Approximations of Lipschitz maps via Ehresmann fibrations and Reeb's sphere theorem for Lipschitz functions[*][†]

To the memory of Kazumasa Kuwada

Kei KONDO[‡]

June 29, 2021

Abstract

We show, as our main theorem, that if a Lipschitz map from a compact Riemannian manifold M to a connected compact Riemannian manifold N, where $\dim M \geq \dim N$, has no singular points on M in the sense of F.H. Clarke, then the map admits a smooth approximation via Ehresmann fibrations. We also show the Reeb sphere theorem for Lipschitz functions, i.e., if a closed Riemannian manifold admits a Lipschitz function with exactly two singular points in the sense of Clarke, then the manifold is homeomorphic to the sphere.

1 Introduction

1.1 Background: Grove–Shiohama theory for distance functions

Armed with the Toponogov Comparison Theorem [45] (see also [6], [39]), K. Grove and K. Shiohama [19] developed a theory for critical points of distance functions on complete Riemannian manifolds that has played a fundamental role in the study of relationships between curvature and topology. Denote by X a complete Riemannian manifold, d its distance function, and T_xX the tangent space at each $x \in X$. Fix $p \in X$, and set $d_p(x) := d(p, x)$ for all $x \in X$. Note that d_p is a 1-Lipschitz function and is smooth on $X \setminus (\{p\} \cup \mathrm{Cut}(p))$ where $\mathrm{Cut}(p)$ indicates the cut locus of p. (For basic definitions in Riemannian geometry see, for example, [6], [9], [39].) Grove and Shiohama gave the following meaningful definition in order to do research into how d_p behaves.

Definition 1.1 ([19]) A point $q \in X \setminus \{p\}$ is said to be *critical for* d_p (or a *critical point of* d_p) *in the sense of Grove–Shiohama* if for each $v \in T_qX \setminus \{o_q\}$ there is a unit speed minimal geodesic segment $\gamma : [0, d_p(q)] \to X$ emanating from $p = \gamma(0)$ to $q = \gamma(d_p(q))$ such that $\angle(-(d\gamma/dt)(d_p(q)), v) \leq \pi/2$ where $\angle(-(d\gamma/dt)(d_p(q)), v)$ denotes the angle between two vectors $-(d\gamma/dt)(d_p(q))$ and v in T_qX. For convenience we also call p a critical point of d_p.

---

[*]2010 Mathematics Subject Classification: Primary 49J52, 53C20; Secondary 57R12, 57R55.

[†]Key words and phrases: convex analysis, Ehresmann fibration, Lipschitz map, nonsmooth analysis, Reeb's sphere theorem, smooth approximation.

[‡]Supported by the Grant-in-Aid for Science Reserch (C), JSPS KAKENHI Grant Numbers 16K05133, 17K05220, 18K03280.

The origins of this definition can be found in the work of M. Berger [2]: the point of maximal distance from a given point $x \in X$ is a critical point of d_x. See the survey articles by J. Cheeger [5] and by K. Grove [17] on critical points of distance functions. Note that any critical point of d_x is also a cut point of x.

Another major development, due to M. Gromov [15], was in the topology of regions free of critical points.

Lemma 1.2 (Gromov's isotopy lemma) *If $0 < R_1 < R_2 \leq \infty$, and if d_p has no critical points on $\overline{B_{R_2}(p)} \setminus B_{R_1}(p)$, then $\overline{B_{R_2}(p)} \setminus B_{R_1}(p)$ is homeomorphic to $\partial B_{R_1}(p) \times [R_1, R_2]$ where each $B_{R_i}(p)$ denotes the metric open ball with center p and radius R_i, and $\overline{B_{R_i}(p)}$ indicates the closure of $B_{R_i}(p)$ $(i = 1, 2)$.*

The Toponogov comparison theorem [45] (see also [6], [39]) together with the isotopy lemma yields the *diameter sphere theorem*:

Theorem 1.3 ([19]) *If the sectional curvature of X is bounded from below by 1, and if the diameter of X is greater than $\pi/2$, then X is homeomorphic to the sphere.*

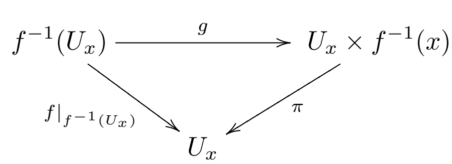
1.2 Critical points of Lipschitz functions

The method of Grove and Shiohama has many applications (see [1], [15], [18], [24], [25], and the survey articles [5], [17]). A natural question to ask is whether it can be extended to general Lipshitz functions.

The purpose of this article is to tackle this question by employing Clarke's nonsmooth analysis. That is, we will extend the notion of critical points of distance functions on Riemannian manifolds to locally Lipschitz maps. In the absence of singular points we will show the existence of a family of Ehresmann fibrations which approximate an arbitrary Lipschitz map between compact manifolds without curvature assumption (Theorem 1.4). Moreover we will show the Reeb sphere theorem for Lipschitz functions on closed Riemannian manifolds (Theorem 1.7) which corresponds to that for smooth ones [36], [31].

1.3 Main theorem

Let M and N be smooth manifolds. A smooth map $f : M \to N$ is called *an Ehresmann fibration* (or a *locally trivial fibration*) if for each $x \in N$ there are an open neighborhood U_x of x and a diffeomorphism $g : f^{-1}(U_x) \to U_x \times f^{-1}(x)$ such that the diagram



commutes where $\pi : U_x \times f^{-1}(x) \to U_x$, $\pi(p, q) := p$, denotes the projection to the first factor. Note that π is a smooth map. Our main theorem is stated as follows:

Theorem 1.4 (Main Theorem) *Let $F : M \to N$ be a Lipschitz map from a compact Riemannian manifold M to a connected compact Riemannian manifold N where $\dim M \geq \dim N$. If F has no singular points on M in the sense of Clarke, then for any $\eta > 0$ there is a constant $\kappa(\eta) > 0$ such that for each $\varepsilon \in (0, \kappa(\eta))$ there is an Ehresmann fibration f_ε from M onto N satisfying* $\max_{x \in M} d_N(f_\varepsilon(x), F(x)) < \eta$.

Remark 1.5 Let us mention remarks on Theorem 1.4:

(i) The definition of a singular point of Lipschitz maps in the sense of Clarke will be given in Section 2.2.

(ii) The author and Tanaka showed the existence of a family of immersions which approximate an arbitrary Lipschitz map between compact manifolds, see [26, Theorem 1.3]. Li [29] announced an another proof of [26, Theorem 1.3]. [26, Corollary 1.15] guarantees that an assumption on [13, Proposition 22] is natural.

(iii) The related result is the Yamaguchi fibration theorem [50]: Let X and Y be complete Riemannian manifolds of $\dim X = n$ and $\dim Y = k$, respectively, where $n \geq k$. Assume that both sectional curvatures are bounded from below by -1, and that the injectivity radius of Y has a lower bound $\delta > 0$. He then showed that there is a constant $\varepsilon(n, \delta) > 0$ such that if $d_{\rm GH}(X, Y) < \varepsilon(n, \delta)$, then there is a fibration $f : X \to Y$ which is an almost Riemannian submersion where $d_{\rm GH}$ indicates the Gromov–Hausdorff distance. He also gave this type of fibration theorem for Alexandrov spaces [51]. Moreover Fujioka [12] showed a locally trivial fibration theorem for Alexandrov spaces assuming a lower positive bound for the volume of the space of directions.

1.4 Reeb's sphere theorem for Lipschitz functions

In the process of proving Theorem 1.4 we obtain the following corollary of a proposition for Lipschitz maps between Riemannian manifolds:

Corollary 1.6 (Corollary 5.5 in Section 5) *Let F be a Lipschitz function on a compact Riemannian manifold M, and $\widetilde{F}_\varepsilon : M \to \mathbb{R}$ the global smooth approximation of F (see Definition 4.9 for $\ell = 1$). If $p \in M$ is nonsingular for F in the sense of Clarke, then there are two constants $\lambda(p) > 0$ and $\varepsilon_0(p) > 0$ such that if $\varepsilon \in (0, \varepsilon_0(p))$, then $\operatorname{grad} \widetilde{F}_\varepsilon \neq 0$ on the metric open ball $B_{\lambda(p)}(p)$ with center p and radius $\lambda(p)$ where $\operatorname{grad} \widetilde{F}_\varepsilon$ denotes the gradient vector field of $\widetilde{F}_\varepsilon$. In particular $\widetilde{F}_\varepsilon$ has no critical points on $B_{\lambda(p)}(p)$ for an $\varepsilon > 0$ sufficiently small.*

Applying this corollary we show Reeb's sphere theorem for Lipschitz functions:

Theorem 1.7 *If a closed Riemannian manifold admits a Lipschitz function with exactly two singular points in the sense of Clarke, then the manifold is homeomorphic to the sphere.*

Remark 1.8 We give two remarks on Theorem 1.7:

(i) Let X be a closed Riemannian manifold, $p \in X$, and d_p the distance function of X given by $d_p(x) := d(p, x)$ for all $x \in X$. We then see that a point $x \in X$ is critical for d_p in the sense of Grove–Shiohama if and only if $x \in X$ is singular for it in that of Clarke (Proposition 2.19 and Lemma 2.20). Theorem 1.7 thus contains Reeb's sphere theorem for distance functions [39, Proposition 2.10], and hence Theorem 1.7 yields Theorem 1.3.

(ii) Since the manifold in Theorem 1.7, denoted by M below, is a twisted sphere, we see, by [4], [23], [32], [34], and [42], that M is diffeomorphic to the standard sphere when $\dim M \leq 6$. Moreover the Weinstein deformation technique for metrics ([47]) shows that M admits a metric such that there is a point whose cut locus consists of a single point. It is the worthy of noting that every exotic sphere of dimension greater than 4 admits such metrics by the Smale h-cobrdism theorem [43], [44] together with the deformation technique.

The article is organized as follows. In Section 2 we define a generalized differential for Lipschitz maps between Riemannian manifolds (Definitions 2.4 and 2.6) and singular points of them in the sense of Clarke (Definition 2.8). Giving intrinsic definitions to them is another aim of this article. That is, although we had given the definitions in [26], the identification, which has often been done in [26], of the set of all linear mappings of tangent spaces and the vector space of matrices seems to have given not only an impression that it is hard to read it, but also a misunderstanding that the definitions depend on the choice of charts. To prevent them we thus employ parallel transports along minimal geodesics in our definitions. This is the big difference between our definitons and those in [26]. Moreover we also define the generalized gradient for Lipschitz functions on Riemannian manifolds (Definition 2.13) and study the relationship between the gradient and the generalized differential of them. As an example of singular points of Lipschitz functions we show finally that critical points of distance functions in the sense of Grove–Shiohama are singular ones of them in that of Clarke (Proposition 2.19 and Lemma 2.20).

In Section 3 we define the adjoint of the generalized differential of Lipschitz maps between Riemannian manifolds (Definition 3.3), and discuss surjectivity and injectivity of the generalized differential near a nonsingular point of a given Lipschitz map (Propositions 3.6 and 3.7, respectively). These propositions show that the set of all singular points of the map is a closed set in its source space (Corollary 3.9).

In Section 4 we first define a local smooth approximation of an arbitrary Lipschitz map between Riemannian manifolds on a strongly convex ball as the Riemannian convolution smoothing (Definiton 4.2), and next define the global smooth approximation of the map via a smooth partition of unity (Definiton 4.9).

In Section 5 we give the proof of Theorem 1.4. For this we first show, broadly speaking, that a global smooth approximation of a Lipschitz map F on a compact manifold carries on surjectivity of the generalized differential of F (Proposition 5.4). As a corollary of Proposition 5.4 we get Corollary 1.6. Using the proposition and the tubular neighborhood theorem, we finally show the main theorem.

In Section 6, making use of Corollary 1.6 with Morse theory, we show Theorem 1.7.

Acknowledgements. The author expresses his sincere thanks to Professors M. Koiso, A. Katsuda, A. Mitsuishi, and Y. Otsu who attended his intensive course on [26], organized by Koiso, at Kyushu university in the summer of 2017 and their many critical remarks, which started him considering intrinsic definitions of the generalized differential and singular points of Lipschitz maps. He wishes to thank Professor N. Nakauchi for helpful discussions about the intrinsic definitions in Section 2. He wishes to express his appreciation to the referee for his cooperation and helpful comments. In particular he has been able to improve Theorem 1.7 due to the comments. And last but not least he is grateful to Professor K. Shiohama for his strong encouragement during the preparation of this work.

2 Nonsmooth analysis in Riemannian geometry

2.1 From Rockafellar to Clarke

Rockafellar [37] was the first to introduce the notion of the subdifferential of a convex function. This was done on of Euclidean space in order to replace assumptions of smoothness with convexity and led to many results. Clarke [7], [8] generalized Rockafellar's work to Lipschitz maps between Euclidean spaces and the subdifferential to the generalized differential (see Definition 2.6).

The two examples below show how Clarke's generalized differential of Lipschitz maps emerges from Rockafellar's subdifferential of convex functions.

(i) We here recall Rockafellar's ideas, that is the subdifferential of convex functions. Let $f_1(x) := |x| - 1$ and $f_2(x) := (x-2)^2 - 1$ for all $x \in \mathbb{R}$. Define the convex function $f : \mathbb{R} \to \mathbb{R}$ by $f(x) := \max\{f_1(x), f_2(x)\}$ $(x \in \mathbb{R})$. Note that f is not differentiable on $\{1, 4\}$. However one-sided limits of f' do exist, i.e., $\lim_{x\uparrow 1} f'(x) = -2$, $\lim_{x\downarrow 1} f'(x) = 1$, $\lim_{x\uparrow 4} f'(x) = 1$, and $\lim_{x\downarrow 4} f'(x) = 4$. Rockafellar's idea is to draw vertical segments between disconnected points of the graph of f' using convex combinations: between $(1, -2)$ and $(1, 1)$ and between $(4, 1)$ and $(4, 4)$. Put differently, for each $\lambda \in [0, 1]$ we have $(1-\lambda)\lim_{x\uparrow 1} f'(x) + \lambda \lim_{x\downarrow 1} f'(x) = 3\lambda - 2$ and $(1-\lambda)\lim_{x\uparrow 4} f'(x) + \lambda \lim_{x\downarrow 4} f'(x) = 3\lambda + 1$, and hence $\partial f(1) := \{3\lambda - 2 \mid \lambda \in [0,1]\} = [-2, 1]$ and $\partial f(4) := \{3\lambda + 1 \mid \lambda \in [0,1]\} = [1, 4]$. Since $0 \in \partial f(1)$, and since f is not monotone near $x = 1$, we can regard $x = 1$ as a critical point of f. In particular f has the minimum value 0 at $x = 1$. On the other hand we can regard $x = 4$ as a noncritical point of f, as $0 \notin \partial f(4)$, and f is increasing near $x = 4$. He called $\partial f(1)$ and $\partial f(4)$ the *subdifferentials* of f at $x = 1, 4$, respectively.

(ii) Let $g : \mathbb{R} \to \mathbb{R}$ be the locally Lipschitz function defined by

$$g(x) = \begin{cases} x^2 \sin \dfrac{1}{x} & (x \neq 0), \\ 0 & (x = 0). \end{cases}$$

Note that g is differentiable on $\mathbb{R}$, but is not C^1 at $x = 0$. Moreover we can not directly apply Rockafellar's idea as in example (i) to one-sided limits of g' at $x = 0$ due to the term $\cos(1/x)$ in $g'(x)$. Clarke's idea is to choose a sequence of lines with the same slope tangent to the graph of g, or, more precisely, for each $\alpha \in [-1, 1]$ we choose a sequence $\{x_i^{(\alpha)}\}_{i\in\mathbb{N}} \subset \mathbb{R}$ which converges to 0 as $i \to \infty$ such that $\lim_{i\to\infty} g'(x_i^{(\alpha)}) = \alpha$, and take the convex hull, denoted by $\mathrm{Conv}(A)$, of the (nonempty) set

$$A := \left\{\alpha \,\middle|\, \exists \{x_i^{(\alpha)}\}_{i\in\mathbb{N}} \subset \mathbb{R} \setminus \{0\} \text{ such that } \lim_{i\to\infty} x_i^{(\alpha)} = 0,\ \lim_{i\to\infty} g'(x_i^{(\alpha)}) = \alpha\right\}$$

For instance, in the case of $\alpha = -1/2$, set $1/x_i^{(-1/2)} := \pi/3 + 2(i-1)\pi$ $(i \in \mathbb{N})$. We then have $\lim_{i\to\infty} x_i^{(-1/2)} = 0$ and $\lim_{i\to\infty} g'(x_i) = -1/2$. So $A = \mathrm{Conv}(A) = [-1, 1]$ (cf. [37, Theorem 2.1]). Since $0 \in \mathrm{Conv}(A)$, we can regard $x = 0$ as a critical point of g. Clarke called $\mathrm{Conv}(A)$ the *generalized differential* of g.

