I}(x)\big) \le \rho_y\big(w, \mathrm{I}(y)\big) \le K\rho_x\big(v, \mathrm{I}(x)\big)$$

and for every $(x,v) \in TM$ with $v \in \mathrm{I}(x)$,

$$K^{-1}\rho_x\big(v, \mathrm{TFL}(x)\big) \le \rho_y\big(w, \mathrm{TFL}(y)\big) \le K\rho_x\big(v, \mathrm{TFL}(x)\big),$$

where $y = \exp_x(v)$ and $w = -d_v \exp_x(v) = -\psi_x(v)$, so in particular $x = \exp_y(w)$.

Proof of Lemma 2.2. The second inequality follows easily by compactness arguments. Let us prove the first inequality. As before, it is sufficient to show the result provided $\rho_x(v, \mathrm{I}(x)) \le \delta$ for some $\delta > 0$. Indeed $\rho_x(v, \mathrm{I}(x)) = 0$ is equivalent to $\rho_y(w, \mathrm{I}(y)) = 0$, so all terms vanish. Let $(x,v) \in TM$ be fixed, set $e_v = \frac{v}{|v|_x}$ and

$$y = \exp_x(v), \quad w = -\psi_x(v), \quad e_w = \frac{w}{|w|_x}, \quad \overline{w} := t_{cut}\,(y, e_w)\, e_w,$$

and in addition

$$\overline{v} := t_{cut}\,(x, e_v)\, e_v, \quad z := \exp_x(\overline{v}), \quad w' := -\psi_x(\overline{v}).$$

Note that since $\overline{v}$ belongs to $\mathrm{TCL}(x)$ the velocity w' belongs to $\mathrm{TCL}(z)$, so it satisfies

$$w' = t_{cut}\,(z, e_{w'})\, e_{w'}.$$

Moreover,

$$\rho_x\big(v, \mathrm{I}(x)\big) = |v - \overline{v}|_x \qquad \text{and} \qquad \rho_y\big(w, \mathrm{I}(y)\big) = |w - \overline{w}|_y\,.$$

Equip TM with any distance d_{TM} which in charts is locally bi-Lipschitz equivalent to the Euclidean distance on $\mathbb{R}^n \times \mathbb{R}^n$. We may assume that $|v|_x$ is bounded. Since the geodesic flow is Lipschitz on compact subsets of TM, there holds

$$d_{TM}\left((y, w), (z, w')\right) \leq K' \big|v - \overline{v}\big|_x,$$

for some uniform constant K'. In fact, if v is close to $\mathrm{I}(x)$ then $\overline{v}$ is close to v, and so also y and z are close to each other, so the above inequality follows from our assumption on d_{TM}. Then, assuming that $\rho_x(v, \mathrm{I}(x)) \leq \delta$ for $\delta > 0$ small enough and taking a local chart in a neighborhood of y if necessary, we may assume that $y, z, w, \overline{w}, w'$ are in $\mathbb{R}^n$. Moreover, up to a bi-Lipschitz transformation which may affect the estimates only up to a uniform multiplicative constant, we may assume for simplicity that d_{TM} coincides with the Euclidean distance on $\mathbb{R}^n \times \mathbb{R}^n$. Since y is perturbed along the geodesic flow, Theorem B.2 gives

$$\begin{aligned}
|w - \overline{w}|_y &= |w|_y - t_{cut}(y, e_w) = |v|_x - |\overline{v}|_x + |\overline{v}|_x - t_{cut}(y, e_w) \\
&= |v|_x - |\overline{v}|_x + \big|w'\big|_z - t_{cut}(y, e_w) \\
&= |v - \overline{v}|_x + t_{cut}(z, e_{w'}) - t_{cut}(y, e_w) \\
&\leq |v - \overline{v}|_x + KK'\,|v - \overline{v}|_x\,.
\end{aligned}$$

□

We are now ready to start the proof of Theorem 1.6.

3. Proof of Theorem 1.6

Let (M, g) be a smooth compact Riemannian manifold of dimension $n \geq 2$ which is nonfocal and satisfies **(MTW)**, and let $K > 0$ be a constant such that all properties of Lemmas 2.1-2.2 are satisfied. For every $\mu > 0$, we set

$$\mathrm{I}^\mu(x) := \Big\{ v \in T_xM \,|\, \rho_x(v, \mathrm{I}(x)) \leq \mu \Big\}.$$

Since M is assumed to be nonfocal, there is $\overline{\mu} > 0$ small enough such that $\mathrm{I}^{\overline{\mu}}(x)$ does not intersect $\mathrm{TFL}(x)$ for any $x \in M$.

Lemma 3.1. *Taking $K > 0$ larger if necessary, we may assume that for every $x \in M$ and any $v_0, v_1 \in \mathrm{I}(x)$ there holds*

$$v_t := (1-t)v_0 + tv_1 \in \mathrm{I}^{K|v_1 - v_0|_x}(x)$$

and

$$\overline{q}_t := -d_{v_t} \exp_x(v_t) \in \mathrm{I}^{K|v_1 - v_0|_x}\big(y_t\big),$$

with $y_t := \exp_x(v_t)$.

Proof of Lemma 3.1. Since the functions $v \in U_x M \mapsto t_{cut}(x,v)$ are uniformly Lipschitz, there is $K > 0$ such that

$$\rho_x\big(v_t, \mathrm{I}(x)\big) \le K|v_1 - v_0|_x \qquad \forall\, v_0, v_1 \in \mathrm{I}(x),\, \forall\, x \in M.$$

The definition of $\mathrm{I}^{K|v_1-v_0|_x}(x)$ together with Lemma 2.2 yield both inclusions. □

Our proof requires the use of the extended MTW tensor which was initially introduced by the first and third author in [8]. To define this extension, we let $x \in M$, $v \in \mathrm{NF}(x)$, and $(\xi, \eta) \in T_xM \times T_xM$. Since $y := \exp_x v$ is not conjugate to x, by the Inverse Function Theorem there exist an open neighbourhood $\mathcal{V}$ of (x,v) in TM, and an open neighbourhood $\mathcal{W}$ of (x,y) in $M \times M$, such that

$$\begin{array}{rcl} \Psi_{(x,v)} : \mathcal{V} \subset TM & \longrightarrow & \mathcal{W} \subset M \times M \\ (x', v') & \longmapsto & \big(x', \exp_{x'}(v')\big) \end{array}$$

is a smooth diffeomorphism from $\mathcal{V}$ to $\mathcal{W}$. Then we may define $\widehat{c}_{(x,v)} : \mathcal{W} \to \mathbb{R}$ by

$$\widehat{c}_{(x,v)}(x', y') := \frac{1}{2}\big|\Psi^{-1}_{(x,v)}(x', y')\big|^2_{x'}, \qquad \forall\, (x', y') \in \mathcal{W}. \tag{3.1}$$

If $v \in \mathrm{I}(x)$ then for y' close to $\exp_x v$ and x' close to x we have $\widehat{c}_{(x,v)}(x', y') = c(x', y') := d(x', y')^2/2$. For every $x \in M$, $v \in \mathrm{NF}(x)$ and $(\xi, \eta) \in T_xM \times T_xM$, the *extended Ma–Trudinger–Wang tensor* at (x,v) is defined by the formula

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

The following lemma may be seen as an "extended" version of [21, Lemma 2.3].

Lemma 3.2. *There exist constants $C, D > 0$ such that, for any $(x,v) \in TM$ with $v \in \mathrm{I}^{\overline{\mu}}(x)$,*

$$\overline{\mathfrak{S}}_{(x,v)}(\xi, \eta) \ge -C\,|\langle \xi, \eta \rangle_x|\,|\xi|_x |\eta|_x - D\rho_x(v, \mathrm{I}(x))|\xi|^2_x |\eta|^2_x \qquad \forall\, \xi, \eta \in T_xM.$$

We also give a local version of this theorem when M is not nonfocal.

Lemma 3.3. *Let $V \subset TM$ and $\mu > 0$ such that*

$$\rho\big(V \cap \mathrm{I}, \mathrm{TFL}\big) := \sup\Big\{\rho_x(v,w) \,|\, x \in M,\, v \in V(x) \cap \mathrm{I}(x), w \in \mathrm{TFL}(x)\Big\} > \mu.$$

Then there exist constants $C, D > 0$ such that, for any $(x,v) \in TM$ with $v \in V(x) \cap \mathrm{I}^\mu(x)$,

$$\overline{\mathfrak{S}}_{(x,v)}(\xi,\eta) \geq -C\,|\langle \xi,\eta\rangle_x|\,|\xi|_x|\eta|_x - D\rho_x(v,\mathrm{I}(x))|\xi|_x^2|\eta|_x^2 \qquad \forall\, \xi,\eta \in T_xM.$$

Proof of Lemma 3.2. The tensors $\mathfrak{S}$ and $\overline{\mathfrak{S}}$ coincide on the sets of $(x,v) \in TM$ such that $v \in \mathrm{I}(x)$, hence

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

Let $\mathrm{I}^{\overline{\mu}}(M)$ be the compact subset of TM defined by

$$\mathrm{I}^{\overline{\mu}}(M) := \cup_{x\in M}\left(\{x\} \times \mathrm{I}^{\overline{\mu}}(x)\right).$$

The mapping

$$(x,v) \in \mathrm{I}^{\overline{\mu}}(M) \longmapsto \big(x, \exp_x(v)\big)$$

is a smooth local diffeomorphism at any $(x,v) \in \mathrm{I}^{\overline{\mu}}(M)$ and the set of (x,v,ξ,η) with $(x,v) \in \mathrm{I}^{\overline{\mu}}(M)$ and $\xi,\eta \in U_xM$ such that $\langle \xi,\eta\rangle_x = 0$ is compact. Then there is $D > 0$ such that

$$\overline{\mathfrak{S}}_{(x,v)}(\xi,\eta) \geq -D\rho_x(v,\mathrm{I}(x)),$$

for every x,v,ξ,η with $(x,v) \in \mathrm{I}^{\overline{\mu}}(M)$ and $\xi,\eta \in U_xM$ such that $\langle \xi,\eta\rangle_x = 0$. By homogeneity we infer that

$$\overline{\mathfrak{S}}_{(x,v)}(\xi,\eta) \geq -D\rho_x(v,\mathrm{I}(x))|\xi|_x^2|\eta|_x^2,$$

for every x,v,ξ,η with $(x,v) \in \mathrm{I}^{\overline{\mu}}(M)$ and $\xi,\eta \in T_xM$ such that $\langle \xi,\eta\rangle_x = 0$. We conclude as in the proof of [21, Lemma 2.3]. □

The proof of Lemma 3.3 follows by the same arguments. The following lemma will play a crucial role.

Lemma 3.4. *Let $h : [0,1] \to [0,\infty)$ be a semiconvex function such that $h(0) = h(1) = 0$ and let $c \geq 0$ be fixed. Assume that there are $t_1 < \ldots < t_N$ in $(0,1)$ such that h is not differentiable at t_i for $i = 1,\ldots,N$, is of class C^2 on $(0,1) \setminus \{t_1,\ldots,t_N\}$, and satisfies*

$$\ddot{h}(t) \geq -|\dot{h}(t)| - c \qquad \forall\, t \in [0,1] \setminus \big\{t_1,\ldots,t_N\big\}. \tag{3.2}$$

Then

$$h(t) \leq c\,t(1-t) \qquad \forall\, t \in [0,1]. \tag{3.3}$$

Moreover, if in addition there exists a constant $\epsilon \geq 0$ such that

$$c \leq \|h\|_\infty + \epsilon, \tag{3.4}$$

then

$$\|h\|_\infty \le \epsilon/3. \tag{3.5}$$

Proof of Lemma 3.4. Let $a > 0$ and $f : [0,1] \to \mathbb{R}$ be the semiconvex function defined by

$$f(t) = h(t) - at(1-t) \qquad \forall\, t \in [0,1].$$

Let $\bar t$ be a maximum point for f. Since f is semiconvex, it has to be differentiable at $\bar t$, so $\bar t \neq t_i$ for $i = 1, \dots, N$. If $\bar t \in (0,1)$, then there holds $\dot f(\bar t) = 0$ and $\ddot f(\bar t) \le 0$. Thus, using (3.2) we get

$$|\dot h(\bar t)| = a|2\bar t - 1| \le a,$$

$$0 \ge \ddot f(\bar t) = \ddot h(\bar t) + 2a \ge -|\dot h(\bar t)| - c + 2a \ge a - c.$$

This yields a contradiction as soon as $a > c$, which implies that in that case f attains its maximum on the boundary of $[0,1]$. Since $f(0) = f(1) = 0$, we infer that

$$h(t) \le at(1-t) \qquad \forall\, t \in [0,1],$$

for every $a > c$. Letting $a \downarrow c$, we get (3.3). Finally, if (3.4) is satisfied, (3.3) implies (recall that h is nonnegative)

$$\|h\|_\infty = \sup_{t\in[0,1]} |h(t)| \le (\|h\|_\infty + \epsilon) \sup_{t\in[0,1]} t(1-t) = (\|h\|_\infty + \epsilon)/4$$

and inequality (3.5) follows easily. □

We recall that given $v_0, v_1 \in \mathrm{I}(x)$, for every $t \in [0,1]$ we set

$$v_t := (1-t)v_0 + tv_1, \qquad y_t := \exp_x(v_t), \qquad \overline{q}_t := -d_{v_t} \exp_x(v_t).$$

In addition, whenever y_t does not belong to $\mathrm{cut}(x)$ (or equivalently $x \notin \mathrm{cut}(y_t)$) we denote by q_t the velocity in $\mathrm{I}(y_t)$ such that

$$\exp_{y_t}(q_t) = x \quad \text{ and } \quad |q_t|_{y_t} = d(x, y_t).$$

The following results follow respectively from [11, Lemma B.2] and [12, Proposition 6.1] and do not need the nonfocality assumption. The idea of Lemma 3.6 goes back to Kim and McCann [17]. Lemma 3.5 is an improvement of [13].

