What Happens to a Manifold Under a Bi-Lipschitz Map?

Armin Eftekhari[*] and Michael B. Wakin[†]

October 3, 2018

Abstract

We study geometric and topological properties of the image of a smooth submanifold of $\mathbb{R}^n$ under a bi-Lipschitz map to $\mathbb{R}^m$. In particular, we characterize how the dimension, diameter, volume, and reach of the embedded manifold relate to the original. Our main result establishes a lower bound on the reach of the embedded manifold in the case where $m \leq n$ and the bi-Lipschitz map is linear. We discuss implications of this work in signal processing and machine learning, where bi-Lipschitz maps on low-dimensional manifolds have been constructed using randomized linear operators.

Keywords. Manifolds, reach, bi-Lipschitz maps, compressive sensing, random projections.

AMS Subject Classification. 28A75, 53A07, 57R40, 68P30, 94A12.

1 Introduction

Let $\mathbb{M} \subset \mathbb{R}^n$ be a k-dimensional smooth *submanifold* of $\mathbb{R}^n$ with finite diameter and volume, and with nonzero *reach*.[1] Consider also $\Phi_{l,u} : \mathbb{R}^n \to \mathbb{R}^m$ that acts as a *bi-Lipschitz* map on $\mathbb{M}$ in the sense that

$$l \cdot \|x - y\|_2 \leq \|\Phi_{l,u}(x) - \Phi_{l,u}(y)\|_2 \leq u \cdot \|x - y\|_2, \qquad \forall x, y \in \mathbb{M}, \tag{1}$$

for some $0 < l \leq 1 \leq u < \infty$. In particular, when $l, u \approx 1$, then $\Phi_{l,u}(\cdot)$ is a *near-isometry* on $\mathbb{M}$, in that it barely distorts the pairwise Euclidean distances between points on $\mathbb{M}$.[2] Such maps naturally arise in a variety of applications in data sciences, often involving dimensionality reduction [2, 3, 6, 8, 9, 11, 12]; we will expand on this remark in Section 5. Let us point out that we only assume the map $\Phi_{l,u}(\cdot)$ to be bi-Lipschitz on $\mathbb{M}$, not on the entire domain $\mathbb{R}^n$. This is a subtle but crucial detail which will be significant, for example, in applications where $\Phi_{l,u}$ is linear and where $m < n$. In such applications $\Phi_{l,u}(\cdot)$ will have a nullspace and map many vectors to 0, but for (1) to be satisfied with $l > 0$ no vector of the form $x - y$ with $x, y \in \mathbb{M}$, $x \neq y$ can be in the nullspace of $\Phi_{l,u}(\cdot)$.

The setup above naturally raises the following question: What happens to $\mathbb{M}$ under the bi-Lipschitz map $\Phi_{l,u}(\cdot)$? It is in fact trivially verified that $\Phi_{l,u}(\mathbb{M}) \subset \mathbb{R}^m$ is itself a smooth submanifold

---

[*]Institute for Computational Engineering and Sciences (ICES), University of Texas at Austin. Email: armin.eftekhari@utexas.edu

[†]Department of Electrical Engineering and Computer Science, Colorado School of Mines. Email: mwakin@mines.edu

[1]Reach measures the regularity of a manifold. This, as well as more specific requirements on $\mathbb{M}$, are described in Section 2.

[2]The deviation $\max[u - 1, 1 - l]$ (if positive) is commonly referred to as the *isometry constant* of the map $\Phi_{l,u}(\cdot)$ on $\mathbb{M}$.

of $\mathbb{R}^m$, where $\Phi_{l,u}(\mathbb{M})$ is the image of $\mathbb{M}$ under $\Phi_{l,u}(\cdot)$. This allows us to define our inquiry more precisely:

- **Question:** How do key characteristics of $\Phi_{l,u}(\mathbb{M})$—namely dimension, diameter, volume, and reach—relate to those of $\mathbb{M}$?