2.2 An intrinsic definition of the generalized differential of Lipschitz maps

The aim of this subsection is to intrinsically define the generalized differential for Lipschitz maps between Riemannian manifolds and their singular points in the sense of Clarke.

We first recall Whitehead's convexity theorem. The theorem not only allows us to intrinsically define the generalized differential for Lipschitz maps between Riemannian manifolds, but also plays an important role in our smooth approximation method for such maps. A proof of the theorem can be found in [49] or [9, Proposition 4.2, pp. 76–77].

Theorem 2.1 (Whitehead's convexity theorem) *Let X be a Riemannian manifold and d_X the distance function on X. Then for each $x \in X$ there is a constant $\alpha(x) > 0$ such that*

(a) *the open ball $B_{\alpha(x)}(x) := \{y \in X \,|\, d_X(x,y) < \alpha(x)\}$ is strongly convex, i.e., for any two points $p, q \in X$ in the closure $\overline{B_{\alpha(x)}(x)}$ there is a unique geodesic segment $\gamma : [0,1] \to X$ emanating from $p = \gamma(0)$ to $q = \gamma(1)$ such that $\gamma(0,1) \subset B_{\alpha(x)}(x)$;*

(b) *the exponential map $\exp_x|_{\mathbb{B}_{\alpha(x)}(o_x)} : \mathbb{B}_{\alpha(x)}(o_x) \to B_{\alpha(x)}(x)$ at x is a diffeomorphism where $\mathbb{B}_{\alpha(x)}(o_x) := \{v \in T_xX \,|\, \|v\| < \alpha(x)\}$ and o_x indicates the origin of the tangent space T_xX at x.*

From now on let M and N be Riemannian manifolds of dimension m and n, respectively, and $F : M \to N$ a locally Lipschitz map. The following lemma is a direct consequence of Theorem 2.1.

Lemma 2.2 *For each $p \in M$ there are two open balls $B_{r(p)}(p) \subset M$ and $B_{t(p)}(F(p)) \subset N$ such that*

(i) *both $B_{r(p)}(p)$ and $B_{t(p)}(F(p))$ satisfy* (a) *and* (b) *of Theorem 2.1;*

(ii) *$F(B_{r(p)}(p)) \subset B_{t(p)}(F(p))$;*

(iii) *$F|_{B_{r(p)}(p)} : B_{r(p)}(p) \to B_{t(p)}(F(p))$ is Lipschitz continuous.*

Using parallel transport we intrinsically define the generalized differential for F:

Definition 2.3 We will use the following notation. If there exists a unique geodesic between $x, y \in M$, then denote parallel transport along that geodesic by $\tau_y^x : T_xM \to T_yM$.

For each $x \in M$ let $\mathcal{L}(T_xM, T_{F(x)}N)$ be the set of all linear mappings of T_xM to $T_{F(x)}N$. Since $\mathcal{L}(T_xM, T_{F(x)}N)$ is isomorphic to the vector space $\mathcal{M}(n, m; \mathbb{R})$ of $n \times m$-matrices with real entries, $\mathcal{L}(T_xM, T_{F(x)}N)$ is an nm-dimensional vector space. We topologize $\mathcal{L}(T_xM, T_{F(x)}N)$ with the operator norm $\|\cdot\|$, so that, throughout this article, we regard $\mathcal{L}(T_xM, T_{F(x)}N)$ as a finite dimensional normed vector space.

Fix $p \in M$. Choose the two balls $B_{r(p)}(p) \subset M$ and $B_{t(p)}(F(p)) \subset N$ satisfying all three properties (i)–(iii) of Lemma 2.2. From Rademacher's theorem [35] there is a set $E_F \subset M$ of Lebesgue measure zero such that the differential dF of F exists on $B_{r(p)}(p) \setminus E_F$. Since F is Lipschitz on $B_{r(p)}(p)$, and since for $q \in B_{r(p)}(p)$ parallel transports τ_q^p and $\tau_{F(p)}^{F(q)}$ are linear isometries on $B_{r(p)}(p)$ and $B_{t(p)}(F(p))$, respectively,

$\{\tau_{F(p)}^{F(q)} \circ dF_q \circ \tau_q^p\}_{q \in B_{r(p)}(p) \setminus E_F}$ is bounded in $\mathcal{L}(T_pM, T_{F(p)}N)$. Since $B_{r(p)}(p) \setminus E_F$ is dense in $B_{r(p)}(p)$, there is a sequence $\{x_i\}_{i\in\mathbb{N}} \subset B_{r(p)}(p) \setminus E_F$ such that $\lim_{i\to\infty} x_i = p$ and $\{\tau_{F(p)}^{F(x_i)} \circ dF_{x_i} \circ \tau_{x_i}^p\}_{i\in\mathbb{N}}$ converges in $\mathcal{L}(T_pM, T_{F(p)}N)$. Hence we can introduce the notion of the "mixture" of the differential of F as follows:

Definition 2.4 (compare [26]) For each $p \in M$ we call the set

$$K_F(p) := \left\{ G \in \mathcal{L}(T_pM, T_{F(p)}N) \;\middle|\; \begin{array}{l} \exists \{x_i\}_{i\in\mathbb{N}} \subset B_{r(p)}(p) \setminus E_F \text{ such that} \\ \lim_{i\to\infty} x_i = p,\ \lim_{i\to\infty} \tau_{F(p)}^{F(x_i)} \circ dF_{x_i} \circ \tau_{x_i}^p = G \end{array} \right\} \tag{2.1}$$

the *mixture of the differential of F at p*. Note here that, from (ii) of Lemma 2.2, parallel transport $\tau_{F(p)}^{F(x_i)} : T_{F(x_i)}N \to T_{F(p)}N$ can be defined.

Remark 2.5 By Definition 2.4, for any $p \in M$, $K_F(p)$ is a nonempty bounded set in $\mathcal{L}(T_pM, T_{F(p)}N)$.

The generalized differential for $F : M \to N$ is now intrinsically defined as follows:

Definition 2.6 For each $p \in M$ we call the set $\partial F(p) := \mathrm{Conv}(K_F(p))$ the *generalized differential of F at p* where again $\mathrm{Conv}(K_F(p))$ denotes the convex hull of the mixture $K_F(p)$ of the differential of F at p.

Remark 2.7 We give remarks on Definition 2.6: Fix $p \in M$.

(i) Clarke [8] originally called $\partial F(p)$ the *generalized Jacobian of F at p* where M and N are Euclidean spaces of the same dimension m. This is because we can use the atlas $\{(\mathbb{R}^m, \mathrm{id}_{\mathbb{R}^m})\}$ with a single chart on $\mathbb{R}^m$ where $\mathrm{id}_{\mathbb{R}^m} : \mathbb{R}^m \to \mathbb{R}^m$ is the identity map, so without referring to independence of the choice of charts we can define $\partial F(p)$ as follows:

$$\partial F(p) := \mathrm{Conv}\left(\left\{ A \in \mathcal{M}(n, m; \mathbb{R}) \;\middle|\; \begin{array}{l} \exists \{x_i\}_{i\in\mathbb{N}} \subset \mathbb{R}^m \setminus E_F \text{ such that} \\ \lim_{i\to\infty} x_i = p,\ \lim_{i\to\infty} (JF)_{x_i} = A \end{array} \right\} \right) \tag{2.2}$$

where $(JF)_{x_i}$ indicates the Jacobian matrix of F at x_i.

(ii) From [37, Theorem 17.2], $\partial F(p)$ is a compact convex subset of $\mathcal{L}(T_pM, T_{F(p)}N)$.

(iii) Although the following fact was mentioned in Section 2.1 (ii), we will mention it again: [37, Theorem 2.1] shows that $\partial F(p)$ is the smallest convex set containing $K_F(p)$, i.e.,

$$\partial F(p) = \mathrm{Conv}(K_F(p)) = \bigcap \{Z \mid Z \text{ is convex in } \mathcal{L}(T_pM, T_{F(p)}N) \text{ with } K_F(p) \subset Z\}.$$

(iv) Since $\dim \mathcal{L}(T_pM, T_{F(p)}N) = nm$, it follows from Carathéodory's theorem (cf. [40, Theorem 1.1.4]) that for any $g \in \partial F(p)$ there are $G_1, G_2, \ldots, G_{nm+1} \in K_F(p)$ such that $g = \sum_{i=1}^{nm+1} a_i G_i$ where $\sum_{i=1}^{nm+1} a_i = 1$ and $a_i \geq 0$ $(i = 1, 2, \ldots, nm+1)$.

(v) If dF_p exists, then $dF_p \in \partial F(p)$. Moreover, if F is of class C^1 on $B_{r(p)}(p)$, then $\partial F(p)$ is a singleton, which means $\partial F(p) = \{dF_p\}$.

(vi) As a direct consequence of Definition 2.6 we observe that for any $\varepsilon > 0$ there is a constant $\mu(p, \varepsilon) \in (0, r(p))$ such that $\tau_{F(p)}^{F(x)} \circ \partial F(x) \circ \tau_x^p \subset \mathcal{U}_\varepsilon(\partial F(p))$ for all $x \in B_{\mu(p,\,\varepsilon)}(p)$ where $\tau_{F(p)}^{F(x)} \circ \partial F(x) \circ \tau_x^p := \{\tau_{F(p)}^{F(x)} \circ g \circ \tau_x^p \mid g \in \partial F(x)\}$ and $\mathcal{U}_\varepsilon(\partial F(p))$ denotes the ε-open neighborhood of $\partial F(p)$ in $\mathcal{L}(T_pM, T_{F(p)}N)$.

Now that we have defined the generalized differential for Lipschitz maps between Riemannian manifolds, it is time to intrinsically define their singular points:

Definition 2.8 A point $p \in M$ is said to be *nonsingular for* F (or a *nonsingular point of* F) *in the sense of Clarke* if every element in $\partial F(p)$ is of maximal rank, i.e., for any $g \in \partial F(p)$, $\mathrm{rank}(g) = \min\{m, n\}$.

Remark 2.9 Clarke [8] first introduced the notion of singular points of Lipschitz maps between Euclidean spaces of the same dimension. Using this notion he extended the inverse function theorem for smooth maps between Euclidean spaces to Lipschitz ones, see [8, Theorem 1].

2.3 The relationship between the generalized differential and the generalized gradient of Lipschitz functions

In this subsection we define the generalized gradient of Lipschitz functions on Riemannian manifolds and study the relationship between their generalized gradient and their generalized differential. Throughout this subsection let M be a Riemannian manifold of dimension m with Riemannian metric $\langle \cdot\,, \cdot \rangle$, $F : M \to \mathbb{R}$ a locally Lipschitz function, and t the standard coordinate on $\mathbb{R}$.

Fix $p \in M$. By Theorem 2.1 there is a strongly convex open ball $B_{r(p)}(p)$ such that $\exp_p|_{\mathbb{B}_{r(p)}(o_p)} : \mathbb{B}_{r(p)}(o_p) \to B_{r(p)}(p)$ is a diffeomorphism. For each $x \in B_{r(p)}(p)$ parallel transport $\tau_x^p : T_pM \to T_xM$ is defined as in Definition 2.3. Note that parallel transport $\tau_y^x : T_x\mathbb{R} \to T_y\mathbb{R}$ is defined for all $x, y \in \mathbb{R}$. Let $K_F(p)$ be the mixture of the differential of F at p defined by Eq. (2.1) for $T_{F(p)}N = T_{F(p)}\mathbb{R}$, and E_F a set of Lebesgue measure zero such that dF exists on $B_{r(p)}(p) \setminus E_F$. The gradient vector field of F denoted by $\mathrm{grad}\, F$ can be defined on $B_{r(p)}(p) \setminus E_F$ because F is differentiable there. Note that

$$\langle (\mathrm{grad}\, F)_x, u \rangle \cdot \frac{d}{dt}\bigg|_{F(x)} = u(F) \cdot \frac{d}{dt}\bigg|_{F(x)} = dF_x(u) \tag{2.3}$$

for all $x \in B_{r(p)}(p) \setminus E_F$ and $u \in T_xM$.

Lemma 2.10 *For any $G \in K_F(p)$ there is a sequence $\{x_i\}_{i\in\mathbb{N}} \subset B_{r(p)}(p) \setminus E_F$ such that $\lim_{i\to\infty} x_i = p$, $\lim_{i\to\infty} \tau_{F(p)}^{F(x_i)} \circ dF_{x_i} \circ \tau_{x_i}^p = G$, and*

$$G(v) = \lim_{i\to\infty} \langle (\mathrm{grad}\, F)_{x_i}, \tau_{x_i}^p(v) \rangle \frac{d}{dt}\bigg|_{F(p)} \qquad (v \in T_pM).$$

Proof. Fix $G \in K_F(p)$. The definition of $K_F(p)$ gives a sequence $\{x_i\}_{i\in\mathbb{N}} \subset B_{r(p)}(p) \setminus E_F$ satisfying $\lim_{i\to\infty} x_i = p$ and $\lim_{i\to\infty} \tau_{F(p)}^{F(x_i)} \circ dF_{x_i} \circ \tau_{x_i}^p = G$. For any $v \in T_pM$, Eq. (2.3) gives

$$G(v) = \Big(\lim_{i\to\infty} \tau_{F(p)}^{F(x_i)} \circ dF_{x_i} \circ \tau_{x_i}^{p}\Big)(v) = \lim_{i\to\infty} \tau_{F(p)}^{F(x_i)}\Big(\tau_{x_i}^{p}(v)(F)\frac{d}{dt}\Big|_{F(x_i)}\Big)$$
$$= \lim_{i\to\infty} \tau_{x_i}^{p}(v)(F)\frac{d}{dt}\Big|_{F(p)} = \lim_{i\to\infty} \langle (\operatorname{grad} F)_{x_i}, \tau_{x_i}^{p}(v)\rangle \frac{d}{dt}\Big|_{F(p)}. \tag{2.4}$$

□

Definition 2.11 Eq. (2.4) defines $\lim_{i\to\infty}(\operatorname{grad} F)_{x_i}$, that is, for each $G \in K_F(p)$ let

$$\langle \lim_{i\to\infty} (\operatorname{grad} F)_{x_i}, v\rangle \frac{d}{dt}\Big|_{F(p)} := G(v) \qquad (v \in T_pM).$$

The following lemma follows directly from Lemma 2.10 together with Definition 2.11.

Lemma 2.12 *The set*

$$\divideontimes_F(p) := \left\{ w \in T_pM \;\middle|\; \begin{array}{l} \exists \{x_i\}_{i\in\mathbb{N}} \subset B_{r(p)}(p) \setminus E_F \quad \text{such that} \\ \lim_{i\to\infty} x_i = p,\ \lim_{i\to\infty} (\operatorname{grad} F)_{x_i} = w \end{array} \right\}$$

is nonempty.

Definition 2.13 We call $\divideontimes_F(p)$ the *mixture of the gradient* of F at p, and the convex set $\circledast_F(p) := \operatorname{Conv}(\divideontimes_F(p))$ the *generalized gradient* of F at p.

Remark 2.14 $\circledast_F(p)$ is compact in T_pM by [37, Theorem 17.2]. Note that Clarke [7] first defined the generalized gradient of Lipschitz functions on $\mathbb{R}^m$.

Lemma 2.15 *Let $\partial F(p)$ be the generalized differential of F at p. Then for any $g \in \partial F(p)$ there is a vector $X^{(g)} \in \circledast_F(p)$ such that $g(v) = \langle X^{(g)}, v\rangle \cdot d/dt|_{F(p)}$ for all $v \in T_pM$.*

Proof. Fix $g \in \partial F(p)$. By Carathéodory's theorem (cf. [40, Theorem 1.1.4]) there are vectors $G_1, G_2, \ldots, G_{m+1} \in K_F(p)$ such that $g = \sum_{k=1}^{m+1} a_k G_k$ where $\sum_{k=1}^{m+1} a_k = 1$ and $a_k \geq 0$ $(k = 1, 2, \ldots, m+1)$. By Lemmas 2.10 and 2.12 for each $k = 1, 2, \ldots, m+1$ there is a vector $w^{(k)} \in \divideontimes_F(p)$ such that $G_k(v) = \langle w^{(k)}, v\rangle (d/dt)|_{F(p)}$ for all $v \in T_pM$, and hence

$$g(v) = \sum_{k=1}^{m+1} a_k G_k(v) = \sum_{k=1}^{m+1} a_k \langle w^{(k)}, v\rangle \frac{d}{dt}\Big|_{F(p)} = \Big\langle \sum_{k=1}^{m+1} a_k w^{(k)}, v \Big\rangle \frac{d}{dt}\Big|_{F(p)}.$$

Since $\sum_{k=1}^{m+1} a_k w^{(k)} \in \circledast_F(p)$, $X^{(g)} := \sum_{k=1}^{m+1} a_k w^{(k)}$ is the desired vector. □

Lemma 2.16 *p is singular for F if and only if $o_p \in \circledast_F(p)$.*

Proof. We first assume that p is singular for F. There is then a point $g_0 \in \partial F(p)$ such that $\operatorname{rank}(g_0) = 0$. By Lemma 2.15 there is a vector $X^{(0)} \in \circledast_F(p)$ such that $g_0(v) = \langle X^{(0)}, v\rangle (d/dt)|_{F(p)}$ holds for all $v \in T_pM$. Since $\operatorname{rank}(g_0) = 0$, $o_{F(p)} = \langle X^{(0)}, v\rangle \cdot d/dt|_{F(p)}$ for all $v \in T_pM$ where $o_{F(p)}$ indicates the origin of $T_{F(p)}\mathbb{R}$, hence $X^{(0)} = o_p$, and finally $o_p = X^{(0)} \in \circledast_F(p)$.