Lemma 3.5. *Let $x \in M$ and $v_0, v_1 \in \mathrm{I}(x)$ be fixed. Then, up to slightly perturbing v_0 and v_1, we can assume that $v_0, v_1 \in \mathrm{I}(x)$ and that the semiconvex function $h : [0,1] \to \mathbb{R}$ defined as*

$$h(t) := \frac{|v_t|_x^2}{2} - \frac{d(x, y_t)^2}{2} \qquad \forall\, t \in [0,1],$$

is of class C^2 outside a finite set of times $0 < t_1 < \ldots < t_N < 1$ and not differentiable at t_i for $i = 1, \dots, N$.

Lemma 3.6. *Let $x \in M$ and $v_0, v_1 \in \mathrm{I}(x)$. Assume that the function h defined above is C^2 outside a finite set of times $0 < t_1 < \ldots < t_N < 1$, and is not differentiable at t_i for $i = 1, \ldots, N$. Furthermore, suppose that $[\overline{q}_t, q_t] \subset \mathrm{NF}(y_t)$ for all $t \in [0,1]$. Then for every $t \in [0,1] \setminus \{t_1, \ldots, t_N\}$ we have*

$$\dot{h}(t) = \big\langle q_t - \overline{q}_t, \dot{y}_t \big\rangle_{y_t}, \tag{3.6}$$

$$\ddot{h}(t) = \frac{2}{3} \int_0^1 (1-s)\, \overline{\mathfrak{S}}_{(y_t, (1-s)\overline{q}_t + s q_t)}(\dot{y}_t, q_t - \overline{q}_t)\, ds. \tag{3.7}$$

The next lemma deals with semiconvexity properties of the sets $I(x)$. We refer the reader to the Appendix A for the main definitions and properties of semiconvex sets.

Lemma 3.7. *There exists a large universal constant $K > 0$ such that the following properties are satisfied for any $x \in M$:*

(i) *Assume there are constants $\omega > 0$ and $\kappa \in (0, \overline{\mu})$ such that*

$$\forall\, v_0, v_1 \in I(x), \quad |v_1 - v_0|_x \le \omega \quad \Longrightarrow \quad \sup_{q \in [\overline{q}_t, q_t]} \Big\{ \rho_{y_t}\big(q, \mathrm{I}(y_t)\big) \Big\} \le \kappa.$$

Then $\overline{\mathrm{I}}(x)$ is $(K\kappa)$-radial-semiconvex.

(ii) *Assume there are constants $\omega, \alpha, \epsilon \ge 0$ such that*

$$\forall\, v_0, v_1 \in I(x), \quad |v_1 - v_0|_x \le \omega$$
$$\Longrightarrow \quad \sup_{q \in [\overline{q}_t, q_t]} \Big\{ \rho_{y_t}\big(q, \mathrm{I}(y_t)\big) \Big\} \le \min\Big\{ \alpha \left(\frac{|v_t|_x^2}{2} - \frac{d(x, y_t)^2}{2} \right) + \epsilon, \overline{\mu} \Big\}.$$

Then $\overline{\mathrm{I}}(x)$ is $(K\epsilon)$-radial-semiconvex.

Proof of Lemma 3.7. We first prove assertion (i). We need to show that there is a uniform constant $K > 0$ and $\nu > 0$ sufficiently small (see Appendix A) such that, for any $v_0, v_1 \in \mathrm{I}(x)$ with $|v_0 - v_1|_x < \nu$,

$$\rho_x\big(v_t, \mathrm{I}(x)\big) \le K\kappa \frac{t(1-t)}{2} \big|v_0 - v_1\big|^2 \qquad \forall\, t \in [0,1].$$

As in Lemma 3.5 we set

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

By Lemma 2.1 it is sufficient to show that

$$h(t) \le K\kappa \frac{t(1-t)}{2} \big|v_0 - v_1\big|^2 \qquad \forall\, t \in [0,1],$$

for some constant $K > 0$. Let $v_0, v_1 \in \mathrm{I}(x)$ and $\nu > 0$ with $|v_1 - v_0|_x < \nu \le \omega$ be fixed. By Lemma 3.5, up to slightly perturbing v_0, v_1 we may assume that $h : [0,1] \to \mathbb{R}$ is

semiconvex, C^2 outside a finite set of times $0 < t_1 < \ldots < t_N < 1$, and not differentiable at t_i for $i = 1, \ldots, N$. By Lemmas 3.2 and 3.6 (observe that $\kappa < \overline{\mu}$ and $\mathrm{I}^{\overline{\mu}}(y_t) \subset \mathrm{NF}(y_t)$),

$$\ddot{h}(t) \geq -C|\dot{h}(t)||\dot{y}_t|_{y_t}|q_t - \overline{q}_t|_{y_t} - D \max_{q \in [\overline{q}_t, q_t]} \Big\{ \rho_{y_t}(q, \mathrm{I}(y_t)) \Big\} |\dot{y}_t|^2_{y_t} |q_t - \overline{q}_t|^2_{y_t},$$

for every $t \in [0,1] \setminus \{t_1, \ldots, t_N\}$. Moreover, by compactness of M, there is a uniform constant $E > 0$ such that

$$\big|\dot{y}_t\big|_{y_t} \leq E \big|v_0 - v_1\big|_x \quad \text{and} \quad \big|q_t - \overline{q}_t\big|_{y_t} \leq E.$$

Hence

$$\ddot{h}(t) \geq -CE^2 |\dot{h}(t)| \big|v_1 - v_0\big|_x - DE^4 \kappa \big|v_1 - v_0\big|^2_x \qquad \forall\, t \in [0,1] \setminus \big\{t_1, \ldots, t_N\big\}. \tag{3.8}$$

Taking $\nu \in (0, \omega)$ small enough yields

$$\ddot{h}(t) \geq -|\dot{h}(t)| - DE^4 \kappa |v_1 - v_0|^2 \qquad \forall\, t \in [0,1] \setminus \big\{t_1, \ldots, t_N\big\},$$

so Lemma 3.4 gives

$$h(t) \leq DE^4 \kappa\, t(1-t) |v_1 - v_0|^2_x \qquad \forall\, t \in [0,1],$$

which shows that $\mathrm{I}(x)$ is $(K\kappa)$-radial-semiconvex where $K > 0$ is a uniform constant. To prove (ii) we note that (3.8) implies

$$\ddot{h}(t) \geq -CE^2 |\dot{h}(t)| \big|v_1 - v_0\big|_x - DE^4 \alpha |h(t)| \big|v_1 - v_0\big|^2_x - DE^4 \epsilon \big|v_1 - v_0\big|^2_x, \tag{3.9}$$

which (by choosing $\nu \in (0, \omega)$ sufficiently small) gives

$$\ddot{h}(t) \geq -|\dot{h}(t)| - \|h\|_\infty - DE^4 \epsilon |v_1 - v_0|^2_x \qquad \forall\, t \in [0,1] \setminus \big\{t_1, \ldots, t_N\big\}.$$

Hence, by the second part of Lemma 3.4 we obtain

$$\|h\|_\infty \leq \frac{DE^4}{3} \epsilon |v_1 - v_0|^2_x.$$

Plugging this information back into (3.9) gives, for ν sufficiently small,

$$\ddot{h}(t) \geq -CE^2 |\dot{h}(t)| \big|v_1 - v_0\big|_x - 2DE^4 \epsilon \big|v_1 - v_0\big|^2_x.$$

We conclude as in the first part of the proof. □

Returning to the proof of Theorem 1.6, we say that the property $\mathcal{P}(r)$ is satisfied if for any $x \in M$ the set $B_x(r) \cap \mathrm{I}(x)$ is convex (here $B_x(r)$ denotes the unit open ball in T_xM with respect to $|\cdot|_x$). If $\mathcal{P}(r)$ is satisfied for any $r \geq 0$, then all the injectivity domains of M are convex. Since $r_0 := \inf_{x \in M,\, v \in \mathrm{TCL}(x)} |v|_x$ is strictly positive, $\mathcal{P}(r)$ is true for any $r \leq r_0$, hence the set of $r \geq 0$ such that $\mathcal{P}(r)$ is satisfied is an interval J with positive length. Moreover, since the convexity property is closed, J is closed. Consequently, in order to prove that $J = [0, \infty)$, it is sufficient to show that J is open.

Lemma 3.8. *The set of r for which $\mathcal{P}(r)$ holds is open in $[0,\infty)$.*

Proof of Lemma 3.8. Assume that $\mathcal{P}(r)$ holds. We want to prove that, if $\beta > 0$ is sufficiently small then $P(r+\beta)$ holds as well. The proof is divided in two steps: first we will show that, for any $\beta \in (0, \overline{\mu}/(2K))$ (here μ and K are as in Lemma 3.7), the sets $B_x(r+\beta) \cap \mathrm{I}(x)$ are $(K\beta)$-radial-semiconvex for any $x \in M$. Then, in Step 2 we show the following "bootstrap-type" result: if the sets $B_x(r+\beta)\cap \mathrm{I}(x)$ are A-radial-semiconvex for all $x \in M$, then they are indeed $(A/2)$-radial-semiconvex. The combination of Steps 1 and 2 proves that, for any $x \in M$ and $\beta > 0$ small, the sets $B_x(r+\beta) \cap \mathrm{I}(x)$ are $(K\beta/2^k)$-radial-semiconvex for any $k \in \mathbb{N}$, hence convex.

Step 1: $\mathrm{I}(x) \cap B_x(r+\beta)$ *is $(K\beta)$-radial-semiconvex for any $\beta \in (0, \overline{\mu}/(2K))$.*

Fix $x \in M$ and $\nu > 0$. Thanks to Lemma 3.1, for any $v_0, v_1 \in \mathrm{I}(x)$ with $|v_0 - v_1|_x < \nu$ we have

$$v_t \in \mathrm{I}^{K\nu}(x) \quad \text{and} \quad \overline{q}_t \in \mathrm{I}^{K\nu}(y_t).$$

Let $\beta > 0$ and $v_0, v_1 \in B_x(r+\beta)\cap \mathrm{I}(x)$ be fixed. By construction

$$|\overline{q}_t|_{y_t} = |v_t|_x < r+\beta, \quad |q_t|_{y_t} \le |v_t|_x < r+\beta, \quad q_t \in \mathrm{I}(y_t).$$

Since $\overline{q}_t \in \mathrm{I}^{K\nu}(y_t)$ we can find $q_t' \in \overline{\mathrm{I}(y_t)} \cap B_{y_t}(r+\beta)$ such that

$$\rho_{y_t}\big(\overline{q}_t, \mathrm{I}(y_t)\big) = |\overline{q}_t - q_t'| \le K\nu.$$

Moreover, using that $\mathrm{I}(y_t)$ is starshaped and that $q_t, q_t' \in B_{y_t}(r+\beta)$, we can find q_t^r, $q_t'^r \in \overline{B}_{y_t}(r) \cap \overline{\mathrm{I}(y_t)}$ such that $\rho_{y_t}(q_t, q_t^r) \le \beta$ and $\rho_{y_t}(q_t', q_t'^r) \le \beta$. Recalling that by assumption $\mathcal{P}(r)$, we have $[q_t^r, q_t'^r] \subset \overline{\mathrm{I}(y_t)}$, which implies (see Figure 1)

$$\begin{aligned}
\max_{q\in[\overline{q}_t,q_t]}\Big\{\rho_{y_t}\big(q,\mathrm{I}(y_t)\big)\Big\} &\le \max_{q\in[\overline{q}_t,q_t]}\Big\{\rho_{y_t}\big(q,[q_t^r,q_t'^r]\big)\Big\}\\
&= \max\Big\{\rho_{y_t}\big(q_t,q_t^r\big),\rho_{y_t}\big(\overline{q}_t,q_t'^r\big)\Big\}\\
&\le \beta + K\nu,
\end{aligned}$$

where at the second line we used that the maximum is attained at one of the extrema of the segment. Thus, Lemma 3.7(i) gives that $B_x(r+\beta)\cap \mathrm{I}(x)$ is $(K\beta + K^2\nu)$-semiconvex for any $\beta, \nu > 0$ such that $\beta + K\nu < \overline{\mu}/K$. We conclude by letting $\nu \downarrow 0$.