When $l, u \approx 1$, we will verify that

$$\dim\left(\Phi_{l,u}(\mathbb{M})\right) = \dim\left(\mathbb{M}\right),$$
$$\operatorname{diam}\left(\Phi_{l,u}(\mathbb{M})\right) \approx \operatorname{diam}\left(\mathbb{M}\right),$$
$$\operatorname{vol}_k\left(\Phi_{l,u}(\mathbb{M})\right) \approx \operatorname{vol}_k\left(\mathbb{M}\right),$$

as formalized in Proposition 3. However, for reach—an important attribute of the manifold—the story is somewhat different as additional restrictions must be imposed on $\Phi_{l,u}(\cdot)$ for reach to be approximately preserved. We briefly outline this peculiar case next and defer the details to Section 3.

Reach, as a measure of regularity of a submanifold of Euclidean space, can be traced back to the pioneering works in *geometric measure theory* [10]. More recently, the concept of reach has proved indispensable in the role that manifold models play in signal processing and machine learning [1,7,12,15,19,22,24]. As elaborated later in Proposition 6, reach is actually preserved under an *exact isometry*, i.e., $\operatorname{rch}(\Phi_{1,1}(\mathbb{M})) = \operatorname{rch}(\mathbb{M})$. However, there is generally no relation between $\operatorname{rch}(\Phi_{l,u}(\mathbb{M}))$ and $\operatorname{rch}(\mathbb{M})$ when $l < 1 < u$, and we provide a concrete example in Proposition 7 to demonstrate this. This example involves a nonlinear choice of $\Phi_{l,u}(\cdot)$. Interestingly, however, if we limit ourselves to the class of *linear* maps, reach is nearly preserved. Specifically, we can lower bound $\operatorname{rch}\left(\Phi_{l,u}(\mathbb{M})\right)$ in terms of $\operatorname{rch}(\mathbb{M})$ when $\Phi_{l,u}(\cdot)$ is a linear near-isometry; we formalize this claim in Theorem 8.

Before attending to the details, however, we present the necessary background in Section 2 to make this paper self-contained. Then, Section 3 makes precise the claims we outlined above, with further discussion and context for these results appearing in Section 4. In Section 5, revisiting the applications of nearly-isometric maps in data sciences, we show how the (seemingly abstract) Question above naturally arises in modern signal processing and machine learning, and we discuss the implications of our results in these contexts. We defer all proofs to the appendices.

2   Necessary Concepts

Let $\mathbb{M} \subset \mathbb{R}^n$ be a smooth, bounded, and boundary-less k-dimensional submanifold of $\mathbb{R}^n$. More formally, $\mathbb{M}$ satisfies the following. Any point $x \in \mathbb{M}$ belongs to some *neighborhood* $\mathbb{U}_x \subseteq \mathbb{M}$ that is *diffeomorphic*[3] to $\mathbb{R}^k$, and the transition between adjacent neighborhoods is smooth [16,17].[4]

To every point $x \in \mathbb{M}$, we assign a *tangent subspace* $\mathbb{T}_x\mathbb{M} \subset \mathbb{R}^n$ comprised of the directions of all smooth curves on $\mathbb{M}$ that pass through x:

$$\mathbb{T}_x\mathbb{M} :=$$
$$\operatorname{span}\left[\left\{\frac{d\gamma}{dt}(0) \,|\, \gamma : [-1,1] \to \mathbb{M} \text{ is a smooth curve and } \gamma(0) = x \in \mathbb{M}\right\}\right].$$

---

[3] A diffeomorphism is a smooth and bijective map with a smooth inverse.