We next assume $o_p \in \circledast_F(p)$. Carathéodory's theorem (cf. [40, Theorem 1.1.4]) shows that there are vectors $w_1, w_2, \ldots, w_{m+1} \in \divideontimes_F(p)$ such that $o_p = \sum_{k=1}^{m+1} \alpha_k w_k$ where $\sum_{k=1}^{m+1} \alpha_k = 1$ and $\alpha_k \geq 0$ $(k = 1, 2, \ldots, m+1)$. Define the linear map $f_0 : T_pM \to T_{F(p)}\mathbb{R}$ by $f_0(v) := \langle \sum_{k=1}^{m+1} \alpha_k w_k, v\rangle \cdot d/dt|_{F(p)}$ for all $v \in T_pM$. We observe $f_0 \in \partial F(p)$. Since $f_0(v) = o_{F(p)}$ for all $v \in T_pM$, $\operatorname{rank}(f_0) = 0$, and hence p is singular for F. $\square$

Remark 2.17 In [7] and [26] a point $p \in M$ is called *noncritical for F* if $o_p \notin \circledast_F(p)$.

2.4 Critical points of distance functions in the sense of Grove–Shiohama are singular points of Clarke.

Throughout this subsection let M be a complete Riemannian manifold of dimension m with Riemannian metric $\langle\,\cdot\,,\,\cdot\,\rangle$ and the distance function d. All geodesics will be normal.

Fix $p \in M$. Define the map $d_p : M \to \mathbb{R}$ by $d_p(x) := d(p, x)$ for all $x \in M$. We then have the following proposition. Note that the proposition appeared as [26, Example 1.9]; however we are sometimes asked the proof, so that we give the details here.

Proposition 2.18 *$q \in M$ is singular for d_p in the sense of Clarke if and only if q is critical for d_p in that of Grove–Shiohama.*

Proof. For each $x \in M$ let $B_{r(x)}(x)$ be a strongly convex open ball, guaranteed by Theorem 2.1, such that $\exp_x|_{\mathbb{B}_{r(x)}(o_x)}$ is a diffeomorphism, and E_{d_p} a set of Lebesgue measure zero such that the differential of d_p exists on $M \setminus E_{d_p}$. This proposition follows from Lemmas 2.19 and 2.20 below. $\square$

Lemma 2.19 *$q \in M \setminus \{p\}$ is singular for d_p in the sense of Clarke if and only if q is critical for d_p in that of Grove–Shiohama.*

Proof. Assume that $q \in M \setminus \{p\}$ is singular for d_p in the sense of Clarke. By Lemma 2.16, $o_q \in \circledast_{d_p}(q)$ holds where $\circledast_{d_p}(q)$ indicates the generalized gradient of d_p at q. From Carathéodory's theorem (cf. [40, Theorem 1.1.4]) there are $w_1, w_2, \ldots, w_{m+1} \in \divideontimes_{d_p}(q)$ such that $o_q = \sum_{k=1}^{m+1} \alpha_k w_k$ where $\sum_{k=1}^{m+1} \alpha_k = 1$ and $\alpha_k \geq 0$ $(k = 1, 2, \ldots, m+1)$, and $\divideontimes_{d_p}(q)$ denotes the mixture of the gradient of d_p at q. Fix $v \in T_qM \setminus \{o_q\}$. We then have $0 = \langle \sum_{k=1}^{m+1} \alpha_k w_k, v\rangle = \sum_{k=1}^{m+1} \alpha_k \langle w_k, v\rangle$. Since $\alpha_k \geq 0$ $(k = 1, 2, \ldots, m+1)$, there is a number $k_0 \in \{1, 2, \ldots, m+1\}$ such that $\langle w_{k_0}, v\rangle \leq 0$. As $w_{k_0} \in \divideontimes_{d_p}(q)$, there is a sequence $\{x_i^{(k_0)}\}_{i\in\mathbb{N}} \subset B_{r(q)}(q) \setminus E_{d_p}$ such that $\lim_{i\to\infty} x_i^{(k_0)} = q$ and $\lim_{i\to\infty} (\operatorname{grad} d_p)_{x_i^{(k_0)}} = w_{k_0}$. Now M is complete, so for each $i \in \mathbb{N}$ there is a minimal geodesic segment $\gamma_i : [0, d_p(x_i^{(k_0)})] \to M$ emanating from p to $x_i^{(k_0)}$, and hence we obtain the sequence $\{\gamma_i\}_{i\in\mathbb{N}}$ of such geodesics. The set $\mathbb{S}_p^{m-1} := \{u \in T_pM \,|\, \|u\| = 1\}$ is compact, so we can assume, by taking a subsequence of $\{(d\gamma_i/dt)(0)\}_{i\in\mathbb{N}} \subset \mathbb{S}_p^{m-1}$ if necessary, that $\lim_{i\to\infty}(d\gamma_i/dt)(0) \in \mathbb{S}_p^{m-1}$ exists. Set $u := \lim_{i\to\infty}(d\gamma_i/dt)(0)$. Since $\lim_{i\to\infty} x_i^{(k_0)} = q$, $\{\gamma_i\}_{i\in\mathbb{N}}$ converges to a minimal geodesic segment $\gamma_\infty : [0, d_p(q)] \to M$ emanating from p to q given by $\gamma_\infty(t) = \exp_p tu$. Moreover, from [39, Proposition 4.8 of Chap. III] we have $(\operatorname{grad} d_p)_{x_i^{(k_0)}} = (d\gamma_i/dt)(d_p(x_i^{(k_0)}))$ for all $i \in \mathbb{N}$. Note that $\|(\operatorname{grad} d_p)_{x_i^{(k_0)}}\| = 1$ for each $i \in \mathbb{N}$. Since $w_{k_0} = \lim_{i\to\infty}(\operatorname{grad} d_p)_{x_i^{(k_0)}} = (d\gamma_\infty/dt)(d_p(q))$ and $\langle w_{k_0}, v\rangle \leq 0$, we see that $0 \geq \langle w_{k_0}, v\rangle = \|v\| \cos\angle((d\gamma_\infty/dt)(d_p(q)), v)$, hence $\angle(-(d\gamma_\infty/dt)(d_p(q)), v) \leq \pi/2$ holds for all $v \in T_qM$, and finally q is therefore critical for d_p in the sense of Grove–Shiohama.

We next assume that q is critical for d_p in the sense of Grove–Shiohama. Fix $v \in \mathbb{S}_q^{m-1} := \{u \in T_qM \mid \|u\| = 1\}$. There is a minimal geodesic segment $\sigma^{(v)} : [0, d_p(q)] \to M$ emanating from p to q such that $\angle(-(d\sigma^{(v)}/dt)(d_p(q)), v) \leq \pi/2$. As $\exp_q|_{\mathbb{B}_{r(q)}(o_q)}$ is a diffeomorphism onto $B_{r(q)}(q)$, we have a unique minimal geodesic $c_v : (-r(q), r(q)) \to B_{r(q)}(q)$ given by $c_v(s) := \exp_q sv$ for all $s \in (-r(q), r(q))$. Also $(\exp_q)^{-1}$ is a diffeomorphism from $B_{r(q)}(q)$ onto $\mathbb{B}_{r(q)}(o_q)$, so it follows from [48, Lemma 6.5] that we can choose a sequence $\{s_i\}_{i\in\mathbb{N}} \subset (-r(q), r(q))$ such that $\lim_{i\to\infty} s_i = 0$ and $c_v(s_i) \in B_{r(q)}(q) \setminus E_{d_p}$. To simplify notation we set $y_i := c_v(s_i)$ for each $i \in \mathbb{N}$. Note that $\lim_{i\to\infty} y_i = q$. Since M is complete, for each $i \in \mathbb{N}$ there is a minimal geodesic segment $\eta_i : [0, d_p(y_i)] \to M$ emanating from p to y_i. By the same argument above we can assume that $\{\eta_i\}_{i\in\mathbb{N}}$ converges to a minimal geodesic segment $\eta_\infty : [0, d_p(q)] \to M$ emanating from p to q. From [39, Proposition 4.8 of Chap. III], $(\operatorname{grad} d_p)_{y_i} = (d\eta_i/dt)(d_p(y_i))$ holds for each $i \in \mathbb{N}$. Now $\lim_{i\to\infty}(\operatorname{grad} d_p)_{y_i} = (d\eta_\infty/dt)(d_p(q))$, set $w^{(v)} := (d\eta_\infty/dt)(d_p(q))$, and $w^{(v)} = \lim_{i\to\infty}(\operatorname{grad} d_p)_{y_i} \in \circledast_{d_p}(q)$. Since $\angle(-(d\sigma^{(v)}/dt)(d_p(q)), v) \leq \pi/2$, [21, Lemma 2.1] shows $\angle(-w^{(v)}, v) = \angle(-\lim_{i\to\infty}(\operatorname{grad} d_p)_{y_i}, v) \leq \pi/2$, and hence we get $\angle(w^{(v)}, v) \geq \pi/2$. Since $v \in \mathbb{S}_q^{m-1}$ was arbitrary, for each $v \in \mathbb{S}_q^{m-1}$ we take $w^{(v)} \in \circledast_{d_p}(q)$ satisfying $\angle(w^{(v)}, v) \geq \pi/2$, and we set $W := \{w^{(v)} \in \circledast_{d_p}(q) \mid v \in \mathbb{S}_q^{m-1}\}$. Then W is not contained in an open half space of T_qM, which implies $o_q \in \operatorname{Conv}(W) \subset \circledast_{d_p}(q)$. Lemma 2.16 gives that q is singular for d_p. □

Lemma 2.20 *p is also singular for d_p in the sense of Clarke.*

Proof. Note that d_p is differentiable on $B_{r(p)}(p) \setminus \{p\}$, for the set has no cut points of p. We first show that $\circledast_F(p) = \mathbb{S}_p^{m-1}$. Indeed, since all the geodesics are normalized, [39, Proposition 4.8 of Chap. III] gives $\circledast_{d_p}(p) \subset \mathbb{S}_p^{m-1}$. Thus it is sufficient to prove $\mathbb{S}_p^{m-1} \subset \circledast_{d_p}(p)$. Fix $v \in \mathbb{S}_p^{m-1}$. Let $\sigma_v : (-r(p), r(p)) \to B_{r(p)}(p)$ be a minimal geodesic defined by $\sigma_v(t) := \exp_p tv$ for all $t \in (-r(p), r(p))$. Let $\{t_i\}_{i\in\mathbb{N}}$ be a sequence of constants $t_i \in (-r(p), r(p)) \setminus \{0\}$ converging to 0 by letting $i \to \infty$. Set $x_i := \sigma_v(t_i)$ for each $i \in \mathbb{N}$. Note that $x_i \in B_{r(p)}(p) \setminus \{p\}$. Combining the Gauss lemma (cf. [39, (1) of Proposition 2.3 of Chap. III]) and [39, Proposition 4.8 of Chap. III] gives that $(\operatorname{grad} d_p)_{\sigma_v(t_i)} = (d\sigma_v/dt)(t_i) = (d\exp_p)_{t_i v} v$. Since $\lim_{i\to\infty}(\operatorname{grad} d_p)_{\sigma_v(t_i)} = \lim_{i\to\infty}(d\exp_p)_{t_i v} v = (d\exp_p)_{o_p} v = v$, we get $v \in \circledast_{d_p}(p)$, i.e., $\mathbb{S}_p^{m-1} \subset \circledast_{d_p}(p)$ holds. Therefore $\circledast_{d_p}(p) = \mathbb{S}_p^{m-1}$.

Since $\circledast_{d_p}(p) = \mathbb{S}_p^{m-1}$, $\circledast_{d_p}(p) = \operatorname{Conv}(\mathbb{S}_p^{m-1}) = \{X \in T_pM \mid \|X\| \leq 1\}$ holds, hence $o_p \in \circledast_{d_p}(p)$, and finally Lemma 2.16 shows that p is singular for d_p in that sense. □

3 The adjoint of the generalized differential of Lipschitz maps

In this section we define the adjoint of the generalized differential of Lipschitz maps between Riemannian manifolds, study surjectivity and injectivity of the generalized differential near their nonsingular points, and finally show that the set of all singular points of the map is closed.

3.1 Definition of the adjoint of the generalized differential

In this subsection we formulate the notion of the adjoint of the generalized differential of Lipschitz maps between Riemannian manifolds. Throughout this subsection let M and

N be Riemannian manifolds with dimension m and n and Riemannian metrics $\langle \cdot, \cdot \rangle_M$ and $\langle \cdot, \cdot \rangle_N$, respectively. $F : M \to N$ is a locally Lipschitz map, and $K_F(p)$ the mixture of the differential of F at $p \in M$.

Fix $p \in M$. Choose two strongly convex open balls $B_{r(p)}(p) \subset M$ and $B_{t(p)}(F(p)) \subset N$ satisfying all three properties (i)–(iii) in Lemma 2.2. Let $\mathcal{L}(T_{F(p)}N, T_pM)$ be the mn-dimensional vector space of all linear mappings of $T_{F(p)}N$ to T_pM topologized with the operator norm $\|\cdot\|$. Consider two nonempty sets

$$\mathrm{adj}(K_F(p)) := \{G^* \in \mathcal{L}(T_{F(p)}N, T_pM) \,|\, G \in K_F(p)\}$$

where G^* denotes the adjoint of G, and

$$\{K_F(p)\}^* := \left\{ H^* \in \mathcal{L}(T_{F(p)}N, T_pM) \;\middle|\; \begin{array}{l} \exists \{x_i\}_{i\in\mathbb{N}} \subset B_{r(p)}(p) \setminus E_F \ \text{ such that} \\ \lim\limits_{i\to\infty} x_i = p, \ \lim\limits_{i\to\infty} \tau_p^{x_i} \circ (dF_{x_i})^* \circ \tau_{F(x_i)}^{F(p)} = H^* \end{array} \right\}$$

where $\tau_p^{x_i}$ and $\tau_{F(x_i)}^{F(p)}$ are parallel transports as in Definition 2.3, and $(dF_{x_i})^*$ denotes the adjoint of dF_{x_i} at each x_i. Note that, from the (ii) of Lemma 2.2, each $\tau_{F(x_i)}^{F(p)} : T_{F(p)}N \to T_{F(x_i)}N$ is defined in that sense.

Lemma 3.1 $\mathrm{adj}(K_F(p)) = \{K_F(p)\}^*$.

Proof. Fix $G^* \in \mathrm{adj}(K_F(p))$. Since $(G^*)^* = G \in K_F(p)$ by definition of $\mathrm{adj}(K_F(p))$, there is a sequence $\{x_i\}_{i\in\mathbb{N}} \subset B_{r(p)}(p) \setminus E_F$ such that $\lim_{i\to\infty} x_i = p$ and $\lim_{i\to\infty} \tau_{F(p)}^{F(x_i)} \circ dF_{x_i} \circ \tau_{x_i}^p = G$. Note that the adjoints of $\tau_{x_i}^p$ and $\tau_{F(p)}^{F(x_i)}$ are their inverses $\tau_p^{x_i}$ and $\tau_{F(x_i)}^{F(p)}$ since parallel transport is an isometry. Fix $u \in T_pM$ and $v \in T_{F(p)}N$. The Riesz representation theorem (cf. [38, Theorem 10.1]) then gives

$$\langle u, G^*(v)\rangle_M = \langle (\lim_{i\to\infty} \tau_{F(p)}^{F(x_i)} \circ dF_{x_i} \circ \tau_{x_i}^p)(u), v\rangle_N = \langle u, (\lim_{i\to\infty} \tau_p^{x_i} \circ (dF_{x_i})^* \circ \tau_{F(x_i)}^{F(p)})(v)\rangle_M,$$

which implies $G^* \in \{K_F(p)\}^*$, and hence $\mathrm{adj}(K_F(p)) \subset \{K_F(p)\}^*$. Our next claim is that $\{K_F(p)\}^* \subset \mathrm{adj}(K_F(p))$. Indeed, for any fixed $H^* \in \{K_F(p)\}^*$ there is a sequence $\{x_i\}_{i\in\mathbb{N}} \subset B_{r(p)}(p) \setminus E_F$ such that $\lim_{i\to\infty} x_i = p$ and that $\lim_{i\to\infty} \tau_p^{x_i} \circ (dF_{x_i})^* \circ \tau_{F(x_i)}^{F(p)} = H^*$. For any $u \in T_pM$ and any $v \in T_{F(p)}N$ we have

$$\langle u, H^*(v)\rangle_M = \langle u, (\lim_{i\to\infty} \tau_p^{x_i} \circ (dF_{x_i})^* \circ \tau_{F(x_i)}^{F(p)})(v)\rangle_M = \langle u, (\lim_{i\to\infty} \tau_{F(p)}^{F(x_i)} \circ dF_{x_i} \circ \tau_{x_i}^p)^*(v)\rangle_M,$$

which implies that $H^* = (\lim_{i\to\infty} \tau_{F(p)}^{F(x_i)} \circ dF_{x_i} \circ \tau_{x_i}^p)^*$. Since the adjoint $(H^*)^*$ of H^* is unique by the Riesz representation theorem (cf. [38, Theorem 10.1]), $\lim_{i\to\infty} \tau_{F(p)}^{F(x_i)} \circ dF_{x_i} \circ \tau_{x_i}^p = (H^*)^* \in K_F(p)$, and hence $\{K_F(p)\}^* \subset \mathrm{adj}(K_F(p))$. Therefore $\mathrm{adj}(K_F(p)) = \{K_F(p)\}^*$. ☐

Lemma 3.2 *Let* $\mathrm{adj}(\partial F(p)) := \{g^* \in \mathcal{L}(T_{F(p)}N, T_pM) \,|\, g \in \partial F(p)\}$ *where* $\partial F(p)$ *is the generalized differential of* F *at* p *and* g^* *denotes the adjoint of each* $g \in \partial F(p)$*. Then*

(i) $\mathrm{adj}(\partial F(p))$ *is a nonempty and compact subset of* $\mathcal{L}(T_{F(p)}N, T_pM)$*;*

(ii) $\mathrm{adj}(\partial F(p)) = \mathrm{Conv}(\{K_F(p)\}^*)$.