Step 2: *If all $\mathrm{I}(x) \cap B_x(r+\beta)$ are A-radial-semiconvex, then they are $(A/2)$-radial-semiconvex.*

We want to prove that the following holds: there exists $\beta_0 > 0$ small such that, if for some $A > 0$ the sets $\mathrm{I}(x) \cap B_x(r+\beta)$ are A-radial-semiconvex for all $x \in M$ and $\beta < \beta_0$, then they are indeed $(A/2)$-radial-semiconvex. To this aim, by the results in Appendix A, we need to prove that there exists $\nu > 0$ sufficiently small such that for every $\beta \in (0, \beta_0)$

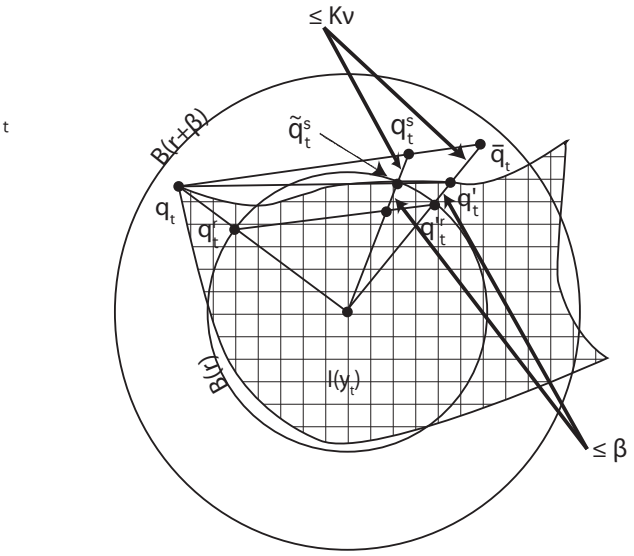


FIGURE 1. Definitions

(β_0 to be fixed later, independently of A) and $v_0, v_1 \in B_x(r+\beta) \cap \mathrm{I}(x)$ with $|v_0 - v_1|_x < \nu$, we have

$$\rho_x\big(v_t, \mathrm{I}(x)\big) \le \frac{A}{2K^*} \frac{t(1-t)}{2} \big|v_0 - v_1\big|^2 \qquad \forall\, t \in [0,1],$$

where K^* is given by Proposition A.4. Let $v_0, v_1 \in \mathrm{I}(x)$ and $\nu > 0$ with $|v_1 - v_0|_x < \nu$, and for $t, s \in [0,1]$ set $q^s_t := (1-s)\overline{q}_t + s q_t$ and denote by $\widetilde{q}^s_t$ the intersection of the segments $[0, q^s_t]$ and $[q_t, q'_t]$ (see Figure 1). We have (by Lemmas 2.1 and 2.2)

$$\begin{aligned}
\rho_{y_t}\big(q^s_t, \mathrm{I}(y_t)\big) &\le \rho_{y_t}\big(q^s_t, \widetilde{q}^s_t\big) + \rho_{y_t}\big(\widetilde{q}^s_t, \mathrm{I}(y_t)\big) \\
&\le \rho_{y_t}\big(\overline{q}_t, q'_t\big) + \rho_{y_t}\big(\widetilde{q}^s_t, \mathrm{I}(y_t)\big) \\
&= \rho_{y_t}\big(\overline{q}_t, \mathrm{I}(y_t)\big) + \rho_{y_t}\big(\widetilde{q}^s_t, \mathrm{I}(y_t)\big) \\
&\le K \rho_x\big(v_t, \mathrm{I}(x)\big) + \rho_{y_t}\big(\widetilde{q}^s_t, \mathrm{I}(y_t)\big) \\
&\le K^2 \left(\frac{|v_t|^2_x}{2} - \frac{d(x, y_t)^2}{2} \right) + \rho_{y_t}\big(\widetilde{q}^s_t, \mathrm{I}(y_t)\big).
\end{aligned}$$

Therefore, for every $t \in [0,1]$ we get

$$\max_{q \in [\overline{q}_t, q_t]} \Big\{ \rho_{y_t}\big(q, \mathrm{I}(y_t)\big) \Big\} \le K^2 \left(\frac{|v_t|^2_x}{2} - \frac{d(x, y_t)^2}{2} \right) + \max_{\widehat{q} \in [q_t, q'_t]} \Big\{ \rho_{y_t}\big(\widehat{q}, \mathrm{I}(y_t)\big) \Big\}. \tag{3.10}$$

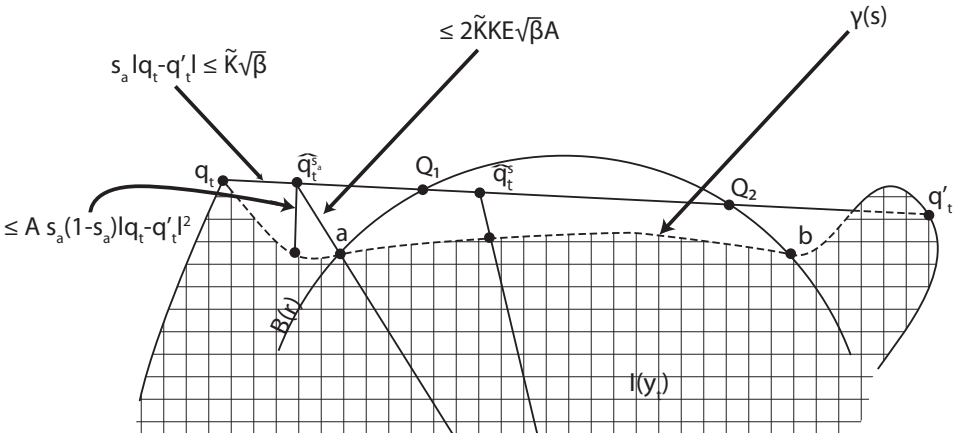


FIGURE 2. Estimations

Set for every $t, s \in [0,1]$, $\widehat{q}_t^s := (1-s)q'_t + sq_t$. By the A-radial-semiconvexity we have

$$\rho_{y_t}\big(\widehat{q}_t^s, \mathrm{I}(y_t)\big) \leq A\frac{s(1-s)}{2}|q_t - q'_t|^2_{y_t}. \tag{3.11}$$

Then, we finally obtain for $\nu > 0$ small enough,

$$\sup_{q\in[\overline{q}_t, q_t]} \Big\{\rho_{y_t}\big(q, \mathrm{I}(y_t)\big)\Big\} \leq \min\left\{K^2\left(\frac{|v_t|_x^2}{2} - \frac{d(x,y_t)^2}{2}\right) + A|q_t - q'_t|^2_{y_t}, \overline{\mu}\right\},$$

for every $t \in [0,1]$. Two cases may appear:

First case: $|q_t - q'_t|^2_{y_t} \leq 1/(2KK^*)$.

In this case, by Lemma 3.7(ii) we deduce that $\mathrm{I}(x) \cap B_x(r+\beta)$ is $(A/2)$-radial-semiconvex.

Second case: $|q_t - q'_t|^2_{y_t} > 1/(2KK^*)$.

We work in the plane generated by $0, q_t, q'_t$ in $T_{y_t}M$, and we define the curve $\gamma : [0,1] \to \mathrm{I}(y_t)$ as (see Figure 2)

$$\gamma(s) := w \quad \text{where} \quad \rho_{y_t}\big(\widehat{q}_t^s, \mathrm{I}(y_t)\big) = |\widehat{q}_t^s - w|_{y_t} \qquad \forall\, s \in [0,1],$$

and denote by $a = \gamma(s_a)$ the first point of γ which enters $\overline{B}_{y_t}(r)$ and $b = \gamma(s_b)$ the last one (see Figure 2). Since both q_t, q'_t belong to $B_{y_t}(r+\beta)$ and $|q_t - q'_t|^2_{y_t} > 1/(2KK^*)$, the intersection of the segment $[q_t, q'_t]$ with $B_{y_t}(r)$ is a segment $[Q_1, Q_2]$ such that

$$\big|Q_1 - q_t\big|_{y_t},\ \big|Q_2 - q'_t\big|_{y_t} \leq \widetilde{K}\sqrt{\beta},$$

for some uniform constant $\widetilde{K} > 0$ and $\beta > 0$ small enough. Since

$$\big|q_t - \widehat{q}_t^{s_a}\big|_{y_t} \le \big|Q_1 - q_t\big|_{y_t} \quad \text{and} \quad \big|q'_t - \widehat{q}_t^{s_b}\big|_{y_t} \le \big|Q_2 - q'_t\big|_{y_t},$$

this implies that both s_a and $1 - s_b$ are bounded by $\frac{\widetilde{K}\sqrt{\beta}}{|q_t - q'_t|_{y_t}} < \sqrt{2KK^*}\widetilde{K}\sqrt{\beta}$. Let us distinguish again two cases:

- On $[s_a, s_b]$, $\mathcal{P}(r)$ is true so $[a, b] \subset \overline{\mathrm{I}(y_t)}$. Hence

$$\sup_{q \in [\widehat{q}_t^{s_a}, \widehat{q}_t^{s_b}]} \Big\{\rho_{y_t}\big(q, \mathrm{I}(y_t)\big)\Big\} \le \max\Big\{\rho_{y_t}\big(\widehat{q}_t^{s_a}, \mathrm{I}(y_t)\big), \rho_{y_t}\big(\widehat{q}_t^{s_b}, \mathrm{I}(y_t)\big)\Big\}.$$

- On $[0, s_a]$ (similarly on $[1 - s_b, 1]$), the A-radial-semiconvexof $\overline{B}_{y_t}(r + \beta) \cap \overline{\mathrm{I}(y_t)}$ yields (by (3.11))

$$\rho_{y_t}\big(\widehat{q}_t^s, \mathrm{I}(y_t)\big) \le A\frac{s(1-s)}{2}|q_t - q'_t|^2_{y_t} \le As_a|q_t - q'_t|^2_{y_t} \le \sqrt{2KK^*}\widetilde{K}E\sqrt{\beta}A,$$

where we used that $|q_t - q'_t|^2_{y_t} \le E$ for some uniform constant $E > 0$. Recalling (3.10) we obtain

$$\sup_{q \in [q_t, \overline{q}_t]} \Big\{\rho_{y_t}\big(q, \mathrm{I}(y_t)\big)\Big\} \le K^2\left(\frac{|v_t|^2_x}{2} - \frac{d(x, y_t)^2}{2}\right) + \sqrt{2KK^*}\widetilde{K}E\sqrt{\beta}A.$$

Hence, if we choose β_0 sufficiently small so that $\sqrt{2KK^*}\widetilde{K}E\sqrt{\beta_0} \le 1/(2KK^*)$ we get

$$\sup_{q \in [q_t, \overline{q}_t]} \Big\{\rho_{y_t}\big(q, \mathrm{I}(y_t)\big)\Big\} \le K^2\left(\frac{|v_t|^2_x}{2} - \frac{d(x, y_t)^2}{2}\right) + \frac{A}{2KK^*},$$

and we conclude again by Lemma 3.7(ii).

As explained above, combining Steps 1 and 2 we infer that, for $\beta > 0$ small enough, all the $\mathrm{I}(x) \cap B_x(r + \beta)$ are convex. This shows that the interval J is open in $[0, \infty)$, concluding the proof of Lemma 3.8 and in turn the proof of Theorem 1.6. □

As we will see in the next section, we can extract from the proof of Theorem 1.6 some ideas which will allow us to treat the case of Riemannian manifolds which do not satisfy the nonfocality assumption. Such a result will play a major role in [6].

4. General version of the proof of Theorem 1.6

Let Z be a compact subset in TM whose fibers are denoted by $Z(x)$. We say that the extended Ma–Trudinger–Wang condition $\big(\overline{\mathbf{MTW}}(-D\rho, C)\big)$ holds on Z if there are constants $C, D > 0$ such that, for any $(x, v) \in TM$ with $v \in Z(x)$,

$$\overline{\mathfrak{S}}_{(x,v)}(\xi, \eta) \ge -C\,|\langle \xi, \eta\rangle_x|\,|\xi|_x|\eta|_x - D\rho_x(v, \mathrm{I}(x))|\xi|^2_x|\eta|^2_x \qquad \forall\, \xi, \eta \in T_xM.$$

The following improvement of Theorem 1.6 can be proved by the same method. Note that we do need assume the manifold to be nonfocal.

Theorem 4.1. *Let (M,g) be a smooth compact Riemannian manifold, assume that the following property holds: For every $r>0$ such that $B_x(r)\cap \mathrm{I}(x)$ is convex for all $x\in M$, there are $\overline{\beta}(r)>0$ and a compact set $Z\subset TM$ with radial fibers (cf. Definition A.2) satisfying the following properties:*

(1) There are $C,D>0$ such that $(\overline{\mathbf{MTW}(-D\rho,C)})$ holds on Z.
(2) There is $K>0$ such that

$$\rho_x\big(v,\mathrm{I}(x)\big)\le K\left(|v|_x^2-d\big(x,\exp_x(v)\big)^2\right)\qquad\forall (x,v)\in Z.$$

(3) $\forall\, x\in M,\forall\,\beta\in(0,\overline{\beta}(r))$, $\mathrm{I}(x)\cap B_x(r+\beta)\subset Z(x)\subset\overline{\mathrm{NF}}(x)$.
(4) $\forall\, x\in M,\forall\,\beta\in(0,\overline{\beta}(r)),\forall\, v_0,v_1\in \mathrm{I}(x)\cap B_x(r+\beta)$, $v_t\in Z(x)$ and $[q_t,\overline{q}_t]\subset Z(y_t)$.

Then all injectivity domains of M are convex.

To prove Theorem 4.1, we will need the following refined version of Lemma 3.4.