[4] By smooth transition we mean the following: If $\phi_x : \mathbb{U}_x \to \mathbb{R}^k$ and $\phi_{x'} : \mathbb{U}_{x'} \to \mathbb{R}^k$ are, respectively, the diffeomorphisms corresponding to the neighborhoods $\mathbb{U}_x$ and $\mathbb{U}_{x'}$ on the manifold $\mathbb{M}$ and $\mathbb{U}_x \cap \mathbb{U}_{x'} \neq \emptyset$, then the composition $(\phi_x \circ \phi_{x'}^{-1})(\cdot)$ must be smooth (wherever defined).
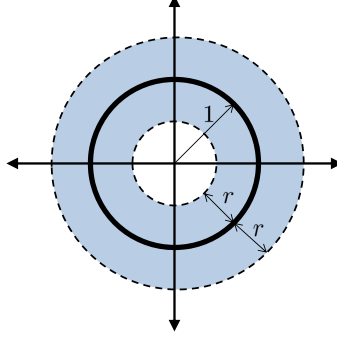
Figure 1: *The reach of the unit circle (in bold) is equal to 1, as the open normal bundle of radius r (shaded) is embedded in $\mathbb{R}^2$ for any $r < 1$. Moreover, any point in $\mathbb{R}^2$ whose distance from the unit circle is less than 1 has a unique nearest point on the unit circle.*

At any $x \in \mathbb{M}$, we note that $\mathbb{T}_x\mathbb{M}$ is a k-dimensional linear subspace of $\mathbb{R}^n$. The *tangent bundle* of $\mathbb{M}$ is the collection of all tangent subspaces:

$$\mathbb{T}\mathbb{M} := \bigcup_{x \in \mathbb{M}} \{x\} \times \mathbb{T}_x\mathbb{M}.$$

The *normal subspace* $\mathbb{N}_x\mathbb{M}$ is the orthogonal complement of $\mathbb{T}_x\mathbb{M}$ with respect to $\mathbb{R}^n$ so that

$$\mathbb{R}^n = \mathbb{T}_x\mathbb{M} \oplus \mathbb{N}_x\mathbb{M},$$

(with $\oplus$ standing for direct sum). Thus, at any $x \in \mathbb{M}$, $\mathbb{N}_x\mathbb{M}$ is an $(n-k)$-dimensional linear subspace of $\mathbb{R}^n$. The *normal bundle* of $\mathbb{M}$ is the collection of all normal subspaces:

$$\mathbb{N}\mathbb{M} := \bigcup_{x \in \mathbb{M}} \{x\} \times \mathbb{N}_x\mathbb{M}.$$

For $r > 0$, we also let $\mathbb{N}^r\mathbb{M} \subset \mathbb{N}\mathbb{M}$ denote the *open normal bundle* of $\mathbb{M}$ of radius r comprised of all normal vectors of length less than r. For example, when $\mathbb{M}$ is the unit circle in $\mathbb{R}^2$ and $r \in (0, 1)$, $\mathbb{N}^r\mathbb{M}$ may be identified with a disc of width $2r$ (around the circle); see Figure 1.

The reach of $\mathbb{M}$ [10], denoted by $rch(\mathbb{M})$ throughout, is a geometric attribute that captures valuable local and global properties of a manifold (that we will highlight shortly).

Definition 1. (Reach) *For a smooth submanifold $\mathbb{M} \subset \mathbb{R}^n$, the reach of $\mathbb{M}$ (denoted by $rch(\mathbb{M})$) is the largest number having the following property: the open normal bundle of $\mathbb{M}$ of radius r is embedded in $\mathbb{R}^n$ for all $r < rch(\mathbb{M})$.*

For example, the reach of a circle of radius ρ is simply ρ itself. This fact is generalized by the following result, which may be of its own independent interest.