Proof. We first show (i). Since $\partial F(p) \neq \emptyset$, the Riesz representation theorem (cf. [38, Theorem 10.1]) guarantees $\mathrm{adj}(\partial F(p)) \neq \emptyset$. Take any sequence $\{g_i^*\}_{i\in\mathbb{N}} \subset \mathrm{adj}(\partial F(p))$ where each g_i^* is the adjoint of $g_i \in \partial F(p)$. Since $\partial F(p)$ is compact, the sequence $\{g_i\}_{i\in\mathbb{N}} \subset \partial F(p)$ contains a subsequence $\{g_{i_k}\}_{k\in\mathbb{N}}$ which converges to some point $h \in \partial F(p)$ as $k \to \infty$. For any $u \in T_pM$ and any $v \in T_{F(p)}N$ we have $\langle u, h^*(v)\rangle_M = \langle h(u), v\rangle_N = \langle(\lim_{k\to\infty} g_{i_k})(u), v\rangle_N = \langle u, (\lim_{k\to\infty} g_{i_k}^*)(v)\rangle_M$, hence $\lim_{k\to\infty} g_{i_k}^* = h^* \in \mathrm{adj}(\partial F(p))$. Since $\{g_{i_k}\}_{k\in\mathbb{N}} \subset \{g_i\}_{i\in\mathbb{N}}$, and since $g_{i_k}^*$ is the adjoint of g_{i_k}, $\{g_i^*\}_{i\in\mathbb{N}}$ contains $\{g_{i_k}^*\}_{k\in\mathbb{N}}$ as a subsequence converging to $h^* \in \mathrm{adj}(\partial F(p))$, which implies that $\mathrm{adj}(\partial F(p))$ is compact.

Next we show (ii). Let $g^* \in \mathrm{adj}(\partial F(p))$ where $g \in \partial F(p)$. From Carathéodory's theorem (cf. [40, Theorem 1.1.4]) there are $g_1, g_2, \ldots, g_{nm+1} \in K_F(p)$ such that $g^* = (\sum_{i=1}^{nm+1} a_i g_i)^*$ where $\sum_{i=1}^{nm+1} a_i = 1$ and $a_i \geq 0$ $(i = 1, 2, \ldots, nm+1)$. Since $g^* = \sum_{i=1}^{nm+1} a_i g_i^*$, and since Lemma 3.1 gives $g_i^* \in \mathrm{adj}(K_F(p)) = \{K_F(p)\}^*$ $(i = 1, 2, \ldots, nm+1)$, $g^* \in \mathrm{Conv}(\{K_F(p)\}^*)$, i.e., $\mathrm{adj}(\partial F(p)) \subset \mathrm{Conv}(\{K_F(p)\}^*)$. The similar discussion shows $\mathrm{Conv}(\{K_F(p)\}^*) \subset \mathrm{adj}(\partial F(p))$, and hence $\mathrm{adj}(\partial F(p)) = \mathrm{Conv}(\{K_F(p)\}^*)$. □

Lemma 3.2 justifies the following definition.

Definition 3.3 We call the set $\{\partial F(p)\}^* := \mathrm{adj}(\partial F(p))$ the *adjoint of* $\partial F(p)$.

Lemma 3.4 *$\{\partial F(p)\}^*$ is a nonempty, compact, and convex subset of $\mathcal{L}(T_{F(p)}N, T_pM)$.*

Proof. This statement follows from Lemma 3.2 and [37, Theorem 17.2]. □

3.2 Surjectivity and injectivity of the generalized differential near a nonsingular point

All notation in this subsection is the same as defined in Section 3.1. First we show surjectivity of the generalized differential of F near a nonsingular point when $m \geq n$, and next show injectivity when $m \leq n$. Finally we see that the set of all singular points of F is a closed set in M.

Lemma 3.5 *For any $p \in M$ and any $\varepsilon > 0$ there is a constant $\mu(p, \varepsilon) \in (0, r(p))$ such that $\tau_p^x \circ \{\partial F(x)\}^* \circ \tau_{F(x)}^{F(p)} \subset \mathcal{U}_\varepsilon(\{\partial F(p)\}^*)$ for all $F(x) \in B_{t(p)}(F(p))$ $(x \in B_{\mu(p,\varepsilon)}(p))$ where $\tau_p^x \circ \{\partial F(x)\}^* \circ \tau_{F(x)}^{F(p)} := \{\tau_p^x \circ g^* \circ \tau_{F(x)}^{F(p)} \mid g^* \in \{\partial F(x)\}^*\}$ and $\mathcal{U}_\varepsilon(\{\partial F(p)\}^*)$ denotes the ε-open neighborhood of $\{\partial F(p)\}^*$ in $\mathcal{L}(T_{F(p)}N, T_pM)$.*

Proof. This is a direct consequence of Definition 3.3. □

Proposition 3.6 *Assume $m \geq n$. If a point $p \in M$ is nonsingular for F, then there are two constants $\lambda(p) > 0$ and $\delta(p) > 0$ satisfying the following properties:*

(i) *$B_{2\lambda(p)}(p)$ satisfies* (a) *and* (b) *of Theorem 2.1;*

(ii) *$F|_{B_{2\lambda(p)}(p)}$ is a Lipschitz map from $B_{2\lambda(p)}(p)$ into $B_{t(p)}(F(p))$;*

(iii) *For any $u \in \mathbb{S}_{F(p)}^{n-1} := \{w \in T_{F(p)}N \mid \|w\| = 1\}$ and any $x \in B_{2\lambda(p)}(p)$ there is a vector $V_x^{(u)} \in \mathbb{S}_x^{m-1} := \{v \in T_xM \mid \|v\| = 1\}$ such that $\langle V_x^{(u)}, (g^* \circ \tau_{F(x)}^{F(p)})(u)\rangle_M \geq \delta(p)$ holds for all $g^* \in \{\partial F(x)\}^*$. In particular $\langle V_x^{(u)}, (dF_x)^*(\tau_{F(x)}^{F(p)}(u))\rangle_M \geq \delta(p)$ for all $x \in B_{2\lambda(p)}(p) \setminus E_F$;*

(iv) *Every $x \in B_{2\lambda(p)}(p)$ is nonsingular for F.*

Proof. Fix $p \in M$ nonsingular for F. By Definition 2.8, $\mathrm{rank}(g) = n$ holds for all $g \in \partial F(p)$. Since $\mathrm{rank}(g^*) = \mathrm{rank}(g) = n$ for all $g^* \in \{\partial F(p)\}^*$, $\{\partial F(p)\}^*$ has maximal rank. Every $g^* \in \{\partial F(p)\}^*$ is therefore injective.

Take $u \in \mathbb{S}^{n-1}_{F(p)}$. Set $\{\partial F(p)\}^* u := \{g^*(u) \mid g^* \in \{\partial F(p)\}^*\} \subset T_pM$. Lemma 3.4 implies that $\{\partial F(p)\}^* u$ is compact and convex in T_pM. Since each $g^* \in \{\partial F(p)\}^*$ is injective as we have seen above, $o_p \notin \{\partial F(p)\}^* u$ holds where o_p indicates the origin of T_pM. Since $\{\partial F(p)\}^* u$ is compact and convex, there is a point $a^{(u)}$ in the boundary $\mathrm{Bd}(\{\partial F(p)\}^* u)$ such that $\|a^{(u)}\| = d_{T_pM}(o_p, \{\partial F(p)\}^* u) > 0$ where d_{T_pM} denotes the distance function of T_pM, i.e., $d_{T_xM}(a, b) := \|a - b\|$ for all $a, b \in T_xM$ $(x \in M)$. Since $\mathbb{S}^{n-1}_{F(p)}$ is compact, there is a constant $\delta(p) > 0$ given by $\delta(p) := \min\{\|a^{(w)}\| \mid w \in \mathbb{S}^{n-1}_{F(p)}\}/2$, and hence $d_{T_pM}(o_p, \{\partial F(p)\}^* u) = \|a^{(u)}\| \geq 2\delta(p)$. By this inequality there is a constant $\varepsilon(p) > 0$ sufficiently small such that

$$d_{T_pM}(o_p, \overline{\mathcal{U}_{\varepsilon(p)}(\{\partial F(p)\}^*)} u) \geq \delta(p) \tag{3.1}$$

where $\overline{\mathcal{U}_{\varepsilon(p)}(\{\partial F(p)\}^*)}$ is the closure of the $\varepsilon(p)$-open neighborhood $\mathcal{U}_{\varepsilon(p)}(\{\partial F(p)\}^*)$ of $\{\partial F(p)\}^*$ in $\mathcal{L}(T_{F(p)}N, T_pM)$. Note that $\overline{\mathcal{U}_{\varepsilon(p)}(\{\partial F(p)\}^*)} u$ is a compact convex subset of T_pM. Indeed, let $\overline{\mathbb{B}_{\varepsilon(p)}(\tilde{o})}$ be a closed ball with centre the origin $\tilde{o}$ of $\mathcal{L}(T_{F(p)}N, T_pM)$ and radius $\varepsilon(p)$. Since $\overline{\mathbb{B}_{\varepsilon(p)}(\tilde{o})}$ and $\{\partial F(p)\}^*$ are convex in $\mathcal{L}(T_{F(p)}N, T_pM)$, and since $\overline{\mathcal{U}_{\varepsilon(p)}(\{\partial F(p)\}^*)} = \{\partial F(p)\}^* + \overline{\mathbb{B}_{\varepsilon(p)}(\tilde{o})}$, [37, Theorem 3.1] shows that $\overline{\mathcal{U}_{\varepsilon(p)}(\{\partial F(p)\}^*)}$ is a compact convex subset in $\mathcal{L}(T_{F(p)}N, T_pM)$, and hence $\overline{\mathcal{U}_{\varepsilon(p)}(\{\partial F(p)\}^*)} u$ is also in T_pM.

Fix $x \in B_{r(p)}(p)$. Since $\|g^*(u)\| \geq \delta(p)$ for all $g^* \in \overline{\mathcal{U}_{\varepsilon(p)}(\{\partial F(p)\}^*)}$ by Eq. (3.1), and since τ^p_x is an isometry, $\|\tau^p_x(g^*(u))\| = \|g^*(u)\| \geq \delta(p)$ holds for all $g^* \in \overline{\mathcal{U}_{\varepsilon(p)}(\{\partial F(p)\}^*)}$, and hence $d_{T_xM}(o_x, \tau^p_x(\overline{\mathcal{U}_{\varepsilon(p)}(\{\partial F(p)\}^*)} u)) \geq \delta(p)$. Since τ^p_x is linear, $\tau^p_x(\overline{\mathcal{U}_{\varepsilon(p)}(\{\partial F(p)\}^*)} u)$ is a compact convex subset of T_xM. There is therefore a point $b^{(u,x)} \in \mathrm{Bd}(\tau^p_x(\overline{\mathcal{U}_{\varepsilon(p)}(\{\partial F(p)\}^*)} u))$ such that

$$\|b^{(u,x)}\| = d_{T_xM}(o_x, \tau^p_x(\overline{\mathcal{U}_{\varepsilon(p)}(\{\partial F(p)\}^*)} u)) \geq \delta(p). \tag{3.2}$$

Define a unit tangent vector $V^{(u)}_x$ at x by $V^{(u)}_x := b^{(u,x)}/\|b^{(u,x)}\| \in \mathbb{S}^{m-1}_x$. Lemma 3.5 shows that for $\varepsilon(p)$ as above there is a constant $\lambda(p) \in (0, r(p)/2)$ for $\mu(p, \varepsilon) = 2\lambda(p) := 2\lambda(p, \varepsilon(p))$ such that $\tau^x_p \circ \{\partial F(x)\}^* \circ \tau^{F(p)}_{F(x)} \subset \mathcal{U}_{\varepsilon(p)}(\{\partial F(p)\}^*)$ holds for all $x \in B_{2\lambda(p)}(p)$. Lemma 2.2 gives assertions (i) and (ii). Since $\mathcal{U}_{\varepsilon(p)}(\{\partial F(p)\}^*) \subset \overline{\mathcal{U}_{\varepsilon(p)}(\{\partial F(p)\}^*)}$, for any $x \in B_{2\lambda(p)}(p)$ we obtain

$$\{\partial F(x)\}^* \tau^{F(p)}_{F(x)}(u) \subset \tau^p_x(\overline{\mathcal{U}_{\varepsilon(p)}(\{\partial F(p)\}^*)} u) \subset T_xM. \tag{3.3}$$

Define the line $\ell : \mathbb{R} \to T_xM$ by $\ell(t) := t\, V^{(u)}_x$. Note that ℓ is passing through $\tau^p_x(\overline{\mathcal{U}_{\varepsilon(p)}(\{\partial F(p)\}^*)} u)$. Fix $g^* \in \{\partial F(x)\}^*$. From Eq. (3.3) there is a unique constant $t_0 > 0$ such that

$$t_0 \geq \|b^{(u,x)}\| \quad \text{and} \quad \angle(\overrightarrow{o_x\, \ell(t_0)}, \overrightarrow{\ell(t_0) g^*(\tau^{F(p)}_{F(x)}(u))}) = \frac{\pi}{2}, \tag{3.4}$$

i.e., $\ell(t_0)$ is the foot of the perpendicular from $g^*(\tau^{F(p)}_{F(x)}(u))$ to the line ℓ. Set $\theta := \angle(\overrightarrow{o_x\, \ell(t_0)}, \overrightarrow{o_x\, g^*(\tau^{F(p)}_{F(x)}(u))})$. Note here that $\theta \in [0, \pi/2)$ because $\tau^p_x(\overline{\mathcal{U}_{\varepsilon(p)}(\{\partial F(p)\}^*)} u)$ is convex in T_xM and $o_x \notin \tau_x^p(\overline{\mathcal{U}_{\varepsilon(p)}(\{\partial F(p)\}^*)u})$. It follows from Eqs. (3.2) and (3.4) that

$$\langle V_x^{(u)}, g^*(\tau_{F(x)}^{F(p)}(u))\rangle_M = \|g^*(\tau_{F(x)}^{F(p)}(u))\| \cos\theta = \|\ell(t_0)\| = t_0 \geq \|b^{(u,x)}\| \geq \delta(p), \tag{3.5}$$

which is assertion (iii).