Lemma 4.2. *Let $h:[0,1]\to[0,\infty)$ be a semiconvex function such that $h(0)=h(1)=0$ and let $c,C>0$ be fixed. Assume that there are $t_1<\ldots<t_N$ in $(0,1)$ such that h is not differentiable at t_i for $i=1,\ldots,N$, is of class C^2 on $(0,1)\setminus\{t_1,\ldots,t_N\}$, and satisfies*

$$\ddot{h}(t)\ge -C|\dot{h}(t)|-c\qquad\forall\, t\in[0,1]\setminus\big\{t_1,\ldots,t_N\big\}. \tag{4.1}$$

Then

$$h(t)\le 4ce^{(1+C)}t(1-t)\qquad\forall\, t\in[0,1]. \tag{4.2}$$

Proof of Lemma 4.2. Given $\mu,\lambda>0$, denote by $f_{\mu,\lambda}:[0,1]\to\mathbb{R}$ the semiconvex function defined

$$f_{\mu,\lambda}(t):=h(t)-\mu\min\left\{1-e^{-\lambda t},1-e^{-\lambda(1-t)}\right\}\qquad\forall\, t\in[0,1].$$

Let $\bar{t}$ be a maximum point for $f_{\mu,\lambda}$. Since $f_{\mu,\lambda}$ is semiconvex, it has to be differentiable at $\bar{t}$, so $\bar{t}\neq 1/2$ and $\bar{t}\neq t_i$ for $i=1,\ldots,N$. If $\bar{t}\in(0,1/2)$, then there holds $\dot{f}_{\mu,\lambda}(\bar{t})=0$ and $\ddot{f}_{\mu,\lambda}(\bar{t})\le 0$. Then using (4.1), we get

$$|\dot{h}(\bar{t})|=\mu\lambda e^{-\lambda\bar{t}},$$

$$0\ge\ddot{f}_{\mu,\lambda}(\bar{t})=\ddot{h}(\bar{t})+\mu\lambda^2e^{-\lambda\bar{t}}\ge -C|\dot{h}(\bar{t})|-c+\mu\lambda^2e^{-\lambda\bar{t}}\ge\mu\lambda(\lambda-C)e^{-\lambda/2}-c.$$

This yields a contradiction provided we choose $\lambda=1+C$ and $\mu=2ce^{1+C}/(1+C)$ and implies that $f_{\mu,\lambda}$ attains its maximum at $t=0$. Repeating the same argument on $[1/2,1]$, since $f(0)=f(1)=0$ we infer that

$$h(t)\le 2ce^{(1+C)}\min\left\{\frac{1-e^{-(1+C)t}}{1+C},\frac{1-e^{-(1+C)(1-t)}}{1+C}\right\}\qquad\forall\, t\in[0,1].$$

Noting that

$$\frac{1-e^{-(1+C)t}}{1+C} \le t \quad \text{and} \quad \min\{t,1-t\} \le 2t(1-t) \qquad \forall\, t \in [0,1],$$

we get the result. □

We are ready to give the proof of Theorem 4.1.

Proof of Theorem 4.1. Let $x \in M$ and $v_0, v_1 \in \mathrm{I}(x)$ be fixed. We keep the same notation as in Section 3.

The following result is a variant of Lemma 3.7.

Lemma 4.3. *Let $r > 0$ be such that $B_x(r) \cap \mathrm{I}(x)$ is convex for all $x \in M$ and $\overline{\beta}(r)$ given by the hypothesis in Theorem 4.1. There exist $\overline{K}$, such that if*

$$\sup_{q \in [\overline{q}_t, q_t]} \Big\{ \rho_{y_t}\big(q, \mathrm{I}(y_t)\big) \Big\} \le \kappa \qquad \forall\, x \in M,\, \forall\, \beta \in (0, \overline{\beta}(r)),\, \forall\, v_0, v_1 \in \mathrm{I}(x) \cap B_x(r+\beta),$$

then $\mathrm{I}(x) \cap B_x(r+\beta)$ is $(\kappa\overline{K})$-radial-semiconvex.

Proof of Lemma 4.3. We need to show that, for any $v_0, v_1 \in \mathrm{I}(x) \cap B_x(r+\beta)$,

$$\rho_x\Big(v_t, \mathrm{I}(x) \cap B_x(r+\beta)\Big) \le \kappa \overline{K} \frac{t(1-t)}{2} \big|v_0 - v_1\big|^2 \qquad \forall\, t \in [0,1].$$

As in Lemma 3.7 we set

$$h(t) := \frac{|v_t|_x^2}{2} - \frac{d(x,y_t)^2}{2} \qquad \forall\, t \in [0,1]$$

with $v_0, v_1 \in \mathrm{I}(x) \cap B_x(r+\beta)$, and up to slightly perturbing v_0, v_1 we may assume that $h : [0,1] \to \mathbb{R}$ is semiconvex, C^2 outside a finite set of times $0 < t_1 < \ldots < t_N < 1$, and not differentiable at t_i for $i = 1, \ldots, N$. Moreover properties (1) and (3)-(4) in Theorem 4.1 yield

$$\ddot{h}(t) \ge -C|\dot{h}(t)||\dot{y}_t|_{y_t}|q_t - \overline{q}_t|_{y_t} - D \max_{q \in [\overline{q}_t, q_t]} \Big\{ \rho_{y_t}(q, \mathrm{I}(y_t)) \Big\} |\dot{y}_t|_{y_t}^2 |q_t - \overline{q}_t|_{y_t}^2,$$

for every $t \in [0,1] \setminus \{t_1, \ldots, t_N\}$. Since by compactness of M, there is a uniform constant $E > 0$ such that

$$\big|\dot{y}_t\big|_{y_t} \le E \big|v_0 - v_1\big|_x \quad \text{and} \quad \big|q_t - \overline{q}_t\big|_{y_t} \le E,$$

we get

$$\ddot{h}(t) \ge -CE^2 |\dot{h}(t)| \big|v_1 - v_0\big|_x - DE^4 \kappa \big|v_1 - v_0\big|_x^2 \qquad \forall\, t \in [0,1] \setminus \big\{t_1, \ldots, t_N\big\}.$$

Thus Lemma 4.2 gives

$$h(t) \le 4e^{(1+CE^2)} DE^4 \kappa\, t(1-t) |v_1 - v_0|_x^2 \qquad \forall\, t \in [0,1],$$

and so by property (2) in Theorem 4.1 we get that $\mathrm{I}(x) \cap B(r)$ is $(\kappa\overline{K})$-radial-semiconvex with $\overline{K} = 2K 4e^{(1+CE^2)} DE^4$. □

We are ready to apply our bootstrap arguments. We recall that the property $\mathcal{P}(r)$ is satisfied if for any $x \in M$ the set $B_x(r) \cap \mathrm{I}(x)$ is convex. As before, in order to conclude the proof of Theorem 4.1 we just need to prove the following result.

Lemma 4.4. *The set of r for which $\mathcal{P}(r)$ holds is open in $[0, \infty)$.*

Proof of Lemma 4.4. Assume that $\mathcal{P}(r)$ holds. The proof is divided in two steps: first we show that there are $\beta_0, K > 0$ such that, for any $\beta \in (0, \beta_0)$, the sets $B_x(r+\beta) \cap \mathrm{I}(x)$ are $((K+1)\overline{K}\beta)$-radial-semiconvex for any $x \in M$. Then in Step 2 we show the following "bootstrap-type" result: if the sets $B_x(r+\beta) \cap \mathrm{I}(x)$ are A-radial-semiconvex for all $x \in M$, then they are indeed $(A/2)$-radial-semiconvex. As before the combination of Steps 1 and 2 proves the convexity of the $B_x(r+\beta) \cap \mathrm{I}(x)$.

Step 1: $\mathrm{I}(x) \cap B_x(r+\beta)$ *is $((K+1)\overline{K}\beta)$-radial-semiconvex for any $\beta \in (0, \beta_0)$.*

Fix $x \in M$ and $\beta \in (0, \overline{\beta}(r))$. Since $B_x(r+\beta) \cap \mathrm{I}(x)$ is starshaped we can find $v_0^r, v_1^r \in \mathrm{I}(x) \cap B(r)$ with, for $i = (0,1)$, $\rho_x(v_i, v_i^r) \leq \beta$. Thus $\mathcal{P}(r)$ implies that $\rho_x(v_t, I(x)) \leq \beta$ for all $t \in [0,1]$, that is $v_t \in \mathrm{I}^\beta(x)$, and it follows from Lemma 2.2 that $\overline{q}_t \in \mathrm{I}^{K\beta}(y_t)$. By construction we also have

$$|\overline{q}_t|_{y_t} = |v_t|_x < r + \beta, \quad |q_t|_{y_t} \leq |v_t|_x < r + \beta, \quad q_t \in \mathrm{I}(y_t).$$

Since $\overline{q}_t \in \mathrm{I}^{K\beta}(y_t)$ we can find $q_t' \in \overline{\mathrm{I}(y_t)} \cap B_{y_t}(r+\beta)$ such that

$$\rho_{y_t}\big(\overline{q}_t, \mathrm{I}(y_t)\big) = |\overline{q}_t - q_t'| \leq K\beta.$$

Moreover, using that $\mathrm{I}(y_t)$ is starshaped and that $q_t, q_t' \in B_{y_t}(r+\beta)$, we can find q_t^r, ${q'}_t^r \in \overline{B}_{y_t}(r) \cap \overline{\mathrm{I}(y_t)}$ such that $\rho_{y_t}(q_t, q_t^r) \leq \beta$ and $\rho_{y_t}(q_t', {q'}_t^r) \leq \beta$. Again $\mathcal{P}(r)$ implies that $[q_t^r, {q'}_t^r] \subset \overline{\mathrm{I}(y_t)}$, so (see Figure 1)

$$\begin{aligned}
\max_{q \in [\overline{q}_t, q_t]} \Big\{ \rho_{y_t}\big(q, \mathrm{I}(y_t)\big) \Big\} &\leq \max_{q \in [\overline{q}_t, q_t]} \Big\{ \rho_{y_t}\big(q, [q_t^r, {q'}_t^r]\big) \Big\} \\
&= \max\Big\{ \rho_{y_t}\big(q_t, q_t^r\big), \rho_{y_t}\big(\overline{q}_t, {q'}_t^r\big) \Big\} \\
&\leq \beta + K\beta,
\end{aligned}$$

where at the second line we used that the maximum is attained at one of the extrema of the segment. Thus, Lemma 4.3 implies that $B_x(r+\beta) \cap \mathrm{I}(x)$ is $((K+1)\overline{K}\beta)$-radial-semiconvex for any $\beta \in \,]0, \overline{\beta}(r)]$.

Step 2: *If all $\mathrm{I}(x) \cap B_x(r+\beta)$ are A-radial-semiconvex, then they are $(A/2)$-radial-semiconvex.*

Let $v_0, v_1 \in \mathrm{I}(x) \cap B_x(r+\beta)$, as before we define in the plane generated by $0, v_0, v_1$ in T_xM the curve $\gamma : [0,1] \to \mathrm{I}(x)$ by (see Figure 2)

$$\gamma(t) = w \quad \text{where} \quad \rho_x\big(v_t, \mathrm{I}(x)\big) = |v_t - w|_x \qquad \forall\, t \in [0,1],$$

and denote by $a=\gamma(t_a)$ the first point of γ which enters $\overline{B}_x(r)$ and $b=\gamma(t_b)$ the last one. Since both v_0, v_1 belong to $B_x(r+\beta)$ and $B_x(r)\cap \mathrm{I}(x)$ is convex, the intersection of the segment $[v_0,v_1]$ with $B_x(r)$ is a segment $[Q_1,Q_2]$ such that

$$\big|Q_1-v_0\big|,\ \big|Q_2-v_1\big|_x \le \widetilde{K}\sqrt{\beta},$$

for some uniform constant $\widetilde{K}>0$ and $\beta>0$ small enough. Since

$$\big|v_{t_a}-v_0\big|_x \le \big|Q_1-v_0\big|_x \quad \text{and} \quad \big|v_{t_b}-v_1\big|_x \le \big|Q_2-v_1\big|_x,$$

both t_a and $1-t_b$ are bounded by $\frac{\widetilde{K}\sqrt{\beta}}{|v_0-v_1|_x}$. Let us distinguish two cases:
- On $[t_a,t_b]$, $\mathcal{P}(r)$ is true so $[a,b]\subset \overline{\mathrm{I}(x)}$. Then

$$\sup_{v\in[v_{t_a},v_{t_b}]}\Big\{\rho_x\big(v,\mathrm{I}(x)\big)\Big\} \le \max\Big\{\rho_x\big(v_{t_a},\mathrm{I}(x)\big),\rho_x\big(v_{t_b},\mathrm{I}(x)\big)\Big\}.$$

- On $[0,t_a]$ (similarly on $[1-t_b,1]$), $\overline{B}_x(r+\beta)\cap\overline{\mathrm{I}(x)}$ is A-radial-semiconvex, so

$$\rho_x\big(v_t,\mathrm{I}(x)\big) \le A\frac{t(1-t)}{2}|v_1-v_0|_x^2 \le AE\widetilde{K}\sqrt{\beta}.$$

Combining these two estimates we get, for all $t\in[0,1]$,

$$\rho_x\big(v_t,\mathrm{I}(x)\big) \le A\frac{t(1-t)}{2}|v_1-v_0|_x^2 \le AE\widetilde{K}\sqrt{\beta}.$$

Then we define as above q_t' such that $\rho_{y_t}\big(\overline{q}_t,\mathrm{I}(y_t)\big)=|\overline{q}_t-q_t'|$. By Lemma 2.2 we get

$$\begin{aligned}\max_{q\in[\overline{q}_t,q_t]}\Big\{\rho_{y_t}\big(q,\mathrm{I}(y_t)\big)\Big\} &\le \rho_{y_t}\big(\overline{q}_t,\mathrm{I}(y_t)\big)+\max_{\widehat{q}\in[q_t',q_t]}\Big\{\rho_{y_t}\big(\widehat{q},\mathrm{I}(y_t)\big)\Big\}\\ &\le KAE\widetilde{K}\sqrt{\beta}+\max_{\widehat{q}\in[q_t',q_t]}\Big\{\rho_{y_t}\big(\widehat{q},\mathrm{I}(y_t)\big)\Big\}.\end{aligned}$$

Since $\overline{B}_x(r+\beta)\cap\overline{\mathrm{I}(x)}$ is A-radial-semiconvex for every $x\in M$, the same argument used above for $[v_0,v_1]$ is also valid on each segment $[q_t',q_t]$, hence

$$\max_{q\in[\overline{q}_t,q_t]}\Big\{\rho_{y_t}\big(q,\mathrm{I}(y_t)\big)\Big\} \le KAE\widetilde{K}\sqrt{\beta}+AE\widetilde{K}\sqrt{\beta}$$

Therefore, if we choose $\overline{\beta}(r)>0$ sufficiently small we get

$$\sup_{q\in[q_t,\overline{q}_t]}\big\{\rho_{y_t}\big(q,\mathrm{I}(y_t)\big)\big\} \le \frac{A}{2\overline{K}},$$

and in turn, by Lemma 4.3,

$$\rho_x\big(v_t,\mathrm{I}(x)\big) \le \frac{A}{2}\frac{t(1-t)}{2}\big|v_0-v_1\big|_x^2 \qquad \forall\, t\in[0,1],$$

which proves the $(A/2)$-radial-semiconvex. □

The proof of Lemma 4.4 concludes the proof of Theorem 4.1. □

We leave the reader to check that if (M, g) is nonfocal, then the properties in Theorem 4.1 are satisfied (take $Z = \mathrm{I}^{\overline{\mu}}$ which was defined in Section 3). As a consequence, Theorem 1.6 can be seen as a corollary of Theorem 4.1.