Proposition 2. (Reach of an ellipsoid) *Let $\mathbb{M}$ be an $(n-1)$-dimensional ellipsoid in $\mathbb{R}^n$ with principal axes $r_1 \geq r_2 \geq \cdots \geq r_n > 0$. Then*

$$rch(\mathbb{M}) = \frac{r_n^2}{r_1}.$$

By definition, any point in $\mathbb{R}^n$ within the distance $rch(\mathbb{M})$ of $\mathbb{M}$ has a unique nearest point on $\mathbb{M}$. The role of reach can be summarized in two key properties (see the Toolbox in [9] for more

details). The first property concerns local regularity of $\mathbb{M}$: the curvature of any unit-speed *geodesic curve*[5] on $\mathbb{M}$ is bounded by $(\mathrm{rch}(\mathbb{M}))^{-1}$. The second property concerns global regularity: at long *geodesic distances*,[6] the reach controls how close the manifold may curve back upon itself. For example, supposing $x, y \in \mathbb{M}$ with geodesic distance $d_{\mathbb{M}}(x, y)$, we have that

$$d_{\mathbb{M}}(x, y) > \mathrm{rch}(\mathbb{M}) \implies \|x - y\|_2 > \frac{\mathrm{rch}(\mathbb{M})}{2}.$$

3 Results

Let $\mathbb{M} \subset \mathbb{R}^n$ be a smooth, bounded, and boundary-less k-dimensional submanifold of $\mathbb{R}^n$. We let $\mathrm{diam}(\mathbb{M})$, $\mathrm{vol}_k(\mathbb{M})$, and $\mathrm{rch}(\mathbb{M}) > 0$ denote, respectively, the diameter, k-dimensional volume, and reach of $\mathbb{M}$. For $0 < l \leq 1 \leq u < \infty$ and for $m \geq k$, let $\Phi_{l,u} : \mathbb{R}^n \to \mathbb{R}^m$ be a smooth bi-Lipschitz map on $\mathbb{M}$ so that

$$l \cdot \|x - y\|_2 \leq \|\Phi_{l,u}(x) - \Phi_{l,u}(y)\|_2 \leq u \cdot \|x - y\|_2, \qquad \forall x, y \in \mathbb{M}. \tag{2}$$

Some basic properties of $\Phi_{l,u}(\mathbb{M})$, the image of $\mathbb{M}$ under $\Phi_{l,u}(\cdot)$, are collected below and proved in Appendix B. One key insight used in the proof is the following: Under $\Phi_{l,u}(\cdot)$, a tangent vector $v \in \mathbb{T}_x\mathbb{M}$ is mapped to a tangent vector at $\Phi_{l,u}(x)$. To be precise, $v \to D\Phi_{l,u}(x) \cdot v$, where $D\Phi_{l,u}(x) \in \mathbb{R}^{m \times n}$ is the derivative of $\Phi_{l,u}(\cdot)$ at x. Moreover, $D\Phi_{l,u}(x)$ acts as a near-isometry on $\mathbb{T}_x\mathbb{M}$ if $\Phi_{l,u}(\cdot)$ is itself a near-isometry on $\mathbb{M}$.

Proposition 3. (Basic properties) *Let the manifold $\mathbb{M}$ and map $\Phi_{l,u}(\cdot)$ be as specified above. Then, $\Phi_{l,u}(\mathbb{M})$ is itself a smooth, bounded, and boundary-less k-dimensional submanifold of $\mathbb{R}^m$. Moreover, it holds that*

$$l \cdot diam(\mathbb{M}) \leq diam\left(\Phi_{l,u}(\mathbb{M})\right) \leq u \cdot diam(\mathbb{M}),$$

$$l^k \cdot vol_k(\mathbb{M}) \leq vol_k\left(\Phi_{l,u}(\mathbb{M})\right) \leq u^k \cdot vol_k(\mathbb{M}).$$

Further discussion of this result appears in Section 4.

Before we study how a nearly isometric map affects the reach of a smooth manifold, some results are in order to build intuition and form insight to the problem. The next result, proved in Appendix C, states that (not surprisingly) the tangent structure of the manifold remains entirely intact under an exact isometry on $\mathbb{M}$, i.e., when $l = u = 1$. The key insight used in the proof is that the angle between any pair of tangent vectors in $\mathbb{M}$ is preserved in $\Phi_{l,u}(\mathbb{M})$ if $\Phi_{l,u}(\cdot)$ is an exact isometry on $\mathbb{M}$.