Furthermore, Eq. (3.5) gives $\delta(p) \leq \langle V_x^{(u)}, (g^* \circ \tau_{F(x)}^{F(p)})(u)\rangle_M \leq \|g^*(\tau_{F(x)}^{F(p)}(u))\|$, which shows that every $g^* \in \{\partial F(x)\}^*$ is injective for all $x \in B_{2\lambda(p)}(p)$. Since $\mathrm{rank}(g) = \mathrm{rank}(g^*) = n$ for all $g \in \partial F(x)$ as $x \in B_{2\lambda(p)}(p)$, any point $x \in B_{2\lambda(p)}(p)$ is nonsingular for F. Assertion (iv) thus holds. □

Proposition 3.7 *Assume $m \leq n$. If a point $p \in M$ is nonsingular for F, then there are two constants $\lambda(p) > 0$ and $\delta(p) > 0$ satisfying the following properties:*

(i) *$B_{2\lambda(p)}(p)$ satisfies* (a) *and* (b) *of Theorem 2.1;*

(ii) *$F|_{B_{2\lambda(p)}(p)}$ is a Lipschitz map from $B_{2\lambda(p)}(p)$ into $B_{t(p)}(F(p))$;*

(iii) *For any $u \in \mathbb{S}_p^{m-1} := \{w \in T_pM \mid \|w\| = 1\}$ and any $x \in B_{2\lambda(p)}(p)$ there is a vector $V_{F(x)}^{(u)} \in \mathbb{S}_{F(x)}^{n-1} := \{v \in T_{F(x)}N \mid \|v\| = 1\}$ such that $\langle (g \circ \tau_x^p)(u), V_{F(x)}^{(u)}\rangle_N \geq \delta(p)$ holds for all $g \in \partial F(x)$;*

(iv) *Every $x \in B_{2\lambda(p)}(p)$ is nonsingular for F.*

Proof. Fix $p \in M$ nonsingular for F, and $u \in \mathbb{S}_p^{m-1}$. Set $\partial F(p)u := \{g(u) \mid g \in \partial F(p)\}$. Remark 2.7 (ii) shows that $\partial F(p)u$ is compact and convex in $T_{F(p)}N$. Since p is nonsingular for F, $\partial F(p)$ has maximal rank m, hence every $g \in \partial F(p)$ is injective, and $o_{F(p)} \notin \partial F(p)u$. Thanks to Remark 2.7 (vi), the same argument as in the proof of Proposition 3.6 works for $\partial F(p)u$. Details are left to the reader. □

Remark 3.8 We give here three remarks on Propositions 3.6 and 3.7.

(i) Proposition 3.6 is a completely new result.

(ii) Clarke first showed the same statement as in Proposition 3.7 in the case both M and N are Euclidean spaces of the same dimension, see [8, Lemma 3].

(iii) Without mentioning Proposition 3.7 we applied it in the proof of [26, Lemma 2.21]. We did not give the proof there. It is provided here.

Corollary 3.9 *The set of all singular points of F is a closed set in M.*

Proof. Let $\mathrm{Sing}(F)$ be the set of all singular points of F. We will show that $M \setminus \mathrm{Sing}(F)$ is open in M. Fix $p \in M \setminus \mathrm{Sing}(F)$. We first assume $m \geq n$. By the property (iv) of Proposition 3.6 we can find a constant $\lambda(p) > 0$ such that any point $x \in B_{2\lambda(p)}(p)$ is nonsingular for F, hence $B_{2\lambda(p)}(p) \subset M \setminus \mathrm{Sing}(F)$, and finally $M \setminus \mathrm{Sing}(F)$ is open. By applying that of Proposition 3.7 the same proof works for $m \leq n$. □

4 Smooth approximation of Lipschitz maps between Riemannian manifolds

Working from results in [14], [16], [19], [20], [22], and [41], we define a smooth approximation of an arbitrary Lipschitz map between Riemannian manifolds. Throughout this section let M be a compact Riemannian manifold of dimension m, N a Riemannian manifold of dimension n, d_M and d_N the distance functions of M and N, respectively, $F : M \to N$ a Lipschitz map, and $\mathrm{inj}(M)$ the injectivity radius of M. Note that $(0, \mathrm{inj}(M)/2) \subset \mathbb{R}$ is not empty because M is compact.

Lemma 4.1 *There is a finite set $\{p_1, p_2, \ldots, p_k\} \subset M$ such that*

(I) *for each $p_i \in \{p_1, p_2, \ldots, p_k\}$ both $B_{r(p_i)}(p_i) \subset M$ and $B_{t(p_i)}(F(p_i)) \subset N$ satisfy the properties (i) and (ii) of Lemma 2.2 for $p = p_i$;*

(II) *$r(p_i) \in (0, \mathrm{inj}(M)/2)$ for all $p_i \in \{p_1, p_2, \ldots, p_k\}$;*

(III) *$M = \bigcup_{i=1}^{k} B_{r(p_i)}(p_i)$.*

Proof. This follows immediately since M is compact. □

By applying the Nash embedding theorem [33], N can be isometrically embedded into the Euclidean space $\mathbb{R}^\ell$ with the canonical Riemannian metric $\langle \cdot, \cdot \rangle$ where $\ell \geq \max\{m, n+1\}$. F can be regarded as a Lipschitz map from M into $\mathbb{R}^\ell$, and hence we set

$$\widetilde{F} := F : M \to N \subset \mathbb{R}^\ell.$$

In the case where $N = \mathbb{R}$, this is not done.

From now on we use the notation $\mathrm{inj}(M)$, $\{B_{r(p_i)}(p_i)\}_{i=1}^k$, $\{B_{t(p_i)}(F(p_i))\}_{i=1}^k$, and $\widetilde{F}$ in the sense above.

4.1 The local smooth approximation of Lipschitz maps

In this subsection we define the local smooth approximation of $\widetilde{F} : M \to N \subset \mathbb{R}^\ell$ on each strongly convex ball $B_{r(p_i)}(p_i) \subset M$ with convolution smoothing.

Fix $p_i \in \{p_1, p_2, \ldots, p_k\}$. Since $\exp_{p_i}|_{\mathbb{B}_{\mathrm{inj}(M)}(o_{p_i})} : \mathbb{B}_{\mathrm{inj}(M)}(o_{p_i}) \to B_{\mathrm{inj}(M)}(p_i)$ is a diffeomorphism, we can define the map $\mathcal{F}^{(i)} : \mathbb{B}_{\mathrm{inj}(M)}(o_{p_i}) \to N \subset \mathbb{R}^\ell$ by

$$\mathcal{F}^{(i)} := \widetilde{F} \circ \exp_{p_i}|_{\mathbb{B}_{\mathrm{inj}(M)}(o_{p_i})}.$$

Choose an orthonormal basis $e_1^{(i)}, e_2^{(i)}, \ldots, e_m^{(i)}$ for $T_{p_i}M$. Using coordinates $(y_1^{(i)}, y_2^{(i)}, \ldots, y_m^{(i)})$ with respect to $e_1^{(i)}, e_2^{(i)}, \ldots, e_m^{(i)}$ on $T_{p_i}M$, we identify $T_{p_i}M$ with $\mathbb{R}^m$. Let $(z_1, z_2, \ldots, z_\ell)$ be the standard coordinates of $\mathbb{R}^\ell$. We then have the coordinate representation $\mathcal{F}^{(i)} = (\mathcal{F}_1^{(i)}, \mathcal{F}_2^{(i)}, \ldots, \mathcal{F}_\ell^{(i)})$ of $\mathcal{F}^{(i)}$ defined by $\mathcal{F}_j^{(i)} := z_j \circ \mathcal{F}^{(i)}$ for each $j \in \{1, 2, \ldots, \ell\}$. Moreover let $\rho^{(i)} : T_{p_i}M \to \mathbb{R}$ be a smooth function given by

$$\rho^{(i)}(y) = \begin{cases} \alpha \cdot e^{-1/(1-\|y\|^2)} & (y \in \mathbb{B}_1(o_{p_i})), \\ 0 & (y \in T_{p_i}M \setminus \mathbb{B}_1(o_{p_i})) \end{cases}$$

where the constant α is chosen so that $\int_{y \in T_{p_i}M} \rho^{(i)}(y)dy = 1$. For an $\varepsilon \in (0, \mathrm{inj}(M)/2)$ the Riemannian mollifier $\rho_\varepsilon^{(i)}$ is then defined by $\rho_\varepsilon^{(i)}(y) := \rho^{(i)}(y/\varepsilon)/\varepsilon^m$ for all $y \in T_{p_i}M$, which is a nonnegative smooth function on $T_{p_i}M$ and satisfies

$$\text{(4.1)} \qquad \operatorname{supp} \rho_\varepsilon^{(i)} = \overline{\mathbb{B}_\varepsilon(o_{p_i})} \quad \text{and} \quad \int_{T_{p_i}M} \rho_\varepsilon^{(i)}(y)dy = 1,$$

see for instance [20], [28], or [52]. We now define the convolution smoothing of $\widetilde{F}$.

Definition 4.2 Fix $p_i \in \{p_1, p_2, \ldots, p_k\}$ and $\varepsilon \in (0, \mathrm{inj}(M)/2)$. The map $\widetilde{F}_\varepsilon^{(p_i)} : B_{r(p_i)}(p_i) \to \mathbb{R}^\ell$ is defined as follows. For any $q \in B_{r(p_i)}(p_i)$,

$$\text{(4.2)} \quad \begin{aligned} \widetilde{F}_\varepsilon^{(p_i)}(q) &:= \int_{y \in T_{p_i}M} \rho_\varepsilon^{(i)}(y)\mathcal{F}^{(i)}(\exp_{p_i}^{-1} q - y)dy \\ &:= \Big(\int_{T_{p_i}M} \rho_\varepsilon^{(i)}(y)\mathcal{F}_1^{(i)}(\exp_{p_i}^{-1} q - y)dy, \ldots, \int_{T_{p_i}M} \rho_\varepsilon^{(i)}(y)\mathcal{F}_\ell^{(i)}(\exp_{p_i}^{-1} q - y)dy \Big). \end{aligned}$$

Remark 4.3 (i) Since

$$\int_{T_{p_i}M} \rho_\varepsilon^{(i)}(y)\mathcal{F}_j^{(i)}(\exp_{p_i}^{-1} q - y)dy = \int_{T_{p_i}M} \rho_\varepsilon^{(i)}(\exp_{p_i}^{-1} q - y)\mathcal{F}_j^{(i)}(y)dy$$

for each $j = 1, 2, \ldots, \ell$ (see for instance [20], [28], or [52]), we have, for any $q \in B_{r(p_i)}(p_i)$,

$$\widetilde{F}_\varepsilon^{(p_i)}(q) = \int_{T_{p_i}M} \rho_\varepsilon^{(i)}(y)\mathcal{F}^{(i)}(\exp_{p_i}^{-1} q - y)dy = \int_{T_{p_i}M} \rho_\varepsilon^{(i)}(\exp_{p_i}^{-1} q - y)\mathcal{F}^{(i)}(y)dy.$$

(ii) Fix $q \in B_{r(p_i)}(p_i)$ and $y \in \mathbb{B}_\varepsilon(o_{p_i})$. We see, by Lemma 4.1 (II), that

$$\text{(4.3)} \qquad \|\exp_{p_i}^{-1} q - y\| \leq \|\exp_{p_i}^{-1} q\| + \|y\| < r(p_i) + \varepsilon < \mathrm{inj}(M),$$

and hence $\mathcal{F}^{(i)}(\exp_{p_i}^{-1} q - y)$ exists. Moreover, since $\operatorname{supp} \rho_\varepsilon^{(i)} = \overline{\mathbb{B}_\varepsilon(o_{p_i})}$,

$$\text{(4.4)} \qquad \widetilde{F}_\varepsilon^{(p_i)}(q) = \int_{y \in \mathbb{B}_\varepsilon(o_{p_i})} \rho_\varepsilon^{(i)}(y)\mathcal{F}^{(i)}(\exp_{p_i}^{-1} q - y)dy$$

holds, and hence $\widetilde{F}_\varepsilon^{(p_i)}(q)$ exists.

(iii) Since each $\int_{T_{p_i}M} \rho_\varepsilon^{(i)}(y)\mathcal{F}_j^{(i)}(\exp_{p_i}^{-1} q - y)dy$ is smooth (see for instance [20], [28], or [52]), $\widetilde{F}_\varepsilon^{(p_i)}$ is smooth.

(iv) In the case where $N = \mathbb{R}$, the convolution smoothing (4.2) of the Lipschitz function $F : M \to \mathbb{R}$ is given by

$$\text{(4.5)} \qquad \widetilde{F}_\varepsilon^{(p_i)}(q) := \int_{y \in T_{p_i}M} \rho_\varepsilon^{(i)}(y)(F \circ \exp_{p_i})(\exp_{p_i}^{-1} q - y)dy$$

for all $q \in B_{r(p_i)}(p_i)$.

Definition 4.4 For each $\varepsilon \in (0, \text{inj}(M)/2)$ let

$$\Lambda(\varepsilon) := \max\{\text{Lip}(\exp_{p_i}|_{\mathbb{B}_{r(p_i)+\varepsilon}(o_{p_i})}) \mid p_i \in \{p_1, p_2, \ldots, p_k\}\} \tag{4.6}$$

where $\text{Lip}(\exp_{p_i}|_{\mathbb{B}_{r(p_i)+\varepsilon}(o_{p_i})})$ is the Lipschitz constant of $\exp_{p_i}|_{\mathbb{B}_{r(p_i)+\varepsilon}(o_{p_i})}$, i.e.,

$$\text{Lip}(\exp_{p_i}|_{\mathbb{B}_{r(p_i)+\varepsilon}(o_{p_i})}) := \sup\left\{ \frac{d_M(\exp_{p_i} u, \exp_{p_i} v)}{\|u - v\|} \,\middle|\, u, v \in \mathbb{B}_{r(p_i)+\varepsilon}(o_{p_i}), u \neq v \right\}.$$

Remark 4.5 Since $r(p_i) + \varepsilon < \text{inj}(M)$ for each $p_i \in \{p_1, p_2, \ldots, p_k\}$, $\exp_{p_i}|_{\mathbb{B}_{r(p_i)+\varepsilon}(o_{p_i})}$ is a diffeomorphism, and hence $\Lambda(\varepsilon)$ converges to a positive constant as $\varepsilon \downarrow 0$.

The next lemma tells us that $\widetilde{F}_\varepsilon^{(p_i)}$ is a local smooth approximation of $\widetilde{F}$ on $B_{r(p_i)}(p_i)$.

Lemma 4.6 ([26, Lemma 2.16]) *For each $\varepsilon \in (0, \text{inj}(M)/2)$ we have $\|\widetilde{F}_\varepsilon^{(p_i)}(q) - \widetilde{F}(q)\| \leq \varepsilon \cdot \Lambda(\varepsilon) \cdot \text{Lip}(F)$ for all $q \in B_{r(p_i)}(p_i)$ ($i \in \{1, 2, \ldots, k\}$) where $\|\cdot\|$ is the Euclidean norm of $\mathbb{R}^\ell$, and $\text{Lip}(F)$ denotes the Lipschitz constant of F, i.e.,*

$$\text{Lip}(F) := \sup\left\{ \frac{d_N(F(x), F(y))}{d_M(x, y)} \,\middle|\, x, y \in M, x \neq y \right\}.$$

Fix $p_i \in \{p_1, p_2, \ldots, p_k\}$ and $\varepsilon \in (0, \text{inj}(M)/2)$. We now construct Jacobi fields on $B_{\text{inj}(M)}(p_i)$ from geodesic variations with the initial point p_i. For each $q \in B_{r(p_i)}(p_i)$ we set $\mathbb{S}_q^{m-1} := \{v \in T_qM \mid \|v\| = 1\}$. Fix $v \in \mathbb{S}_q^{m-1}$. For $\delta > 0$ sufficiently small let $c_v : (-\delta, \delta) \to B_{r(p_i)}(p_i)$ be the minimal geodesic segment defined by $c_v(s) := \exp_q sv$. Since $c_v(s) \in B_{r(p_i)}(p_i)$ for all $s \in (-\delta, \delta)$, we observe, by the same argument as in Eq. (4.3), that

$$\exp_{p_i}^{-1} c_v(s) - y \in \mathbb{B}_{\text{inj}(M)}(o_{p_i}) \tag{4.7}$$

for all $s \in (-\delta, \delta)$ and $y \in \mathbb{B}_\varepsilon(o_{p_i})$. Since $\exp_{p_i}(\exp_{p_i}^{-1} c_v(s) - y) \in B_{\text{inj}(M)}(p_i)$ holds for all $s \in (-\delta, \delta)$ and $y \in \mathbb{B}_\varepsilon(o_{p_i})$ from Eq. (4.7), for each $y \in \mathbb{B}_\varepsilon(o_{p_i})$ we can define the smooth map $\varphi_y^{(v)} : [0, 1] \times (-\delta, \delta) \to B_{\text{inj}(M)}(p_i)$ by $\varphi_y^{(v)}(t, s) := \exp_{p_i} t[\exp_{p_i}^{-1} c_v(s) - y]$. The map $\varphi_y^{(v)}$ is a geodesic variation with the initial point p_i of the minimal geodesic segment

$$\varphi_y^{(v)}(t, 0) = \exp_{p_i} t(\exp_{p_i}^{-1} q - y) \tag{4.8}$$

emanating from $\varphi_y^{(v)}(0, 0) = p_i$ and ending at $\varphi_y^{(v)}(1, 0) = \exp_{p_i}(\exp_{p_i}^{-1} q - y)$. We get the Jacobi field

$$J_y^{(v)}(t) := \frac{\partial \varphi_y^{(v)}}{\partial s}(t, 0) \tag{4.9}$$

along $\varphi_y^{(v)}(t, 0)$, which satisfies the initial conditions $J_y^{(v)}(0) = o_{p_i}$ and $(DJ_y^{(v)}/dt)(0) = (d[\exp_{p_i}^{-1} c_v(s) - y]/ds)(0)$. For simplicity of notation we set

$$q_i(y) := \varphi_y^{(v)}(1, 0) = \exp_{p_i}(\exp_{p_i}^{-1} q - y). \tag{4.10}$$