5. Conclusion and perspectives

We can develop our proof further to cover all the results obtained in [21], namely modifying just a bit Lemma 4.2 we can prove that $\mathbf{(MTW}(\kappa_0, \infty)\mathbf{)}$ for $\kappa_0 > 0$ gives κ-uniform convexity for some $\kappa > 0$. For a definition of κ uniform convexity we refer to [21] or Appendix A.

Lemma 5.1. *[Modified lemma] Let $h : [0,1] \to [0,\infty)$ be a semiconvex function such that $h(0) = h(1) = 0$ and let $c, C > 0$ be two fixed constants. Assume that there are $t_1 < \ldots < t_N$ in $(0,1)$ such that h is not differentiable at t_i for $i = 1, \ldots, N$, is of class C^2 on $(0,1) \setminus \{t_1, \ldots, t_N\}$, and satisfies*

$$\ddot{h}(t) \geq -C|\dot{h}(t)| + c \qquad \forall\, t \in [0,1] \setminus \big\{t_1, \ldots, t_N\big\}. \tag{5.1}$$

Then

$$h(t) \leq -4ce^{(1+C)}t(1-t) \qquad \forall\, t \in [0,1]. \tag{5.2}$$

It leads to the following theorem.

Theorem 5.2. *Let (M, g) be a nonfocal Riemannian manifold satisfying $\mathbf{(MTW}(\kappa_0, \infty)\mathbf{)}$, with $\kappa_0 > 0$. Then there exists $\kappa > 0$ such that all injectivity domains of M are κ uniformly convex.*

Sketch of the proof . Thanks to Theorem 1.6 we know that for all $x \in M$, $\mathrm{I}(x)$ is convex. Therefore we can define $\forall v_0, v_1 \in \partial \mathrm{I}(x)$, $v_t = (1-t)v_0 + tv_1 \in \mathrm{I}(x)$, $q_t = \exp_x(t_c(v_t)v_t)$ and

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

According to [21] we deduce from $\mathbf{(MTW}(\kappa_0, \infty)\mathbf{)}$ that M satisfies $\mathbf{(MTW}(\kappa_0, C)\mathbf{)}$, where $C > 0$. We conclude thanks to Lemmas 3.6 and 5.1. □

Theorem 4.1 is very general, it can be extended to κ uniform convexity. We only need to find a domain satisfying the control condition (1)–(4) of Theorem 4.1. For this construction we face two difficulties located around the purely focal points, the first one is to give a sign to the extended tensor near these points, The second one is to isolate them. To be done we need to better understand the repartition of purely focal points, and the behavior of the tensor near them. We adopt this strategy for an analytic manifold of dimension 2 in [6]. If one succeed in proving the Villani's conjecture, it will give a very nice formulation of necessary and sufficient conditions for regularity of optimal transport maps [11].

A. Semiconvexity

Following [21] we recall several equivalent definitions for semiconvex functions.

Definition A.1 (Semiconvexity)**.** Let O be a convex subset of $\mathbb{R}^n$. A function $f : O \to \mathbb{R}$ is said to be δ-semiconvex if equivalently, for any x, y in $\mathbb{R}^n$ and t in $[0,1]$,

(i) $f((1-t)x + t(y)) \leq (1-t)f(x) + tf(y) + \delta t(1-t)\frac{|x-y|^2}{2}$;
(ii) $f + \delta\frac{|x|^2}{2}$, is convex;
(iii) $\nabla^2 f \geq -\delta$.

Here (iii) has to be understood in a distributional sense where f is not differentiable. The equivalent of (i), (ii), and (iii) is a classical convexity result. Note that (iii) tells us that as convexity, semiconvexity may be seen as a local property. When $\delta < 0$ we find the uniform convexity.

Definition A.2. An open set $V \subset \mathbb{R}^{n+1}$ is a Lipschitz radial set if it is starshaped around 0 and its boundary is Lipschitz.

Here and in the sequel, ρ denotes the radial distance as defined in Section 2.

Definition A.3. A radial set V is said to be

- δ**-distance-semiconvex** if $\mathrm{dist}(\cdot, \overline{V})$ is δ-semiconvex, that is for any $x, y \in V$, the function $h(t) := \mathrm{dist}((1-t)x + ty, \overline{V})$ is δ-semiconvex on $[0,1]$.
- **locally** δ**-distance-semiconvex** if there exists $\nu > 0$ such that for any $x, y \in V$ with $|x-y| < \nu$, the function $h(t) := \mathrm{dist}((1-t)x+ty, \overline{V})$ is δ-semiconvex on $[0,1]$.
- δ**-radial-semiconvex** if ρ is δ semiconvex, that is for any $x, y \in V$ the function $h(t) := \rho((1-t)x + ty, \overline{V})$ is δ-semiconvex on $[0,1]$.
- **locally** δ**-radial-semiconvex** if there exists $\nu > 0$ such that for any $x, y \in V$ with $|x-y| < \nu$, the function $h(t) := \rho((1-t)x + ty, \overline{V})$ is δ-semiconvex on $[0,1]$.

These definitions are very much inspired by the definition of κ–uniform convexity in [21]. To obtain both notions in one definition one need to consider the signed distance function $\mathrm{dist}_{sign}(\cdot, \partial V)$ (resp. $\rho_{sign}(\cdot, \partial V)$) instead of $\mathrm{dist}(\cdot, \overline{V})$ (resp. ρ): we take the distance with the negative sign when we are inside V.

Proposition A.4. *If a radial set V is (locally) δ-distance-semiconvex then it is (locally) $K^*\delta$-radial-semiconvex for some $K^* > 0$. Reciprocally if V is (locally) δ-radial-semiconvex then it is (locally) δ-distance-semiconvex.*

Proof. Equation (A.4) of [21] provides a constant $K^* > 0$ depending on the dimension, the Lipschitz regularity, and the diameter of V, such that

$$\mathrm{dist}(\cdot, \partial V) \leq \rho(\cdot, \partial V) \leq K^* \mathrm{dist}(\cdot, \partial V).$$

□

Proposition A.5. *If a radial set V is 0-radial-semiconvex then it is convex.*

Proof. For any $x, y \in V$ we have $\rho((1-t)x + ty, \overline{V}) \leq 0$, that is $[x, y] \in \overline{V}$. □

Proposition A.6. *Let V be a radial set which is locally δ-distance-semiconvex, then V is δ-distance-semiconvex. If V is locally δ-radial-semiconvex then V is $K^*\delta$-radial-semiconvex, where K^* is given by Proposition A.4.*

Proof. The first assertion can be deduced from Proposition A.4 of [21]. The second follows from our Proposition A.4 □

B. The tangent cut loci are Lipschitz continuous

B.1. **Introduction.** Let (M, g) be a smooth compact Riemannian manifold of dimension $n \geq 2$. We know that the function t_{cut} defined in Section 2 is bounded from below by the injectivity radius of M and bounded from above by the diameter of M.

In the spirit of the definition of t_{cut} and t_f we define, for any subset O of TM with starshaped fibers, the boundary function $t_b : UM \to \mathbb{R}^+$ by

$$t_b(x, v) := \sup\Big\{t \geq 0 \,|\, tv \in O_x\Big\}$$

We then give the notion of κ–Lipschitz continuity for O.

Definition B.1 (κ–Lipschitz continuity)**.** Let $O \subset TM$ be such that, for any $x \in M$, the fiber O_x is starshaped. The set O is κ–Lipschitz continuous if for any $(\overline{x}, \overline{v}) \in UM$, there exists a κ–Lipschitz continuous function τ defined on a neighbourhood in UM of $(\overline{x}, \overline{v})$ such that $t_b(x, v) \leq \tau(x, v)$ and $t_b(\overline{x}, \overline{v}) = \tau(\overline{x}, \overline{v})$, where t_b is the boundary function for O.

This definition implies that the boundary of O_x is locally a κ–Lipschitz continuous function. Our aim is to prove the following theorem:

Theorem B.2 (Lipschitz continuity of the tangent cut loci)**.**

(1) There exists $\kappa > 0$ such that for each $x \in M$ the set $I(x)$ is κ-Lipschitz continuous. Moreover for any $(x, v) \in UM$ and $(y, w) \in U_{\exp_x(\mathbb{R}v)}M$ we have

$$|t_{cut}(y, w) - t_{cut}(x, v)| \leq \kappa\, d_{TM}\left((x, v), (y, w)\right).$$

We call this property the Lipschitz continuity in the geodesic direction.

(2) If M satisfies $\delta(TM) > 0$ then there exists $\kappa > 0$ such that $\{(x, p) \,|\, x \in M,\, p \in I(x)\}$ is κ-Lipschitz continuous.

(3) If M has dimension 2 then there exists $\kappa > 0$ such that $\{(x, p) \,|\, x \in M,\, p \in I(x)\}$ is κ-Lipschitz continuous.

To prove this theorem, we first prove the two following results:

Theorem B.3 (Lipschitz continuity of the tangent focal loci)**.** *There exists a constant κ such that $\{(x, p) \,|\, x \in M,\, p \in \mathrm{NF}(x)\}$ is κ-Lipschitz continuous.*

Theorem B.4 (Semiconcavity of the tangent focal loci)**.** *The set* $\{(x,p)\,|\,x\in M,\ p\in \mathrm{NF}(x)\}$ *is semiconcave.*

The definition of semiconcavity is similar as the definition B.1, where we ask τ to be semiconcave instead of Lipschitz continuous.

Remark B.5. The first item of Theorem B.2 is a result due to Li-Nirenberg, Itoh-Tanaka, and Castelpietra-Rifford [16, 19, 4], while the second and third ones are new.

B.2. **Proof of Theorem B.3: Lipschitz continuity of the tangent focal loci.** The proof uses the Hamiltonian structure hidden in the Jacobi field equation. It is based on the one given in the paper of Castelpietra and Rifford [4], the main difference is that we adopt here a Lagrangian point of view whereas Castelpietra and Rifford used an Hamiltonian point of view.