Remark 4.7 The differential $(d\widetilde{F}_\varepsilon^{(p_i)})_q(v)$ of $\widetilde{F}_\varepsilon^{(p_i)}$ at q is given for any $q \in B_{r(p_i)}(p_i)$ and $v \in \mathbb{S}_q^{m-1}$ by $(d\widetilde{F}_\varepsilon^{(p_i)})_q(v) = \int_{y \in \mathbb{B}_\varepsilon(o_{p_i})} \rho_\varepsilon^{(i)}(y) d\widetilde{F}_{q_i(y)}(J_y^{(v)}(1))dy$. Indeed, fix $q \in B_{r(p_i)}(p_i)$ and $v \in \mathbb{S}_q^{m-1}$. Since $\widetilde{F} : M \to N \subset \mathbb{R}^\ell$, it follows from the definition of the differential of smooth maps (cf. [46]), Eqs. (4.4), and (4.10) that

$$
\begin{aligned}
(d\widetilde{F}_\varepsilon^{(p_i)})_q(v) &= \frac{d}{ds}\bigg|_0 (\widetilde{F}_\varepsilon^{(p_i)} \circ c_v)(s) = \frac{d}{ds}\bigg|_0 \int_{\mathbb{B}_\varepsilon(o_{p_i})} \rho_\varepsilon^{(i)}(y) \widetilde{F}(\varphi_y^{(v)}(1,s))dy \\
&= \int_{\mathbb{B}_\varepsilon(o_{p_i})} \rho_\varepsilon^{(i)}(y) d\widetilde{F}_{q_i(y)} \left(\frac{\partial \varphi}{\partial s}(1,0) \right) dy = \int_{\mathbb{B}_\varepsilon(o_{p_i})} \rho_\varepsilon^{(i)}(y) d\widetilde{F}_{q_i(y)}(J_y^{(v)}(1))dy.
\end{aligned}
$$

Lemma 4.8 *There is a constant $\omega(M) \in (0, \mathrm{inj}(M)/2)$ such that if $\varepsilon \in (0, \omega(M))$, then parallel transport $\tau_{q_i(y)}^q : T_qM \to T_{q_i(y)}M$ is defined as in Definition 2.3 for all $y \in \mathbb{B}_\varepsilon(o_{p_i})$ and $q \in B_{r(p_i)}(p_i)$ $(i \in \{1, 2, \ldots, k\})$.*

Proof. We see, by Remark 4.5, that for $\mathrm{inj}(M)/2$ there is a constant $\omega(M) \in (0, \mathrm{inj}(M)/2)$ such that $\varepsilon \cdot \Lambda(\varepsilon) < \mathrm{inj}(M)/2$ for all $\varepsilon \in (0, \omega(M))$. Fix $\varepsilon \in (0, \omega(M))$, and let $y \in \mathbb{B}_\varepsilon(o_{p_i})$ and $q \in B_{r(p_i)}(p_i)$. Since $\exp_{p_i}^{-1} q - y \in \mathbb{B}_{r(p_i)+\varepsilon}(o_{p_i})$ by Eq. (4.3), and since $\exp_{p_i}^{-1} q \in \mathbb{B}_{r(p_i)+\varepsilon}(o_{p_i})$, Eq. (4.10) gives $d_M(q, q_i(y)) \leq \Lambda(\varepsilon) \cdot \varepsilon < \mathrm{inj}(M)/2$, and hence there is a unique minimal geodesic segment emanating from q to $q_i(y)$. The map $\tau_{q_i(y)}^q : T_qM \to T_{q_i(y)}M$ is therefore defined as claimed. □

4.2 The global smooth approximation of Lipschitz maps

In this subsection we define the global smooth approximation $\widetilde{F}_\varepsilon$ of $\widetilde{F}$ using local smooth approximations and a partition of unity argument.

Since M is compact, there is a smooth partition of unity $\{\psi_i\}_{i=1}^k$ subordinate to $\{B_{r(p_i)}(p_i)\}_{i=1}^k$ (cf. [46]).

Definition 4.9 Fix $\varepsilon \in (0, \mathrm{inj}(M)/2)$. We define the smooth map $\widetilde{F}_\varepsilon : M \to \mathbb{R}^\ell$ by

$$
\widetilde{F}_\varepsilon(q) := \sum_{i=1}^k \psi_i(q) \widetilde{F}_\varepsilon^{(p_i)}(q) \qquad (q \in M) \tag{4.11}
$$

where each $\widetilde{F}_\varepsilon^{(p_i)}$ is the local smooth approximation of $\widetilde{F}$ on $B_{r(p_i)}(p_i)$.

The following lemma says that $\widetilde{F}_\varepsilon$ is the global smooth approximation of $\widetilde{F}$.

Lemma 4.10 ([26, Lemma 2.17]) *For each $\varepsilon \in (0, \mathrm{inj}(M)/2)$,*

$$
\|\widetilde{F}_\varepsilon(q) - \widetilde{F}(q)\| \leq \varepsilon \cdot \Lambda(\varepsilon) \cdot \mathrm{Lip}(F)
$$

holds for all $q \in M$ where $\|\cdot\|$ denotes the Euclidean norm of $\mathbb{R}^\ell$ and $\Lambda(\varepsilon)$ is the constant given by Eq. (4.6).

5 Proof of Main Theorem (Theorem 1.4)

5.1 Preliminaries

In this section it is shown that the smooth approximation of a Lipschitz map defined in subsection 4.2 is surjective near points that are nonsingular in the sense of Clarke. Throughout this subsection let M be a compact Riemannian manifold of dimension m, N a Riemannian manifold of dimension n with $m \geq n$, and $F : M \to N$ a Lipschitz map. Note here that we do not assume that N is connected or compact.

Via the Nash embedding theorem [33], we isometrically embed N into Euclidean space $\mathbb{R}^\ell$ with the canonical Riemannian metric $\langle \cdot, \cdot \rangle$ where $\ell \geq \max\{m, n+1\}$. Let $\widetilde{F} := F : M \to N \subset \mathbb{R}^\ell$, which is Lipschitz. Moreover we will use the same notation as in Section 4, e.g.,

- $\mathrm{inj}(M)$ is the injectivity radius of M,
- $\{B_{r(p_i)}(p_i)\}_{i=1}^k$, $\{B_{t(p_i)}(\widetilde{F}(p_i))\}_{i=1}^k$ are families of a finite number k of strongly convex balls $B_{r(p_i)}(p_i) \subset M$, $B_{t(p_i)}(\widetilde{F}(p_i)) \subset N$ satisfying (I)–(III) in Lemma 4.1,
- $\{\psi_i\}_{i=1}^k$ is the smooth partition of unity subordinate to $\{B_{r(p_i)}(p_i)\}_{i=1}^k$,
- $\widetilde{F}_\varepsilon^{(p_i)} : B_{r(p_i)}(p_i) \to \widetilde{F}_\varepsilon^{(p_i)}(B_{r(p_i)}(p_i)) \subset \mathbb{R}^\ell$ is the local smooth approximation of $\widetilde{F}$, defined by Eq. (4.2),
- $\widetilde{F}_\varepsilon : M \to \widetilde{F}_\varepsilon(M) \subset \mathbb{R}^\ell$ is the global smooth one of $\widetilde{F}$, done by Eq. (4.11), etc.

In what follows let $p \in M$ be nonsingular for $\widetilde{F}$, and let $\lambda(p)$ be the positive constant as in Proposition 3.6. Fix $q \in B_{\lambda(p)}(p)$. We can then choose $i \in \{1, 2, \ldots, k\}$ satisfying $q \in \operatorname{supp} \psi_i$. Note that $\operatorname{supp} \psi_i \subset B_{r(p_i)}(p_i)$.

Lemma 5.1 *Set $\varepsilon^{(i)}(p) := \min\{r(p_i), \omega(M), \lambda(p)/\operatorname{Lip}(\exp_{p_i}|_{\mathbb{B}_{2r(p_i)}(o_{p_i})})\}$ where $\omega(M) \in (0, \mathrm{inj}(M)/2)$ denotes the constant as in Lemma 4.8. Then for any $y \in \mathbb{B}_{\varepsilon^{(i)}(p)}(o_{p_i})$ we have $q_i(y) \in B_{2\lambda(p)}(p)$ where each $q_i(y)$ is the point defined by Eq. (4.10). In particular for any $y \in \mathbb{B}_{\varepsilon^{(i)}(p)}(o_{p_i})$ parallel transport $\tau_{\widetilde{F}(q_i(y))}^{\widetilde{F}(p)} : T_{\widetilde{F}(p)}N \to T_{\widetilde{F}(q_i(y))}N$ along a unique minimal geodesic of N emanating from $\widetilde{F}(p)$ to $\widetilde{F}(q_i(y))$ is defined in the sense of Definition 2.3.*

Note here that $\tau_{\widetilde{F}(q_i(y))}^{\widetilde{F}(p)}$ is not parallel translation along a line segment of $\mathbb{R}^\ell$ joining the two points.

Proof. Fix $y \in \mathbb{B}_{\varepsilon^{(i)}(p)}(o_{p_i})$. Since $\varepsilon^{(i)}(p) \leq r(p_i)$, the triangle inequality gives

$$\|\exp_{p_i}^{-1} q - y\| \leq \|\exp_{p_i}^{-1} q\| + \|y\| < r(p_i) + \varepsilon^{(i)}(p) \leq 2r(p_i),$$

and hence $\exp_{p_i}^{-1} q - y \in \mathbb{B}_{2r(p_i)}(o_{p_i})$. Since $q \in B_{r(p_i)}(p_i)$, it is clear that $\exp_{p_i}^{-1} q \in \mathbb{B}_{2r(p_i)}(o_{p_i})$. Note that $\exp_{p_i}|_{\mathbb{B}_{2r(p_i)}(o_{p_i})}$ is a diffeomorphism, as $2r(p_i) < \mathrm{inj}(M)$, see Lemma 4.1 (II). We then see, by the triangle inequality, that $d_M(p, q_i(y)) \leq d_M(p, q) + d_M(q, q_i(y)) < \lambda(p) + \varepsilon^{(i)}(p) \cdot \operatorname{Lip}(\exp_{p_i}|_{\mathbb{B}_{2r(p_i)}(o_{p_i})}) \leq 2\lambda(p)$. Hence we get $q_i(y) \in B_{2\lambda(p)}(p)$ as claimed. Moreover, since $p, q_i(y) \in B_{2\lambda(p)}(p)$, it follows from Proposition 3.6 (ii) that $\widetilde{F}(p), \widetilde{F}(q_i(y)) \in B_{t(p)}(\widetilde{F}(p)) \subset N$. Along the minimal geodesic of N emanating from $\widetilde{F}(p)$ to $\widetilde{F}(q_i(y))$ parallel transport $\tau_{\widetilde{F}(q_i(y))}^{\widetilde{F}(p)} : T_{\widetilde{F}(p)}N \to T_{\widetilde{F}(q_i(y))}N$ is defined as in Definition 2.3. $\square$

Remark 5.2 Since $\varepsilon^{(i)}(p) \leq \omega(M)$, and since $q \in B_{r(p_i)}(p_i)$, from Lemma 4.8 we have parallel transport $\tau_q^{q_i(y)} : T_{q_i(y)}M \to T_qM$, as in Definition 2.3, for all $y \in \mathbb{B}_{\varepsilon^{(i)}(p)}(o_{p_i})$. We use this in the next lemma.

From now on $\delta(p) > 0$ indicates the constant as in Proposition 3.6, and for each $x \in M$ let $\mathbb{S}_x^{m-1} := \{u \in T_xM \mid \|u\| = 1\}$ and $\mathbb{S}_{\widetilde{F}(x)}^{n-1} := \{v \in T_{\widetilde{F}(x)}N \mid \|v\| = 1\}$.

Lemma 5.3 (Key Lemma) *Fix $\varepsilon \in (0, \varepsilon^{(i)}(p))$. For any $y \in \mathbb{B}_\varepsilon(o_{p_i})$ and any $\tilde{u} \in \mathbb{S}_{\widetilde{F}(q)}^{n-1}$ there is a vector $V_{q_i(y)}^{(\tilde{u})} \in \mathbb{S}_{q_i(y)}^{m-1}$ such that*

$$\big\langle (d\widetilde{F}_\varepsilon^{(p_i)})_q(\tau_q^{q_i(y)}(V_{q_i(y)}^{(\tilde{u})})), \tilde{u}\big\rangle$$
$$\geq \delta(p) - \mathrm{Lip}(F)\big(\sup_{y\in\mathbb{B}_\varepsilon(o_{p_i})} \|J_y^{(\tau_q^{q_i(y)}(V_{q_i(y)}^{(\tilde{u})}))}(1) - V_{q_i(y)}^{(\tilde{u})}\| + \sup_{y\in\mathbb{B}_\varepsilon(o_{p_i})} \|\tilde{u} - (\tau_{\widetilde{F}(q_i(y))}^{\widetilde{F}(p)} \circ \tau_{\widetilde{F}(p)}^{\widetilde{F}(q)})(\tilde{u})\|\big).$$

Here $J_y^{(\tau_q^{q_i(y)}(V_{q_i(y)}^{(\tilde{u})}))}$ is the Jacobi field, defined by Eq. (4.9) for $v = \tau_q^{q_i(y)}(V_{q_i(y)}^{(\tilde{u})}) \in \mathbb{S}_q^{m-1}$, along the geodesic $\varphi_y^{(\tau_q^{q_i(y)}(V_{q_i(y)}^{(\tilde{u})}))}(t, 0)$ given by Eq. (4.8) joining p_i to $q_i(y)$.

Proof. By Lemma 5.1, $q_i(y) \in B_{2\lambda(p)}(p)$ holds for all $y \in \mathbb{B}_\varepsilon(o_{p_i})$. Fix $\tilde{u} \in \mathbb{S}_{\widetilde{F}(q)}^{n-1}$. It follows from Proposition 3.6 (iii) for $u = \tau_{\widetilde{F}(p)}^{\widetilde{F}(q)}(\tilde{u}) \in \mathbb{S}_{\widetilde{F}(p)}^{n-1}$ and $x = q_i(y)$ that *for almost all* $y \in \mathbb{B}_\varepsilon(o_{p_i})$ there is a vector $V_{q_i(y)}^{(\tilde{u})} := V_{q_i(y)}^{(\tau_{\widetilde{F}(p)}^{\widetilde{F}(q)}(\tilde{u}))} \in \mathbb{S}_{q_i(y)}^{m-1}$ such that

$$\big\langle V_{q_i(y)}^{(\tilde{u})}, (d\widetilde{F}_{q_i(y)})^*(\tau_{F(q_i(y))}^{\widetilde{F}(p)}(\tau_{\widetilde{F}(p)}^{\widetilde{F}(q)}(\tilde{u})))\big\rangle_M \geq \delta(p) \tag{5.1}$$

where $(d\widetilde{F}_{q_i(y)})^*$ is the adjoint of the differential $d\widetilde{F}_{q_i(y)} : T_{q_i(y)}M \to T_{\widetilde{F}(q_i(y))}N$, and $\langle\cdot,\cdot\rangle_M$ denotes Riemannian metric of M. Since N is isometrically embedded into $\mathbb{R}^\ell$, we see, by the Riesz representation theorem (cf. [38, Theorem 10.1]) and Eq. (5.1), that

$$\begin{aligned}\langle d\widetilde{F}_{q_i(y)}(V_{q_i(y)}^{(\tilde{u})}), (\tau_{\widetilde{F}(q_i(y))}^{\widetilde{F}(p)} \circ \tau_{\widetilde{F}(p)}^{\widetilde{F}(q)})(\tilde{u})\rangle &= \langle V_{q_i(y)}^{(\tilde{u})}, (d\widetilde{F}_{q_i(y)})^*(\tau_{F(q_i(y))}^{\widetilde{F}(p)}(\tau_{\widetilde{F}(p)}^{\widetilde{F}(q)}(\tilde{u})))\rangle_M \\ &\geq \delta(p)\end{aligned} \tag{5.2}$$

for almost all $y \in \mathbb{B}_\varepsilon(o_{p_i})$. For simplicity of notation we set $v := \tau_q^{q_i(y)}(V_{q_i(y)}^{(\tilde{u})}) \in \mathbb{S}_q^{m-1}$. We then see, by Eq. (4.1), Remark 4.7, and the Cauchy–Schwarz inequality, that

$$\begin{aligned}&\langle (d\widetilde{F}_\varepsilon^{(p_i)})_q(v), \tilde{u}\rangle \\ &\geq -\mathrm{Lip}(F) \sup_{y\in\mathbb{B}_\varepsilon(o_{p_i})} \|J_y^{(v)}(1) - V_{q_i(y)}^{(\tilde{u})}\| + \int_{\mathbb{B}_\varepsilon(o_{p_i})} \rho_\varepsilon^{(i)}(y)\langle d\widetilde{F}_{q_i(y)}(V_{q_i(y)}^{(\tilde{u})}), \tilde{u}\rangle dy.\end{aligned} \tag{5.3}$$

Moreover, we see, by Eqs. (4.1), (5.2), and the Cauchy–Schwarz inequality, that

$$\begin{aligned}&\int_{\mathbb{B}_\varepsilon(o_{p_i})} \rho_\varepsilon^{(i)}(y)\big\langle d\widetilde{F}_{q_i(y)}(V_{q_i(y)}^{(\tilde{u})}), \tilde{u}\big\rangle dy \\ &\geq -\int_{\mathbb{B}_\varepsilon(o_{p_i})} \rho_\varepsilon^{(i)}(y)\|d\widetilde{F}_{q_i(y)}(V_{q_i(y)}^{(\tilde{u})})\| \cdot \|\tilde{u} - (\tau_{\widetilde{F}(q_i(y))}^{\widetilde{F}(p)} \circ \tau_{\widetilde{F}(p)}^{\widetilde{F}(q)})(\tilde{u})\|dy + \delta(p) \\ &\geq -\mathrm{Lip}(F) \sup_{y\in\mathbb{B}_\varepsilon(o_{p_i})} \|\tilde{u} - (\tau_{\widetilde{F}(q_i(y))}^{\widetilde{F}(p)} \circ \tau_{\widetilde{F}(p)}^{\widetilde{F}(q)})(\tilde{u})\| + \delta(p).\end{aligned} \tag{5.4}$$

Substituting Eq. (5.4) into Eq. (5.3), we obtain the desired inequality. □

We can apply an argument almost identical to the proof of [26, Lemmas 2.23, 2.24] and Lemma 5.3 to show the following. We omit the proof.

Proposition 5.4 *There is a constant $\varepsilon_0(p) > 0$ satisfying the following: For any $\tilde{u} \in \mathbb{S}^{n-1}_{\widetilde{F}(q)}$ there is a vector $w^{(\tilde{u})} \in \mathbb{S}^{m-1}_q$ such that for any $\varepsilon \in (0, \varepsilon_0(p))$,*

$$\big\langle (d\widetilde{F}_\varepsilon)_q(w^{(\tilde{u})}), \tilde{u} \big\rangle \geq \frac{1}{3}\delta(p). \tag{5.5}$$

Corollary 5.5 *If $N = \mathbb{R}$, and $p \in M$ is nonsingular for the Lipschitz function $F : M \to \mathbb{R}$, then there are two constants $\lambda(p) > 0$ and $\varepsilon_0(p) > 0$ such that* $\operatorname{grad} \widetilde{F}_\varepsilon \neq 0$ *on $B_{\lambda(p)}(p)$ for all $\varepsilon \in (0, \varepsilon_0(p))$, hence $\widetilde{F}_\varepsilon$ has no critical points on $B_{\lambda(p)}(p)$.*

Proof. Let $\lambda(p)$ be the positive constant as in Proposition 3.6. Fix $x \in B_{\lambda(p)}(p)$ and $u \in \mathbb{S}^0_{F(x)}$. By Proposition 5.4 there is a constant $\varepsilon_0(p) > 0$ such that there is a vector $w^{(u)} \in \mathbb{S}^{m-1}_x$ satisfying $\langle (d\widetilde{F}_\varepsilon)_x(w^{(u)}), u\rangle \geq \delta(p)/3$ for all $\varepsilon \in (0, \varepsilon_0(p))$. Fix $\varepsilon \in (0, \varepsilon_0(p))$. Since $\tau^{F(x)}_{\widetilde{F}_\varepsilon(x)}(u) = (d/dt)|_{\widetilde{F}_\varepsilon(x)}$, we have $\delta(p)/3 \leq \langle (d\widetilde{F}_\varepsilon)_x(w^{(u)}), \tau^{F(x)}_{\widetilde{F}_\varepsilon(x)}(u)\rangle = \langle w^{(u)}(\widetilde{F}_\varepsilon)(d/dt)|_{\widetilde{F}_\varepsilon(x)}, (d/dt)|_{\widetilde{F}_\varepsilon(x)}\rangle = w^{(u)}(\widetilde{F}_\varepsilon) \cdot 1 = \langle (\operatorname{grad} \widetilde{F}_\varepsilon)_x, w^{(u)}\rangle_M \leq \|(\operatorname{grad} \widetilde{F}_\varepsilon)_x\|$, which shows the first assertion. The second assertion follows from the first one. □

5.2 Proof of Theorem 1.4

We follow assumptions and notation of Section 5.1. In addition we assume that N is connected and compact, and that the Lipschitz map $\widetilde{F} : M \to N \subset \mathbb{R}^\ell$ has no singular points on M.

Since N can be isometrically embedded into $\mathbb{R}^\ell$, it follows from the tubular neighborhood theorem (cf. [20], [27]) via the normal exponential map $\exp^\perp : TN^\perp \to \mathbb{R}^\ell$ that there is a constant $\mu_0 > 0$ such that $\exp^\perp$ is a diffeomorphism from an open neighborhood $\mathcal{U}_{\mu_0}(O(TN^\perp)) := \{X \in TN^\perp \,|\, \|X\| < \mu_0\}$ of the zero section $O(TN^\perp) = \{o_x \in T_xN^\perp \,|\, x \in N\}$ onto an open one $\mathcal{U}_{\mu_0}(N) := \exp^\perp[\mathcal{U}_{\mu_0}(O(TN^\perp))]$ of N in $\mathbb{R}^\ell$, which we will call the tubular neighborhood of N, where o_x is the origin of $T_xN^\perp$. Since $\exp^\perp|_{\mathcal{U}_{\mu_0}(O(TN^\perp))}$ is bijective, for any $y \in \mathcal{U}_{\mu_0}(N)$ there is a unique point $(z, v) \in \mathcal{U}_{\mu_0}(O(TN^\perp))$ such that $y = \exp^\perp(z, v)$. For such a pair $(y, (z, v))$ we have the smooth projection $\pi_N : \mathcal{U}_{\mu_0}(N) \to N$ given by $\pi_N(y) = \pi_N(\exp^\perp(z, v)) := z$. Note that the first variation formula yields $\|y - \pi_N(y)\| = \inf_{x\in N} \|y - x\|$ for all $y \in \mathcal{U}_{\mu_0}(N)$. For any $z \in N$ the definition of π_N gives $(T_zN)^\perp = \operatorname{Ker}(d\pi_N)_z$.

Since every $p \in M$ is nonsingular for $\widetilde{F}$, there are two positive constants $\delta(p)$ and $\varepsilon_0(p)$ obtained in Propositions 3.6 and 5.4, which satisfy Eq. (5.5). Set $\delta_0 := \min\{\delta(p) \,|\, p \in M\}$ and $\varepsilon_0 := \min\{\varepsilon_0(p) \,|\, p \in M\}$. Moreover Lemma 4.10 shows that for μ_0 above there is a constant $\varepsilon(\mu_0) \in (0, \operatorname{inj}(M)/2)$ such that if $\varepsilon \in (0, \varepsilon(\mu_0))$, then

$$\widetilde{F}_\varepsilon(M) \subset \mathcal{U}_{\mu_0}(N). \tag{5.6}$$

Set $\varepsilon_1 := \min\{\varepsilon_0, \varepsilon(\mu_0)\}$. It then follows from Proposition 5.4 for $q = p$ that for any $p \in M$ and any $\tilde{u} \in \mathbb{S}^{n-1}_{\widetilde{F}(p)}$ there is a vector $w^{(\tilde{u})} \in \mathbb{S}^{m-1}_p$ such that for any $\varepsilon \in (0, \varepsilon_1)$,

$$\big\langle (d\widetilde{F}_\varepsilon)_p(w^{(\tilde{u})}), \tilde{u} \big\rangle \geq \frac{1}{3}\delta_0. \tag{5.7}$$

For any $x, y \in \mathbb{R}^\ell$ let $P_y^x : T_x\mathbb{R}^\ell \to T_y\mathbb{R}^\ell$ be the parallel translation along the line segment in $\mathbb{R}^\ell$ joining x to y, and let $P_x^y := \big(P_y^x\big)^{-1}$.

Lemma 5.6 *Fix $p \in M$ and $\varepsilon \in (0, \varepsilon_1)$. For any $\widehat{u} \in P^{\widetilde{F}(p)}_{\widetilde{F}_\varepsilon(p)}\big(\mathbb{S}^{n-1}_{\widetilde{F}(p)}\big)$ there is a vector $\widehat{w} \in \mathbb{S}^{m-1}_p$ such that $\angle((d\widetilde{F}_\varepsilon)_p(\widehat{w}), \widehat{u}\,) < \pi/2$ holds where $\angle\big((d\widetilde{F}_\varepsilon)_p(\widehat{w}), \widehat{u}\,\big)$ is the angle between $(d\widetilde{F}_\varepsilon)_p(\widehat{w})$ and $\widehat{u}$ at the origin $o_{\widetilde{F}_\varepsilon(p)}$ of $T_{\widetilde{F}_\varepsilon(p)}\mathbb{R}^\ell$.*

Proof. Fix $\widehat{u} \in P^{\widetilde{F}(p)}_{\widetilde{F}_\varepsilon(p)}\big(\mathbb{S}^{n-1}_{\widetilde{F}(p)}\big)$. By Proposition 5.4, for $\tilde{u} := P^{\widetilde{F}_\varepsilon(p)}_{\widetilde{F}(p)}\big(\widehat{u}\big)(\in \mathbb{S}^{n-1}_{\widetilde{F}(p)})$ there is a vector $\widehat{w} := w^{(\tilde{u})} \in \mathbb{S}^{m-1}_p$ with Eq. (5.7). Since $\langle (d\widetilde{F}_\varepsilon)_p(\widehat{w}), \tilde{u}\,\rangle = \langle (d\widetilde{F}_\varepsilon)_p(\widehat{w}), P^{\widetilde{F}(p)}_{\widetilde{F}_\varepsilon(p)}(\tilde{u})\,\rangle$, we see $0 < \delta_0/3 \le \langle (d\widetilde{F}_\varepsilon)_p(\widehat{w}), \tilde{u}\,\rangle = \langle (d\widetilde{F}_\varepsilon)_p(\widehat{w}), \widehat{u}\,\rangle = \|(d\widetilde{F}_\varepsilon)_p(\widehat{w})\| \cos(\angle((d\widetilde{F}_\varepsilon)_p(\widehat{w}), \widehat{u}\,))$, and finally $\angle\big((d\widetilde{F}_\varepsilon)_p(\widehat{w}), \widehat{u}\,\big) < \pi/2$. ☐

Since every $p \in M$ is nonsingular for $\widetilde{F}$, $\mathrm{rank}(g) = n$ holds for all $g \in \partial\widetilde{F}(p)$, and hence for each $\varepsilon \in (0, \varepsilon_1)$ we see, by Lemma 4.10 and Eq. (5.7), that

$$\mathrm{Im}(d\widetilde{F}_\varepsilon)_p \cap P^{\widetilde{F}(p)}_{\widetilde{F}_\varepsilon(p)}(\mathrm{Ker}(d\pi_N)_{\widetilde{F}(p)}) = \{o_{\widetilde{F}_\varepsilon(p)}\} \qquad (p \in M). \tag{5.8}$$

Moreover, by virtue of Eq. (5.6), for each $\varepsilon \in (0, \varepsilon_1)$ we can define the smooth map $f_\varepsilon : M \to N$ by

$$f_\varepsilon(p) := (\pi_N \circ \widetilde{F}_\varepsilon)(p), \qquad (p \in M).$$

Lemma 5.7 *For any $\eta > 0$ there is a constant $\kappa(\eta) \in (0, \varepsilon_1)$ such that if $\varepsilon \in (0, \kappa(\eta))$, then $d_N(f_\varepsilon(p), \widetilde{F}(p)) < \eta$ and $\mathrm{Im}(d\widetilde{F}_\varepsilon)_p \cap \mathrm{Ker}(d\pi_N)_{\widetilde{F}_\varepsilon(p)} = \{o_{\widetilde{F}_\varepsilon(p)}\}$ hold for all $p \in M$.*

Proof. Fix $p \in M$. By Lemma 4.10, $\lim_{\varepsilon\downarrow 0} \|\widetilde{F}_\varepsilon(p) - \widetilde{F}(p)\| = 0$, and since $\pi_N(\widetilde{F}(p)) = \widetilde{F}(p)$, we have $\lim_{\varepsilon\downarrow 0} \|f_\varepsilon(p) - \widetilde{F}(p)\| = 0$. From this for any $\eta > 0$ there is a constant $\alpha_1(p, \eta) \in (0, \varepsilon_1)$ such that if $\varepsilon \in (0, \alpha_1(p, \eta))$, then

$$\|f_\varepsilon(p) - \widetilde{F}(p)\| < \frac{\eta}{\eta + 1}. \tag{5.9}$$

Fix $\eta > 0$. Since N is isometrically embedded into $\mathbb{R}^\ell$, $\lim_{\varepsilon\downarrow 0} \|f_\varepsilon(p) - \widetilde{F}(p)\| = 0$ also implies that there is a constant $\alpha_2(p, \eta) \in (0, \varepsilon_1)$ such that if $\varepsilon \in (0, \alpha_2(p, \eta))$, then

$$\left| \frac{d_N(f_\varepsilon(p) - \widetilde{F}(p))}{\|f_\varepsilon(p) - \widetilde{F}(p)\|} - 1 \right| < \eta. \tag{5.10}$$

Let $\beta_1(p, \eta) := \min\{\alpha_1(p, \eta), \alpha_2(p, \eta)\}$. Eqs. (5.9) and (5.10) show that if $\varepsilon \in (0, \beta_1(p, \eta))$, then

$$d_N(f_\varepsilon(p), \widetilde{F}(p)) < \eta. \tag{5.11}$$

For each $\varepsilon \in (0, \beta_1(p, \eta))$ let $\gamma_\varepsilon : [0, \mu_0) \to \mathcal{U}_{\mu_0}(N)$ be a unit speed minimal geodesic emanating perpendicularly from $f_\varepsilon(p)$ and passing through $\widetilde{F}_\varepsilon(p)$. Eq. (5.11) shows that by letting $\varepsilon \downarrow 0$, γ_ε converges to a unit speed minimal geodesic $\gamma_0 : [0, \mu_0) \to \mathcal{U}_{\mu_0}(N)$ emanating perpendicularly from $\widetilde{F}(p)$. Since $\lim_{\varepsilon\downarrow 0} \|\widetilde{F}_\varepsilon(p) - \widetilde{F}(p)\| = 0$, we see

$\lim_{\varepsilon\downarrow 0} \mathrm{Ker}(d\pi_N)_{\widetilde{F}_\varepsilon(p)} = \mathrm{Ker}(d\pi_N)_{\widetilde{F}(p)}$. Since $\mathrm{Im}(d\widetilde{F}_\varepsilon)_p \cap P^{\widetilde{F}(p)}_{\widetilde{F}_\varepsilon(p)}(\mathrm{Ker}(d\pi_N)_{\widetilde{F}(p)}) = \{o_{\widetilde{F}_\varepsilon(p)}\}$ by Eq. (5.8), we see that for η there is a constant $\beta_2(p,\eta) \in (0, \beta_1(p,\eta)$ such that if $\varepsilon \in (0, \beta_2(p,\eta))$, then

$$\mathrm{Im}(d\widetilde{F}_\varepsilon)_p \cap \mathrm{Ker}(d\pi_N)_{\widetilde{F}_\varepsilon(p)} = \{o_{\widetilde{F}_\varepsilon(p)}\}. \tag{5.12}$$

By setting $\kappa(\eta) := \min\{\beta_2(p,\eta) \,|\, p \in M\}$, Eqs. (5.11) and (5.12) complete the proof. □

Fix $\varepsilon \in (0, \kappa(\eta))$. Lemma 5.7 shows $\mathrm{rank}(d\pi_N|_{\mathrm{Im}(d\widetilde{F}_\varepsilon)_p}) = n$ for all $p \in M$, and hence $\mathrm{rank}((df_\varepsilon)_p) = n$ for all $p \in M$, which proves that f_ε is a smooth submersion from M to N. Note that f_ε is an open map, because f_ε is locally equivalent to the canonical projection on some coordinate neighborhood of each point of M, see [46]. Since f_ε is continuous, and since M is compact, $f_\varepsilon(M)$ is compact in N. $f_\varepsilon(M)$ is thus closed in N, for N is Hausdorff. Since M is open in M, $f_\varepsilon(M)$ is open in N. Connectedness of N shows that f_ε is surjective. Let K be any compact set in N. By virtue of the compactness of N, K is closed in N. From the continuity of f_ε on M, $f_\varepsilon^{-1}(K)$ is closed in M. Since M is compact, $f_\varepsilon^{-1}(K)$ is also, and hence f_ε is proper. Since f_ε is a proper and surjective submersion between compact, smooth manifolds, Ehresmann's lemma [11] shows that f_ε is a locally trivial fibration, i.e., an Ehresmann fibration. □

6 Proof of Reeb's sphere theorem for Lipschitz functions (Theorem 1.7)

Throughout this section let M be a closed Riemannian manifold of dimension m, and we assume that M admits a Lipschitz function $F : M \to \mathbb{R}$ with exactly two singular points in the sense of Clarke, denoted by $z_1, z_2 \in M$.