B.2.1. *Focalization and Jacobi fields.* Let $(x,v)\in TM$, and consider the geodesic path $\gamma_0 : t\in\mathbb{R}^+\mapsto \exp_x(tv)$. We choose an orthonormal basis of T_xM given by $(v,e_2,...,e_i,...,e_n)$ and define by parallel transport an orthonormal basis of $T_{\exp_x(tv)}M$:

$$B(t)=(e_1(t),e_2(t),...,e_i(t),...,e_n(t))\,.$$

We identify $T_{\exp_x(tv)}M$ with $\mathbb{R}^n$ thanks to the basis $B(t)$. By definition the Jacobi field equation along γ_0 is given by [14, 24]

$$\begin{aligned}&\ddot J(t)+R(t)J(t)=0,\quad t\in\mathbb{R}^+,\\ &J(0)=h,\quad h\in T_xM,\\ &\dot J(0)=p,\quad p\in T_xM,\end{aligned}\tag{B.1}$$

where $R(t)$ is the symmetric operator given, in the basis $B(t)$, by $R(t)_{ij}=\langle R(e_i,e_j)e_i,e_j\rangle$, where R is the Riemann tensor. The Jacobi fields describe how a small perturbation of the geodesic path evolves along it. Since a focal point is related to the size of the neighborhood one can "visit" by perturbing the geodesic path, one can understand that both notions are linked. The Jacobi field equation (B.1) is a linear equation of order two, hence we define $J_0^1 : t\mapsto M_n(\mathbb{R})$ as the solution of the following matricial Jacobi field equation

$$\begin{aligned}&\ddot J(t)+R(t)J(t)=0,\quad t\in\mathbb{R}^+,\\ &J(0)=I_n,\\ &\dot J(0)=0.\end{aligned}$$

We similarly define J_1^0 as the solution of

$$\begin{aligned}&\ddot J(t)+R(t)J(t)=0,\quad t\in\mathbb{R}^+,\\ &J(0)=0,\\ &\dot J(0)=I_n.\end{aligned}$$

Any solution J of the Jacobi field equation (B.1) can be written for any $t \in \mathbb{R}^+$

$$J(t) = J_0^1(t)J(0) + J_1^0(t)\dot{J}(0). \tag{B.2}$$

Let us now exhibit two very particular families of Jacobi fields. For any $h \in T_xM$ we define the path

$$\gamma_\alpha(s,t) = \exp_{\exp_x(sh)}(tv), \qquad (s,t) \in [0,1] \times \mathbb{R}^+, \tag{B.3}$$

$$\gamma_\beta(s,t) = \exp_x(t(v+sh)), \qquad (s,t) \in [0,1] \times \mathbb{R}^+. \tag{B.4}$$

It leads to the following families of Jacobi fields

$$J_\alpha(t) := \left.\frac{d}{ds}\right|_{s=0} \gamma_\alpha(s,t) = (d_x \exp_\cdot(tv)) \cdot (h), \tag{B.5}$$

$$J_\beta(t) := \left.\frac{d}{ds}\right|_{s=0} \gamma_\beta(s,t) = (d_{p=tv} \exp_x) \cdot (th). \tag{B.6}$$

Notice that the Jacobi field J_β is nothing but $J_1^0(\cdot)h$, since $J_\beta(0) = 0$ and $\dot{J}_\beta(0) = h$. Analogously the Jacobi field J_α is equal to $J_0^1(\cdot)h$. The link with focalization is enclosed in the following lemma.

Lemma B.6. *Let $(x,v) \in U_xM$ then*

$$t_f(x,v) = \inf\left\{t \in \mathbb{R}^+ \,|\, \exists\, q \in U_xM \text{ with } J_1^0(t)q = 0.\right\} \tag{B.7}$$

The direction q is called a focal direction at (x,v).

Proof. The proof is a direct consequence of (B.6): for any $t > 0$, $J_1^0(t)h = (d_{p=tv} \exp_x) \cdot (th)$. □

B.2.2. *Proof of Theorem B.3.* We start with some remarks on the symplectic structure coming with a Riemannian manifold.

Definition B.7 (The symplectic form)**.** Let M be a Riemannian manifold of dimension n. For any $x \in M$ we fix a base $\mathcal{B}$ of $(T_xM \times T_xM)$, and we define the symplectic form σ as

$$\begin{aligned} \sigma : \quad & (T_xM \times T_xM)^2 \to \mathbb{R}, \\ & (h,q)_{\mathcal{B}}, (h',q')_{\mathcal{B}} \mapsto \langle h, q' \rangle - \langle h', q \rangle = (h,q)^t \, \mathbb{J} \, (h',q'), \end{aligned}$$

where the matrix $\mathbb{J} = \begin{pmatrix} 0 & I_n \\ -I_n & 0 \end{pmatrix}$. A change of coordinates given by a matrix P is symplectic if $P^t \mathbb{J} P = J$. In this case in the new base $\mathcal{B}'$ we have

$$\begin{aligned} \sigma : \quad & (T_xM \times T_xM)^2 \to \mathbb{R}, \\ & (a,b)_{\mathcal{B}'}, (a',b')_{\mathcal{B}'} \mapsto \langle a, b' \rangle - \langle a', b \rangle = (a,b)^t P^t \, \mathbb{J} \, P(a',b'), \end{aligned}$$

Definition B.8 (Lagrangian subspace.)**.** A subspace $L \in T_xM \times T_xM$ is said to be Lagrangian if $\dim(L) = n$ and $\sigma|_{L\times L}$ is equal to 0.

For example the vertical subspace $\{0\}\times T_xM \subset T_xM\times T_xM$ and the horizontal subspace $T_xM \times \{0\} \subset T_xM \times T_xM$ are Lagrangian. The matrix J_1^0 and J_0^1 are the fundamental solutions of the Jacobi field equation (B.1) on those subspaces.

Lemma B.9. *Let L be a Lagrangian subspace and E, F be two vectorial spaces of dimension n such that $E \overset{\perp}{\oplus} F = T_xM \times T_xM$ and the change of coordinates matrix is symplectic. Suppose that $L \cap E \times \{0\} = \{0\}$. Then there exist a symmetric matrix S such that*

$$L = \left\{ (Sq, q)_{E,F} \mid q \in F \right\}.$$

We say that L is a graph above F.

Proof. The matrix S exists since L has dimension n and no direction in E. To see that S is symmetric we look at the symplectic form on two vectors of L: let $q, q' \in F$. Then by definition

$$\begin{aligned} 0 &= \sigma\left((Sq, q), (Sq', q')\right) \\ &= \langle Sq, q' \rangle - \langle Sq', q \rangle \\ &= \langle Sq, q' \rangle - \langle q, Sq' \rangle . \end{aligned}$$

□

An important link between the symplectic form and the Jacobi field is that the symplectic form is preserved along the flow of the Jacobi field equation.

Lemma B.10. *Let J_1 and J_2 be two solution of the Jacobi field equation* (B.1)*. Then for any $t > 0$*

$$\sigma\left(\left(J_1(t), \dot{J}_1(t)\right), \left(J_2(t), \dot{J}_2(t)\right)\right) = \sigma\left(\left(J_1(0), \dot{J}_1(0)\right), \left(J_2(0), \dot{J}_2(0)\right)\right).$$

Equivalently, defining $M(t) = \begin{pmatrix} J_1(t) & J_2(t) \\ \dot{J}_1(t) & \dot{J}_2(t) \end{pmatrix}$ we have $M^t(t)\mathbb{J}M(t) = \mathbb{J}$. In this case we say that $M(t)$ is symplectic.

We now define a particular Lagrangian subspace in order to find a new formulation for t_f.

Definition B.11. Let $(x, v) \in UM$. We define:

- the horizontal subspace at $\exp_x(tv)$:

$$H_{t,v} := T_{exp_x(tv)M} \times \{0\} \subset T_{exp_x(tv)M} \times T_{exp_x(tv)M}.$$

- the vertical subspace at $\exp_x(tv)$:
$$V_{t,v} := \{0\} \times T_{exp_x(tv)M} \subset T_{exp_x(tv)M} \times T_{exp_x(tv)M}.$$
- the subspace $L_{t,v}$ of initial conditions such that at time t the Jacobi field is equal to 0:
$$L_{t,v} := \left\{(h,q) \in T_xM \times T_xM \,|\, J_0^1(t)h + J_1^0(t)q = 0\right\}.$$

Notice that $L_{t,v}$ can be equivalently defined as $L_{t,v} = M^{-1}(t)V_{t,v}$ where
$$M(t) := \begin{pmatrix} J_0^1(t) & J_1^0(t) \\ \dot{J}_0^1(t) & J_1^0(t) \end{pmatrix}.$$

Proposition B.12. *The space $L_{t,v}$ is a Lagrangian subspace of $T_xM \times T_xM$.*

Proof. Since $M^t(t)\mathbb{J}M(t) = \mathbb{J}$ the matrix $M(t)$ is invertible, therefore $L_{t,v}$ is a vectorial subspace of dimension n.

To see that it is Lagrangian we use that σ is preserved along the flow: let $(h,q),(h',q') \in L_{t,v}$, and denote by $J_{h,q}$ the solution of the Jacobi field equation (B.1) with $J_{h,q}(0) = h$ and $\dot{J}_{h,q}(0) = q$. Then, for any $t > 0$,
$$\begin{aligned}\sigma\left((h,q),(h',q')\right) &= \sigma\Big((J_{h,q}(t),\dot{J}_{h,q}(t)),(J_{h',q'}(t),\dot{J}_{h',q'}(t))\Big)\\ &= \sigma\Big((0,\dot{J}_{h,q}(t)),(0,\dot{J}_{h',q'}(t))\Big) = 0.\end{aligned}$$
□

We can now give a new formulation of Lemma B.6.

Lemma B.13. *Let $(x,v) \in U_xM$. Then*
$$t_f(x,v) = \inf\left\{t \in \mathbb{R}^+ \,|\, L_{t,v} \cap V_{0,v} \neq \{0\}\right\}. \tag{B.8}$$
The set $L_{t,v} \cap V_{0,v}$ is called the focal set at (x,v).

Proof. Let $q \in U_xM \setminus \{0\}$ satisfy $(0,q) \in L_{t,v} \cap V_{0,v}$. Then $J_{0,q}(t) = J_1^0(t)q = 0$ and Lemma B.6 concludes the proof. □

We recall that we identify $T_{\exp_x(tv)}M$ with $\mathbb{R}^n$ through the basis
$$B(t) = (e_1(t),...,e_i(t),...,e_n(t)).$$
According to Lemma B.9 the obstruction to see $L_{t_f(x,v),v}$ as a graph above $V_{0,v}$ comes from the intersection of $L_{t_f(x,v),v}$ with the horizontal space. By definition we have
$$L_{t_f(x,v),v} \cap H_{0,v} = \operatorname{Ker} J_0^1(t_f(x,v)).$$
Let us identify, for any $u \geq 0$, $H_{u,v}$ with $\operatorname{Vect}(e_1'(u),...,e_i'(u),...,e_n'(u))$ and $V_{u,v}$ with $\operatorname{Vect}(f_1(u),...,f_i(u),...,f_n(u))$, where $e_i'(u) = e_i(u) \times \{0\} \in T_{\exp_x(uv)} \times T_{\exp_x(uv)}$ and $f_i(u) = \{0\} \times e_i(u) \in T_{\exp_x(uv)} \times T_{\exp_x(uv)}$. With this notation, without loss of generality

we can suppose there exists an index $l > 1$ such that $\operatorname{Ker} J_0^1(t_f(x,v)) = \operatorname{Vect}(e'_l, ...e'_n)$. Therefore, for any $i \geq l$ we can change $e'_i(u)$ by $f_i(u)$ and $f_i(u)$ by $-e'_i(u)$ to get two new orthonormal spaces of dimension n:

$$E(u) = \operatorname{Vect}\left(e'_1(u), ..., e'_{l-1}(u), f_l(u), ..., f_n(u)\right)$$
$$F(u) = \operatorname{Vect}\left(f_1(u), ..., f_{l-1}(u), -e'_l(u), ..., -e'_n(u)\right).$$

Remark B.14. The change of coordinates is symplectic, that is $P^t \mathbb{J} P = J$, where P is the change of basis matrix. Therefore for any $(z,w),(z',w') \in E \times F$ we have

$$\sigma\left((z,w),(z',w')\right) = \left\langle z, w' \right\rangle - \left\langle z', w \right\rangle.$$

By construction for any $u \geq 0$ we have

(1) $E(u) \overset{\perp}{\oplus} F(u) = T_{\exp_x(uv)} \times T_{\exp_x(uv)}$
(2) $L_{t_f(x,v),v} \cap E(0) = \{0\}$.

Since $L_{u,v'}$ is smooth with respect to (x',v',u), there exist a neighbourhood $O_{x,v,t_f(x,v)} \subset TM \times \mathbb{R}^+$ of $(x,v,t_f(x,v))$ such that, for any $(x',v',t) \in O_{x,v,t_f(x,v)}$, we have

$$L_{t,v'} \cap E(0) = \{0\}. \tag{B.9}$$

Moreover Lemma B.9 implies that there exist a smooth function

$$\begin{aligned} S: \quad & O_{x,v,t_f(x,v)} \to S_n(\mathbb{R}) \\ & \left(x',v',t\right) \mapsto S(t), \end{aligned}$$

such that, for any $w \in F(0)$, we have $S(t)w \in E(0)$ and

$$L_{t,v'} = \left\{ \left(S(t)w, w\right)_{E(0)\times F(0)} \text{ with } w \in F(0) \right\}.$$

Remark B.15. Notice that the matrix $S(t)$ as well as the subspaces $E(u)$ and $F(u)$ depend on (x',v',t), but the indices l used to define $E(u)$ and $F(u)$ for any $(x',v',t) \in O_{x,v,t_f(x,v)}$ only depend on $x, v, t_f(x,v)$.

The following lemma is the key tool to apply later the Implicit Function Theorem.

Lemma B.16. *let $(x,v) \in TM$.*

(1) Let $q \in U_x M$ satisfy $(\{0\}, q) \in L_{t_f(x,v),v} \cap V_{0,v}$. Then $q \in F(0)$ and $q^t\, S(t_f(x,v))q = 0$.
(2) There exists $\delta > 0$ such that, for any $(x,v) \in TM$, $||\dot{S}(t_f(v))|| \geq \delta$.
(3) For any $(x',v',t) \in O_{x,v,t_f(x,v)}$, if $q^t\, S(t)q = 0$ then $t_f(v') \leq t$.

Notice that q is defined only in $T_x M$, but for any x' close to x we can see it also as an element of $T_{x'} M$ thanks to the identification with the coordinates. The dot always stands for the derivative along the Jacobi Field ($\frac{d}{dt}$).

Proof. Let $q \in L_{t_f(x,v),v} \cap V_{0,v}$. Since $L_{t_f(x,v),v} \cap H_{0,v} = \text{Vect}\,(e'_l, ...e'_n)$, using the symplectic form σ we find that $q_i = 0$ for any $i = l, \dots, n$. This gives that $q \in F(0)$. Moreover $S(t_f(x,v))q \in V_{0,v}$ thus $(S(t_f(x,v))q)_i = 0$ for any $i = 1, \dots, l-1$. Consequently $q^t\, S(t_f(x,v))q = 0$.