Since $z_1, z_2 \in M$ are singular for F, we see, by Lemma 2.16, that $o_{z_i} \in \circledast_F(z_i)$ $(i = 1, 2)$ where $\circledast_F(z_i)$ indicates the generalized gradient of F at z_i (see Definition 2.13). From the maximum and minimum values theorem we can therefore assume, without loss of generality, that $F(z_1) = \min_{x \in M} F(x)$ and $F(z_2) = \max_{y \in M} F(y)$. For simplicity of notation let $a_i := F(z_i)$ for $i = 1, 2$. Note that $a_1 < a_2$.

Lemma 6.1 *For any $r > 0$ with $B_r(z_1) \cap B_r(z_2) = \emptyset$ there is a constant $b_i(r) \in (a_1, a_2)$ such that $F^{-1}(b_i(r)) \subset B_r(z_i)$ for each $i = 1, 2$.*

Proof. We prove this lemma only in the case of $i = 1$. Suppose not. There is then $r_0 > 0$ such that for any $\lambda \in (a_1, a_2)$, $F^{-1}(\lambda) \not\subseteq B_{r_0}(z_1)$ holds. For each $n \in \mathbb{N}$ there is $x_n \in F^{-1}(a_1 + (a_2 - a_1)/2n)$ such that $x_n \notin B_{r_0}(z_1)$, and hence we get a sequence $\{x_n\}_{n\in\mathbb{N}}$ of such points x_n. Since M is compact, $\{x_n\}_{n\in\mathbb{N}}$ has a convergent subsequence $\{x_{n_j}\}_{j\in\mathbb{N}}$. Let $\bar{x} := \lim_{j\to\infty} x_{n_j}$. Since F is continuous on M, we see that $F(\bar{x}) = \lim_{j\to\infty} F(x_{n_j}) = \lim_{j\to\infty}\{a_1 + (a_2 - a_1)/2n_j\} = a_1$. Now $\bar{x} \neq z_1$, and $F(\bar{x}) = a_1$, hence $\bar{x}$ is a critical point of F, which is a contradiction. □

Fix $r > 0$ with $B_r(z_1) \cap B_r(z_2) = \emptyset$. For $\varepsilon \in (0, \mathrm{inj}(M)/2)$ let $\widetilde{F}_\varepsilon : M \to \mathbb{R}$ be the global smooth approximation of F defined by Eq. (4.11).

Lemma 6.2 *There is an open set V of M and a constant $\varepsilon_0 \in (0, \mathrm{inj}(M)/2)$ such that if $\varepsilon \in (0, \varepsilon_0)$, then $F^{-1}([b_1(r), b_2(r)]) \subset V$, and $\widetilde{F}_\varepsilon$ has no critical points on V.*

Proof. For simplicity of notation let $M' := F^{-1}([b_1(r), b_2(r)])$. We first show $M' \subset V$. Since z_1 and z_2 are the only two critical points of F, and since $a_1 < b_1(r) < b_2(r) < a_2$, F has no critical points on M'. It follows from Lemma 2.16 and Corollary 5.5 that for each $p \in M'$ there are two constants $\lambda(p) > 0$ and $\bar{\varepsilon}(p) > 0$ such that for each $\varepsilon \in (0, \bar{\varepsilon}(p))$, $\operatorname{grad} \widetilde{F}_\varepsilon \neq 0$ on $B_{\lambda(p)}(p)$. Since M' is compact, there is a finite set $\{p_1, p_2, \ldots, p_k\} \subset M'$ such that $M' \subset \bigcup_{i=1}^k B_{\lambda(p_i)}(p_i)$. Since $\bigcup_{i=1}^k B_{\lambda(p_i)}(p_i)$ is open in M, setting $V := \bigcup_{i=1}^k B_{\lambda(p_i)}(p_i)$, we get the first assertion.

We next show the second assertion. Set $\varepsilon_0 := \min\{\bar{\varepsilon}(p_1), \bar{\varepsilon}(p_2), \ldots, \bar{\varepsilon}(p_k)\}$. Since $F^{-1}([b_1(r), b_2(r)]) \subset V = \bigcup_{i=1}^k B_{\lambda(p_i)}(p_i)$ and $\operatorname{grad} \widetilde{F}_\varepsilon \neq 0$ on V, $\widetilde{F}_\varepsilon$ has no critical points on V for all $\varepsilon \in (0, \varepsilon_0)$. □

Lemma 6.3 *There is a constant $\varepsilon_1 \in (0, \varepsilon_0]$ such that if $\varepsilon \in (0, \varepsilon_1)$, then for any $c \in (b_1(r), b_2(r))$, $\widetilde{F}_\varepsilon^{-1}([b_1(r), b_2(r)])$ is diffeomorphic to $\widetilde{F}_\varepsilon^{-1}(c) \times [b_1(r), b_2(r)]$.*

Proof. Fix $i \in \{1, 2\}$. Since $B_r(z_i) \cap V$ is an open neighborhood of $F^{-1}(b_i(r))$, we see, by Lemma 4.10, that there is a constant $\hat{\varepsilon}_i \in (0, \operatorname{inj}(M)/2)$ such that if $\varepsilon \in (0, \hat{\varepsilon}_i)$, then $\widetilde{F}_\varepsilon^{-1}(b_i(r)) \subset B_r(z_i) \cap V$. Let $\varepsilon_1 := \min\{\varepsilon_0, \hat{\varepsilon}_1, \hat{\varepsilon}_2\}$, and fix $\varepsilon \in (0, \varepsilon_1)$. Since $\widetilde{F}_\varepsilon^{-1}(b_i(r)) \subset B_r(z_i) \cap V$, $\widetilde{F}_\varepsilon^{-1}([b_1(r), b_2(r)]) \subset V$ holds, and hence we see, by Lemma 6.2, that $\widetilde{F}_\varepsilon$ has no critical points on $\widetilde{F}_\varepsilon^{-1}([b_1(r), b_2(r)])$. From this, for each $t \in [b_1(r), b_2(r)]$, $\widetilde{F}_\varepsilon^{-1}(t)$ is an $(m-1)$-dimensional compact regular submanifold of M. $\widetilde{F}_\varepsilon^{-1}([b_1(r), b_2(r)])$ is therefore diffeomorphic to $\widetilde{F}_\varepsilon^{-1}(b_1) \times [b_1(r), b_2(r)]$ by a well-known theorem in Morse theory ([30, Theorem 2.31], or [31, Theorem 3.1]). Fix $c \in (b_1(r), b_2(r))$. Since $\widetilde{F}_\varepsilon^{-1}(s)$ and $\widetilde{F}_\varepsilon^{-1}(t)$ are diffeomorphic for all $s, t \in [b_1(r), b_2(r)]$, $\widetilde{F}_\varepsilon^{-1}(b_1)$ is diffeomorphic to $\widetilde{F}_\varepsilon^{-1}(c)$, which yields our assertion. □

Now we give the proof of Theorem 1.7. From Lemma 6.1 we observe $\lim_{r \downarrow 0} b_i(r) = a_i$ $(i = 1, 2)$. Fix $c \in (b_1(r), b_2(r)) \subseteq (a_1, a_2)$. Let $r \downarrow 0$. Lemma 6.3 then shows that M is homeomorphic to the suspension, denoted by Σ, of the compact regular submanifold $\widetilde{F}_\varepsilon^{-1}(c)$ of M for a sufficiently small $\varepsilon > 0$. It follows easily from a result of Brown [3] (see also [10, Introduction]) that Σ is homeomorphic to the m-sphere for a sufficiently small $\varepsilon > 0$, and M is also. □

References

[1] U. Abresch and D. Gromoll, *On complete manifolds with nonnegative Ricci curvature*, J. Amer. Math. Soc. **3** (1990), No. 2, 355–374.

[2] M. Berger, *Les variétés Riemanniennes (1/4)-pincées*, Ann. Scuola Norm. Sup. Pisa (3) **14** (1960), 161–170.

[3] M. Brown, *The monotone union of open n-cells is an open n-cell*, Proc. Amer. Math. Soc. **12** (1961), 812–814.

[4] J. Cerf, Sur les difféomorphismes de la sphère de dimension trois ($\Gamma_4 = 0$), Lecture Notes in Math., **53**. Springer-Verlag, 1968.

[5] J. Cheeger, *Critical points of distance functions and applications to geometry*, in Geometric topology: recent developments (Montecatini Terme, 1990), 1–38, Lecture Notes in Math., **1504**. Springer, Berlin, 1991.

[6] J. Cheeger and D.G. Ebin, "Comparison Theorems in Riemannian Geometry", North-Holland Mathematical Library, **9**. North-Holland Publishing Co., Amsterdam-Oxford; American Elsevier Publishing Co., Inc., New York, 1975.

[7] F.H. Clarke, *Generalized gradients and applications*, Trans. Amer. Math. Soc. **205** (1975), 247–262.

[8] F.H. Clarke, *On the inverse function theorem*, Pacific J. Math. **64** (1976), 97–102.

[9] M.P. do Carmo, "Riemannian Geometry", Birkhäuser, Boston, MA, 1992. Translated from the second Portuguese edition by F. Flaherty.

[10] R.D. Edwards, *Suspensions of homology spheres*, `arXiv:0610573`

[11] C. Ehresmann, *Les connexions infinitésimales dans un espace fibré différentiable*, Colloque de topologie (espaces fibrés), Bruxelles, 1950, pp. 29–55. Georges Thone, Liége; Masson et Cie., Paris, 1951.

[12] T. Fujioka, *A fibration theorem for collapsing sequences of Alexandrov spaces*, to appear in J. Topol. Anal.; `arXiv:1905.05484`

[13] P. Goldstein and P. Hajłasz, *Topological obstructions to continuity of Orlicz–Sobolev mappings of finite distortion*, Ann. Mat. Pura Appl. (4) **198** (2019), no. 1, 243–262.

[14] R.E. Greene and H. Wu, C^∞ *approximations of convex, subharmonic, and plurisubharmonic functions*, Ann. Sci. École Norm. Sup. (4) **12** (1979), no. 1, 47–84.

[15] M. Gromov, *Curvature, diameter and Betti numbers*, Comment. Math. Helv. **56** (1981), 179–195.

[16] M. Gromov, "Partial Differential Relations", Ergeb. Math. Grenzgeb. (3), **9**. Springer–Verlag, Berlin, 1986.

[17] K. Grove, *Critical point theory for distance functions*, in Differential geometry: Riemannian geometry (Los Angeles, CA, 1990), 357–385, Proc. Sympos. Pure Math., **54**, Part 3. Amer. Math. Soc., Providence, RI, 1993.

[18] K. Grove and P. Petersen, *Bounding homotopy types by geometry*, Ann. of Math. (2) **128** (1988), 195–206.

[19] K. Grove and K. Shiohama, *A generalized sphere theorem*, Ann. of Math. (2) **106** (1977), 201–211.

[20] M.W. Hirsch, "Differential Topology", Grad. Texts in Math., **33**. Springer–Verlag, New York, 1994. Corrected reprint of the 1976 original.

[21] J. Itoh and M. Tanaka, *The Lipschitz continuity of the distance function to the cut locus*, Trans. Amer. Math. Soc. **353** (2001), no. 1, 21–40.

[22] H. Karcher, *Riemannian center of mass and mollifier smoothing*, Comm. Pure Appl. Math. **30** (1977), 509–541.

[23] M. Kervaire and J. Milnor, *Groups of homotopy spheres. I*, Ann. of Math. (2) **77** (1963), 504–537.

[24] K. Kondo and S. Ohta, *Topology of complete manifolds with radial curvature bounded from below*, Geom. Funct. Anal. **17** (4) (2007) 1237–1247.

[25] K. Kondo and M. Tanaka, *Total curvatures of model surfaces control topology of complete open manifolds with radial curvature bounded below, II*, Trans. Amer. Math. Soc. **362** (2010), 6293–6324.

[26] K. Kondo and M. Tanaka, *Approximations of Lipschitz maps via immersions and differentiable exotic sphere theorems*, Nonlinear Anal. **155** (2017), 219–249.

[27] J.M. Lee, "Introduction to Smooth Manifolds" (Second edition), Grad. Texts in Math., **218**. Springer, New York, 2013.

[28] G. Leoni, "A First Course in Sobolev Spaces" (Second edition), Grad. Stud. Math., **181**. Amer. Math. Soc., Providence, RI, 2017.

[29] S. Li, *Cartan–Whitney presentation, non-smooth analysis and smoothability of manifolds: On a theorem of Kondo–Tanaka*, `arXiv:1904.00515`

[30] Y. Matsumoto, "An Introduction to Morse Theory", Translations of Mathematical Monographs, **208**. Amer. Math. Soc., Providence, RI, 2002. Translated from the 1997 Japanese original by Kiki Hudson and Masahico Saito.

[31] J. Milnor, "Morse theory", Annals of Mathematics Studies, **51**. Princeton University Press, Princeton, N.J., 1963. Notes by M. Spivak and R. Wells.

[32] J. Milnor, "Lectures on the h-cobordism Theorem", Princeton University Press, Princeton, N.J., 1965. Notes by L. Siebenmann and J. Sondow.

[33] J. Nash, *The imbedding problem for Riemannian manifolds*, Ann. of Math. **63**, No. 1 (1956), 20–63.

[34] R. Palais, *Extending diffeomorphisms*, Proc. Amer. Math. Soc. **11** (1960), 274–277.

[35] H. Rademacher, *Über partielle und totale differenzierbarkeit von Funktionen mehrerer Variabeln und über die Transformation der Doppelintegrale*, Math. Ann. **79** (1919), 340–359.

[36] G. Reeb, *Sur certaines propriétés topologiques des variétés feuilletées*, Publ. Inst. Math. Univ. Strasbourg 11, pp. 5–89, 155–156. Actualités Sci. Ind., no. 1183, Hermann & Cie., Paris, 1952.

[37] R.T. Rockafellar, "Convex Analysis", Princeton Landmarks in Mathematics, Princeton University Press, Princeton, NJ, 1997. Reprint of the 1970 original.

[38] S. Roman, "Advanced Linear Algebra" (Third edition), Grad. Texts in Math. **135**, Springer-Verlag New York, 2008.

[39] T. Sakai, "Riemannian Geometry", Translations of Mathematical Monographs **149**, Amer. Math. Soc., Providence, RI, 1996. Translated from the 1992 Japanese original by the author.

[40] R. Schneider, "Convex Bodies: The Brunn–Minkowski Theory", Encyclopedia of Mathematics and its Applications, **44**. Cambridge University Press, Cambridge, 1993.

[41] Y. Shikata, *On a distance function on the set of differentiable structures*, Osaka J. Math. **3** (1966), 65–79.

[42] S. Smale, *Diffeomorphisms of the* 2*-sphere*, Proc. Amer. Math. Soc. **10** (1959) 621–626.

[43] S. Smale, *Generalized Poincaré's conjecture in dimensions greater than four*, Ann. of Math. **74** (1961), 391–406.

[44] S. Smale, *On the structure of manifolds*, Amer. J. Math. **84** (1962), 387-399.

[45] V.A. Toponogov, *Riemann spaces with curvature bounded below* (Russian), Uspehi Mat. Nauk **14** (1959) no. 1 (85), 87–130.

[46] L.W. Tu, “An Introduction to Manifolds” (Second edition), Universitext, Springer, New York, 2011.

[47] A.D. Weinstein, *The cut locus and conjugate locus of a riemannian manifold*, Ann. of Math. (2) **87** (1968), 29–41.

[48] R.L. Wheeden and A. Zygmund, “Measure and Integral”, Marcel Decker, New York, 1977.

[49] J.H.C. Whitehead, *On the covering of a complete space by the geodesics through a point*, Ann. of Math. (2) **36** (1935), 679–704.

[50] T. Yamaguchi, *Collapsing and pinching under a lower curvature bound*, Ann. of Math. (2) **133** (1991), no. 2, 317–357.

[51] T. Yamaguchi, *A convergence theorem in the geometry of Alexandrov spaces*, Actes de la Table Ronde de Géométrie Différentielle (Luminy, 1992), 601–642, Sémin. Congr. **1**, Soc. Math. France, Paris, 1996.

[52] K. Yosida, “Functional Analysis”, Classics in Mathematics, Springer–Verlag, Berlin, 1995. Reprint of the sixth (1980) edition.

K. Kondo
Department of Mathematics, Faculty of Science, Okayama University
Okayama City, Okayama Pref. 700-8530, Japan
e-mail: `keikondo@math.okayama-u.ac.jp`