To compute the derivative with respect to t we again use the symplectic form. Let $(0,z) \in V_{t,v}$ for any t such that $(x,v,t) \in O_{x,v,t_f(x,v)}$. There exists $\phi(t) = (h_t, q_t) = (S(t)w_t, w_t)_{E(0)\times F(0)} \in L_{t,v}$ such that $M(t)\phi(t) = (0,z)$. On one side

$$\begin{aligned}\sigma\left(\phi(t), \dot\phi(t)\right) &= \sigma\left((S(t)w_t, w_t)_{E(0)\times F(0)}, (\dot S(t)w_t + S(t)\dot w_t, \dot w_t)_{E(0)\times F(0)}\right)\\ &= \sigma\left((S(t)w_t, w_t), (S(t)\dot w_t, \dot w_t)\right) + \sigma\left((S(t)w_t, w_t), (\dot S(t)w_t, 0)\right)\\ &= \left\langle \dot S(t)w_t, w_t\right\rangle.\end{aligned}$$

On the other side,

$$\sigma\left(\phi(t), \dot\phi(t)\right) = \sigma\left(M(t)\phi(t), M(t)\dot\phi(t)\right).$$

Since $M(t)\phi(t) = (0,z)$ we have $M(t)\dot\phi(t) = -\dot M(t)\phi(t)$. Moreover

$$\dot M(t) = \begin{pmatrix} 0 & I_n \\ -R(t) & 0\end{pmatrix},$$

thus

$$\begin{aligned}\sigma\left(\phi(t), \dot\phi(t)\right) &= -\sigma\left(M(t)\phi(t), \dot M(t)\phi(t)\right)\\ &= -\sigma\left(M(t)\phi(t), \begin{pmatrix} 0 & I_n \\ -R(t) & 0\end{pmatrix} M(t)\phi(t)\right)\\ &= -\sigma\left((0,z), (-z,0)\right) = -|z|^2.\end{aligned}$$

Hence

$$\left\langle \dot S(t)w_t, w_t\right\rangle = -|z|^2,$$

and by compactness we deduce the existence of a constant $\delta > 0$ such that $||\dot S(t_f(x,v)|| \geq \delta$.

For the third item we reason by contradiction: we take $(x', v', t') \in O_{x,v,t_f(x,v)}$ and suppose that $q^t\, S(t')q = 0$ and $t' < t_f(x', v')$. By definition $q \in V_{0,v'} \cap F(0)$, thus $q_i = 0$ for any $i = l, \dots, n$. Since $t' < t_f(x', v')$ the space $L_{t'(x,v),v}$ is a graph on the horizontal space. More precisely, according to (B.2), for any $t \in (0, t_f(x', v'))$ we have

$$L_{t_f(x,v),v} = \left\{(h, \left(J_1^0(t_f(x,v))\right)^{-1} J_0^1(t)h) \,|\, h \in H_{0,v}\right\}.$$

We denote $\left(J_1^0(t_f(x,v))\right)^{-1} J_0^1(t) = K(t)$. Then the exact same computation done above proves that

$$\left\langle \dot{K}(t)h, h \right\rangle < 0 \qquad \forall\, h \in H_{0,v}.$$

Since $t\left(J_1^0(t_f(x,v))\right)^{-1}$ converges to I_n when t goes to zero, we deduce that for t small enough K is symmetric positive definite.
For any $h \in H_{0,v}$ and $q' \in V_{0,v}$ we denote $h = (h_1, h_2)$, where $h_1 \in H_{0,v} \cap E(0)$, $h_2 \in H_{0,v} \cap F(0)$, and $q' = (q_1', q_2')$ with $q_1' \in V_{0,v} \cap E(0)$ and $q_2' \in V_{0,v} \cap F(0)$. With this notation we have $((h, q') = (h_1, q_2'), (q_1', h_2))_{E\times F}$ and we define the matrices $S_i(t)$, $K_i(t)$, $i = 1, \dots, 4$, such that

$$(q_1', q_2') = (K_1(t)h_1 + K_2(t)h_2, K_3(t)h_1 + K_4(t)h_2)\,,$$

and

$$(h_1, q_2') = \left(S_1(t)q_1' + S_2(t)h_2, S_3(t)q_1' + S_4(t)h_2\right).$$

Since by hypothesis $L_{t',v'}$ is a graph over $H_{0,v}$ and $F(0)$, we deduce that $S_1(t') = K_1^{-1}(t')$ and in particular we see that $K_1(t')$ is invertible. In the focal direction $q \in F(0) \cap V_{0,v'}$ we have $q = (q_1, 0)_{F(0)}$ and

$$0 = q^t\, S(t')\, q = q_1^t\, S_1(t')\, q_1 = h_1^t\, K_1(t')\, h_1,$$

where $h_1(t) = K_1^{-1}(t)q_1$. To get a contradiction we just have to remark that, for any $A > 0$, up to taking $O_{x,v,t_f(x,v)}$ smaller we have that, for any $(x', v', t) \in O_{x,v,t_f(x,v)}$ with $t \le t_f(x', v')$,

$$h_1^t(t)K_1(t)h_1(t) \le -Ah_1^t(t)h_1(t).$$

Also, in the direction (x, v), for any $t \le t_f(x, v)$ we have

$$\left((S_1\,(t)\,q_1, S_3\,(t)\,q_1)\,, (q_1, 0)\right)_{E\times F} = \left((h_1\,(t)\,, 0)\,, (K_1\,(t)\,h_1\,(t)\,, K_3\,(t)\,h_1\,(t))\right) \in L_{t,v}.$$

By definition of q we have $S_1(t)q_1 = h_1(t) \to 0$ when $t \to t_f(x, v)$ and $K_1(t)h_1(t) = q_1$. Assuming with no loss of generality that $K_1(t)$ is diagonal, we see that any eigenvalue $\lambda_i(t)$ corresponding to an eigenvector $q_i \neq 0$ goes to $-\infty$ (notice that it cannot goes to $+\infty$ since we proved that $t \mapsto K(t)$ decreases). Hence, being the eigenvalues continuous with respect to (x', v', t), by further shrinking $O_{x,v,t_f(x,v)}$ if needed we have $h_1^t(t)K_1(t)h_1(t) \le -Ah_1^t h_1$. □

Remark B.17. The last proof just says that, before focalization, when the Lagrangian space $L_{t,v}$ has a vertical component it cannot be at the same time a graph above the horizontal space and above F.

To conclude the proof of Theorem B.3 we apply the Implicit Function Theorem in order to find the function τ needed in Definition B.1. Let $(\overline{x}, \overline{v}) \in UM$ and $q \in U_xM$ be the

focal direction associated. Then the function

$$\begin{aligned}\Psi : \quad & UM \times \mathbb{R}^+ \to \mathbb{R} \\ & (x, v, t) \mapsto q^t S\,(t)\, q\end{aligned}$$

is well defined on a neighbourhood of $(\overline{x}, \overline{v}, t_f(\overline{x}, \overline{v}))$. Moreover $\Psi(\overline{x}, \overline{v}, t_f(\overline{x}, \overline{v})) = 0$ and by Lemma B.16 we have:

$$|\partial_t \Psi(\overline{x}, \overline{v}, t_f(\overline{x}, \overline{v}))| = \left| q^t \dot{S}\,(t_f(x, v))\, q \right| \geq \delta.$$

Hence, by the Implicit Function Theorem we get a function τ defined in a neighborhood $O_{\overline{x},\overline{v}}$ of $(\overline{x}, \overline{v})$ such that $\Psi(x, v, \tau(x, v)) = 0$. By Lemma B.16 we find that $t_f(x, v) \leq \tau(x, v)$, and it only remains to check that τ is Lipschitz continuous. This follows from the fact that, by compactness, there exist $K > 0$ such that

$$\begin{aligned}|d_{\overline{x},\overline{v}}\tau| &= \left| \frac{1}{\partial_t \Psi(\overline{x}, \overline{v}, t_f(\overline{x}, \overline{v}))} d_{\overline{x},\overline{v}}\Psi \right| \\ &\leq \frac{K}{\delta}.\end{aligned}$$

It concludes proof of Theorem B.2.

Remark B.18. This method also proves Theorem B.4, as we easily see that the second differential of τ at (x, v) is bounded from above.

B.3. **Proof of Theorem B.2: Lipschitz continuity of the tangent cut loci.**

Proof of Theorem B.2. Let $x \in M$, $e_v \in U_x M$, $v = t_{cut}(e_v) e_v$. We want to find the function τ needed in Theorem B.1 using the Implicit Function Theorem. The construction of the function τ will depends on x, v, and $\delta(v)$.

B.3.1. *At the intersection with the tangent focal locus.* If $v \in \mathrm{TFL}(x) \cap \mathrm{TCL}(x)$ then $t_{cut}(x, e_v) = t_f(x, e_v)$ and for any $(y, e_w) \in U_x M$ we have $t_{cut}(y, e_w) \leq t_f(y, e_w)$. Notice that by Theorem B.3 the function t_f is κ–Lipschitz continuous, so the choice $\tau = t_f$ works.

B.3.2. *Far from the tangent focal locus.* If $v \notin \mathrm{TFL}(x) \cap \mathrm{TCL}(x)$ then $\delta(v) > 0$. Let $\overline{v} \in \bar{\mathrm{I}}(x)$ such that $|v - \overline{v}| = \delta(v)$ and $\exp_x v = \exp_x \overline{v} = y$. Let $K \subset TM$ be a compact neighborhood of the geodesic path $t \in [0, 1] \mapsto \exp_x(t\overline{v})$ and $0 < \epsilon < t_{inj}$ such that $B(y, \epsilon) \subset K(y)$. For any $\eta \in T_y S$ with $z = \exp_y \eta \in B(y, \epsilon)$, we construct a path $s, t \in [0, \epsilon] \times [0, 1] \mapsto \gamma(s, t)$ satisfying the following conditions for any $(s, t) \in [0, \epsilon] \times [0, 1]$:

(1) $\gamma(0, t) = \gamma(t) = \exp_x(t\overline{v})$.
(2) $\gamma(s, 1) = \exp_y(s\eta) = z_s$.
(3) $\gamma(s, 0) = x$.
(4) $\gamma(\cdot, \cdot) \in C^1([0, 1]^2, M)$.
(5) $(\gamma(s, t), \dot{\gamma}(s, t)) \in K$.

Working in smooth charts this construction is easy to realize. Note that $s \leq \epsilon \leq t_{inj}$ implies that $s \mapsto \exp_y(s\eta)$ is a minimizing geodesic path, therefore $d^2(y, z_s) = s^2$ and $z_s \in B(y, \epsilon)$. However $t \mapsto \gamma(s, t)$ and $s \mapsto \gamma(s, t)$ are not necessarily geodesic paths away from $s = 0$ and $t = 1$. Anyway the first variation formula applied to γ provides a constant K such that

$$\begin{aligned} d^2(x, z_s) \leq A(\gamma(s,t)) &\leq A(\gamma(0,t)) + s\,\langle d_{\overline{v}} \exp_x \overline{v}, \eta\rangle + \frac{s^2}{2}K \\ &\leq d^2(x,y) + s\,\langle d_{\overline{v}} \exp_x \overline{v}, \eta\rangle + \frac{K}{2} d^2(y, z_s). \end{aligned} \tag{B.10}$$

We can similarly add a perturbation of x. Hence we define $u : B(x,\epsilon) \times B(y,\epsilon) \to \mathbb{R}^+$ by

$$\begin{aligned} u(x', z) := d^2(x,y) + \langle d_{\overline{v}} \exp_x \overline{v}, (\exp_y)^{-1}(z)\rangle - \langle \overline{v}, (\exp_x)^{-1}(x')\rangle \\ + K\left(d^2(x,x') + d^2(y,z)\right). \end{aligned} \tag{B.11}$$

Note that if we compare the above expression to the right hand side of (B.10) we have changed $\frac{1}{2}K$ to K; this modification shows that $d^2(x', z) = u(x', z)$ if and only if $z = y$ and $x' = x$ otherwise $d^2(x', z) < u(x', z)$. Moreover u is C^1 and

$$(d_{x'=x, z=y} u) \cdot (\zeta, \eta) = -\langle \overline{v}, \zeta\rangle + \langle d_{\overline{v}} \exp_x \overline{v}, \eta\rangle.$$

By continuity of $\exp_x$ there exits $\epsilon > 0$ such that for any $(x', w) \in B((x,v),\epsilon) \subset TM$, $\exp_{x'}(w) = z \in B(y,\epsilon)$. Let $\gamma(x', w, \theta) = \exp_{x'}(\theta w)$ we define $\Phi : B((x,v),\epsilon) \to \mathbb{R}$ by

$$w \mapsto u(\exp_{x'}(w)) - A(\gamma(x', w, \theta)).$$

According to the first variation formula , Φ is C^1 on $B((x,v),\epsilon)$ and the differential at x, v in the direction ζ, ξ (*i.e.* $x' = \exp_x(r\zeta)$, $w = v + s\xi$) is given by

$$\begin{aligned} (d_{x,v}\Phi)(\zeta,\xi) &= \langle d_{p=\overline{v}} \exp_x \overline{v}, d_{p=v} \exp_x \xi\rangle_y - \langle d_{p=v} \exp_x v, d_{p=v} \exp_x \xi\rangle_y + \langle v - \overline{v}, \zeta\rangle \\ &= \langle q - \overline{q}, \eta\rangle_y + \langle v - \overline{v}, \zeta\rangle, \end{aligned} \tag{B.12}$$

where $d_{\overline{v}} \exp_x \overline{v} = -\overline{q}$, $d_v \exp_x v = -q$ and $d_v \exp_x \xi = \eta$.

The set $O_{x,v} := \{(x', v', t) \in UM \times \mathbb{R}^+ \,|\, (x', tv') \in B((x,v),\epsilon)\}$ is an open subset of $UM \times \mathbb{R}^+$, and $(x, e_v, t_{cut}(e_v)) \in O_{x,v}$. We define Ψ by

$$\begin{aligned} \Psi :&\quad O_{x,v} \to \mathbb{R} \\ &\Psi(x', v', t) \mapsto \Phi(x', tv'). \end{aligned}$$

By definition $\Psi(x, e_v, t_{cut}(e_v)) = u(x,y) - A(\gamma(x, v, \theta)) = 0$ and for $(x', v', t) \neq (x, e_v, t_{cut}(e_v))$ if $\Psi(x', v', t) = 0$ then (B.10) implies

$$d^2(x', \exp_{x'}(tv')) < A(\gamma(x', v', t)),$$

hence $t > t_{cut}(x', v')$. Furthermore we compute

$$\frac{\partial}{\partial t}\Psi(x, e_v, t_{cut}(e_v)) = d_{p=v}\Phi(x, e_v) = \langle q - \overline{q}, -\frac{1}{t_{cut}(e_v)}q\rangle_y.$$

Since the geodesic flow is Lipschitz continuous, there exists $A > 0$ such that

$$\frac{1}{A} \leq |q - \overline{q}|_y \leq A|v - \overline{v}|_x.$$

Since $|q|_y^2 = |\overline{q}|_y^2$, and t_{cut} is bounded by a constant C on TM, we have

$$\frac{1}{t_{cut}(e_v)}|\langle q - \overline{q}, q\rangle_y| = \frac{1}{t_{cut}(e_v)}|q - \overline{q}|^2 \geq \frac{1}{2AC}\delta(v)^2 > 0. \tag{B.13}$$

Therefore

$$\left|\frac{\partial}{\partial t}\Psi(x, e_v, t_{cut}(e_v))\right| \geq \frac{1}{2C'}\delta(v)^2 > 0. \tag{B.14}$$

Consequently we can apply the implicit function theorem to $\Psi(x', v', t) = 0$ at $(x, e_v, t_{cut}(e_v))$, to find a neighborhood of $O'_{x,v} \subset UM$ of (x, e_v) and a function $\tau \in C^1(O'_{x,v}, \mathbb{R}^+)$ such that

$$\forall (x', v') \in O'_{x,v} \quad t_{cut}(x', v') \leq \tau(x', v'), \qquad t_{cut}(x, e_v) = \tau(x, e_v). \tag{B.15}$$

The implicit function theorem also gives the differential of τ:

$$\begin{aligned} d_{x'=x,v'=e_v}\tau(\zeta, \xi) &= -\frac{1}{d_{p=v}\Phi(x, e_v)}d_{x'=x,p=v}\Phi(\zeta, \xi) \\ &= \frac{t_{cut}(e_v)}{\langle q - \overline{q}, q\rangle_y}\left[\langle q - \overline{q}, \eta\rangle_y + \langle v - \overline{v}, \zeta\rangle_x\right] \\ &\leq \frac{C''\left(|\eta|_y + |\zeta|_x\right)}{\delta(v)}. \end{aligned} \tag{B.16}$$

We fix a small constant $\overline{\delta} > 0$ and distinguish two cases.

Case 1: $\delta(v) \geq \overline{\delta}$. In this case (B.16) becomes:

$$|d_{x'=x,v'=e_v}\tau(\zeta, \xi)| \leq \frac{C}{\overline{\delta}}\left(|\zeta| + |\xi|\right).$$

Therefore the function τ is κ Lipschitz–continuous, near (x, e_v), for any $\kappa \leq \frac{C}{2\overline{\delta}}$. In this case we are done. We remark that we proved the Lipschitz continuity of t_{cut} for any perturbation of (x, v), so in particular we obtained also the second item of Theorem B.2 in the case $\delta(v) \geq \overline{\delta}$. So we are only left to understand the case of speeds near a purely focal point.

Case 2: $\delta(v) \leq \overline{\delta}$. In this case v is near a purely focal point, and we need to be slightly more precise regarding the estimate of $|d_{x'=x,v'=e_v}\tau(\zeta,\xi)|$. First of all we can rewrite (B.13) as

$$\left|\frac{\partial}{\partial t}\Psi(x,e_v,t_{cut}(e_v))\right| \geq \frac{1}{2C'}|v-\overline{v}|^2. \tag{B.17}$$

Since the symplectic form is preserved along the Jacobi field we have, for any $t>0$,

$$\begin{aligned}\sigma\left((0,v-\overline{v}),(\zeta,\xi)\right)&\\ = \sigma\left((J_1^0(t)(v-\overline{v}),\dot{J}_1^0(t)(v-\overline{v})),(J_0^1(t)\zeta+J_1^0(t)\xi,\dot{J}_0^1(t)\zeta+\dot{J}_1^0(t)\xi)\right)&\end{aligned} \tag{B.18}$$

thus

$$\begin{aligned}-\langle v-\overline{v},\zeta\rangle_x - \left\langle J_0^1(t)\zeta+J_1^0(t)\xi,\dot{J}_1^0(t)(v-\overline{v})\right\rangle_y =&\\ \left\langle J_1^0(t)(v-\overline{v}),\dot{J}_0^1(t)\zeta+\dot{J}_1^0(t)\xi\right\rangle_y.&\end{aligned} \tag{B.19}$$

A Taylor formula together with the fact that $\exp_x(v)=\exp_x(\overline{v})$ gives that there exists $A\in\mathbb{R}_+$ such that

$$|d_{p=v}\exp_x(v-\overline{v})|_y = |J_1^0(t_{cut}(e_v))(v-\overline{v})|_y \leq A|v-\overline{v}|^2. \tag{B.20}$$

Thus the right hand side of (B.19) is smaller then $A|v-\overline{v}|^2$. Thanks to (B.16), we can show the Lipschitz continuity separately on each variable; we conclude by examining three different cases. The first case is a perturbation along the variable v. The second and third cases deal with a perturbation along the variable x.

- If we only consider a perturbation along the speed ($\zeta=0$) then (B.19) and (B.20) give

$$\left|\left\langle \eta,\dot{J}_1^0(t)(v-\overline{v})\right\rangle_y\right| \leq A|v-\overline{v}|^2. \tag{B.21}$$

Moreover a Taylor formula on $q-\overline{q}=d_{p=v}\exp_x(v)-d_{p=\overline{v}}\exp_x(\overline{v})$ gives, for $\delta(v)$ small enough,

$$\dot{J}_1^0(t)(v-\overline{v}) = q-\overline{q}+o(|v-\overline{v}|^2). \tag{B.22}$$

We deduce that there exist $C>0$ and $\overline{\delta}>0$ such that for any $x\in M$ and $v\in \mathrm{I}(x)$ with $\delta(v)\leq\overline{\delta}$ we have, according to (B.12),

$$\left|d_{x'=x,p=v}\Phi(0,\xi)\right| = \left|\langle\eta,q-\overline{q}\rangle_y\right| \leq C|v-\overline{v}|^2.$$

Together with (B.17) and (B.16), we obtain

$$d_{x'=x,v'=e_v}\tau(0,\xi) \leq \frac{2C'C|v-\overline{v}|^2}{|v-\overline{v}|^2} \leq C,$$

which proves the Lipschitz continuity in the v variable. We recall that the constant C can change in each inequality but is uniform on TM.

We now want to look for the Lipschitz continuity in the x variable.

• If the perturbation ζ is collinear to v ($\zeta = \pm v$) then (B.12) rewrites

$$|d_{x'=x,p=v}\Phi(\zeta,0)| = |\,\langle v, v-\overline{v}\rangle_x\,| = \frac{1}{2}|v-\overline{v}|_x^2.$$

Together with (B.17) in (B.16) we obtain that

$$d_{x'=x,v'=e_v}\tau(\zeta,0) \leq C.$$

This is exactly the Lipschitz continuity at (x,v) in the x variable along the geodesic direction given by v, and this concludes the proof of the first item of Theorem B.2.

• If the perturbation ζ belongs to $\operatorname{Ker} J_0^1(t_{cut}(e_v))$ and $\xi = 0$, then (B.19) becomes

$$-\langle v-\overline{v}, \zeta\rangle_x = \left\langle J_1^0(t_{cut}(e_v))(v-\overline{v}), \dot{J}_0^1(t_{cut}(e_v))\,\zeta \right\rangle_y, \tag{B.23}$$

and together with (B.12) and (B.20) we get $|d_{x'=x,p=v}\Phi(\zeta,0)| \leq A|v-\overline{v}|_x^2$. By this estimate combined to (B.17) in (B.16) we obtain a constant $C>0$ such that

$$d_{x'=x,v'=e_v}\tau(\zeta,0) \leq C.$$

Therefore the function t_{cut} is Lipschitz continuous along these directions.

Notice that in dimension two, for any $(x,v)\in M$ we can take a basis with one direction along e_v and the other one in $\operatorname{Ker} J_0^1(t_{cut}(e_v))$, and we deduce that t_{cut} is Lipschitz continuous on UM. This concludes the proof of Theorem B.2. □

Remark B.19. We do not know if in any dimension the function t_{cut} is Lipschitz continuous on UM. However, for any n–dimensional Riemannian manifold, such that

$$\dim\left[\operatorname{Ker} J_0^1(t_{cut}(e_v))\right] = n-1,$$

we proved that t_{cut} is Lipschitz continuous on UM. It is for example the case of $\mathbb{S}^n$. More generally we proved the following theorem:

Theorem B.20 (Lipschitz continuity of the tangent cut loci II)**.** *There exists $\kappa>0$ such that for each $x\in M$ the set $I(x)$ is κ-Lipschitz continuous. Moreover, for any $(x,v)\in UM$, $\zeta\in\{\operatorname{Ker} J_0^1(t_{cut}(e_v))\}\cup\{\pm v\}$, and $(y,w)\in U_{\exp_x(\mathbb{R}\zeta)}M$, we have*

$$|t_{cut}(y,w) - t_{cut}(x,v)| \leq \kappa\, d_{TM}((x,v),(y,w)).$$

References

[1] E. Bierstone and P.D. Milman. Semianalytic and subanalytic sets. *Inst. Hautes Études Sci. Publ. Math.*, 67:5–42, 1988.

[2] M.A. Buchner. Simplicial structure of the real analytic cut locus. *Proc. Amer. Math. Soc.*, 64:118–121, 1977.

[3] 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, 2004.

[4] 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.

[5] A. Figalli. Regularity of optimal transport maps (after Ma–Trudinger–Wang and Loeper). *Séminaire Bourbaki*. Vol. 2008/2009, Exp. No. 1009.

[6] A. Figalli, T. Gallouët, and L. Rifford. On the convexity of injectivity domains on surfaces. *Work in progress.*

[7] A. Figalli and G. Loeper, Grégoire. C^1 regularity of solutions of the Monge-Ampère equation for optimal transport in dimension two. *Calc. Var. Partial Differential Equations.*

[8] 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.

[9] 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.

[10] A. Figalli, L. Rifford and C. Villani. Tangent cut loci on surfaces. *Differential Geom. Appl.*, 29(2):154–159, 2011.

[11] A. Figalli, L. Rifford and C. Villani. Necessary and sufficient conditions for continuity of optimal transport maps on Riemannian manifolds. *Tohoku Math. J.*, 63(4):855–876, 2011.

[12] A. Figalli, L. Rifford and C. Villani. Nearly round spheres look convex. *Amer. J. Math.*, 134(1):109-139, 2012.

[13] 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.

[14] S. Gallot, D. Hulin and J. Lafontaine. *Riemannian geometry*, second ed. Universitext. Springer-Verlag, Berlin, 1990.

[15] T. Gallouët. Transport Optimal : Régularité et applications. PhD Thesis, ENS Lyon, 2012.

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

[17] K. Young-Heon and R.J. McCann. Continuity, curvature, and the general covariance of optimal transportation *J. Eur. Math. Soc. (JEMS)*,12(4) ,1009–1040, 2010.

[18] K. Young-Heon and R.J. McCann. Towards the smoothness of optimal maps on Riemannian submersions and Riemannian products (of round spheres in particular) *J. Reine Angew. Math.*,664,1–27,2012

[19] 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.

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

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

[22] X. N. Ma, N. S. Trudinger and X. J. Wang. Regularity of potential functions of the optimal transportation problem. *Arch. Ration. Mech. Anal.*, 177 (2005), no. 2, 151–183.

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

[24] T. Sakai. Riemannian geometry. *Translations of Mathematical Monographs*, Vol. 149. American Mathematical Society, 1996.

[25] C. Villani. Optimal transport, old and new. *Grundlehren des mathematischen Wissenschaften*, Vol. 338. Springer-Verlag, Berlin, 2009.

[26] C. Villani. Regularity of optimal transport and cut locus: From nonsmooth analysis to geometry to smooth analysis. *Discrete Contin. Dyn. Sys.*, 30(2):559–571, 2011.