Proposition 4. (Tangent bundle under exact isometry) *For the manifold $\mathbb{M}$ specified earlier, take $p, q \in \mathbb{M}$. When $l = u = 1$, it holds that*

$$\angle\left[\mathbb{T}_{\Phi_{l,u}(p)}\Phi_{l,u}(\mathbb{M}), \mathbb{T}_{\Phi_{l,u}(q)}\Phi_{l,u}(\mathbb{M})\right] = \angle\left[\mathbb{T}_p\mathbb{M}, \mathbb{T}_q\mathbb{M}\right],$$

with $\angle[\mathbb{A}, \mathbb{B}]$ standing for the angle between the subspaces $\mathbb{A}$ and $\mathbb{B}$.[7]

---

[5]Loosely speaking, a geodesic curve between a pair of points on $\mathbb{M}$ is a curve on $\mathbb{M}$ that locally minimizes the distances [16]. For example, an arc on a circle is a geodesic curve between its end points. Note that the geodesic curve between two points is not necessarily unique.

[6]The geodesic distance between a pair of points on $\mathbb{M}$ is the length of the shortest geodesic curve connecting those two points [16].

[7]Angle between subspaces generalizes the notion of angle between lines. The (largest principle) angle between the two subspaces $\mathbb{A}$ and $\mathbb{B}$ is defined such that $\cos(\angle[\mathbb{A}, \mathbb{B}]) := \min_a \max_b \cos(\angle[a, b])$, where the optimization is over all vectors $a \in \mathbb{A}$ and $b \in \mathbb{B}$. See [13] for more details.

To extend Proposition 4 to a general bi-Lipschitz map on $\mathbb{M}$ with $l \leq 1 \leq h$, we restrict ourselves to linear maps. The next result is proved in Appendix D.

Proposition 5. (Tangent bundle under linear near-isometry) *With the manifold $\mathbb{M}$ as before, take $p, q \in \mathbb{M}$. If the map $\Phi_{l,u} : \mathbb{R}^n \to \mathbb{R}^m$ is linear on $\mathbb{R}^n$ and bi-Lipschitz on $\mathbb{M}$ (i.e., satisfies (2)), then*

$$\left| \cos \left(\angle [\mathbb{T}_{\Phi_{l,u}(p)} \Phi_{l,u} (\mathbb{M}), \mathbb{T}_{\Phi_{l,u}(q)} \Phi_{l,u} (\mathbb{M})] \right) - \cos \left(\angle [\mathbb{T}_p \mathbb{M}, \mathbb{T}_q \mathbb{M}] \right) \right|$$

$$\leq \frac{25}{4l^2} \cdot \frac{\|p - q\|_2^2}{rch(\mathbb{M})^2} \cdot \left(\max(\|\Phi_{l,u}\|^2, u^2) + 1 \right)^2 + \frac{18 \max \left(\Delta_{l,u}, \sqrt{\Delta_{l,u}} \right)}{l^2},$$

where $\Delta_{l,u} := \max \left[1 - l^2, u^2 - 1 \right]$ and $\|\Phi_{l,u}\|$ is the spectral norm of the matrix representation of the linear operator $\Phi_{l,u}(\cdot)$.

In words, the angles between tangent spaces on the manifold remain *nearly* unchanged under a linear near-isometry ($l, u \approx 1$), for sufficiently close points on the manifold (those within a distance of approximately $\|\Phi_{l,u}\|^{-2} rch(\mathbb{M})$). The proof of this fact involves sampling two additional points from the manifold, one near p and one near q, controlling the distances between all four of these points under the bi-Lipschitz map $\Phi_{l,u}$, and relating these distances to the angles between the tangent spaces at p and q.

Finally, what can be said about $rch(\Phi_{l,u}(\mathbb{M}))$? The next result, proved in Appendix E, deals with the special case of an exact isometry.

Proposition 6. (Reach under exact isometry) *Consider the manifold $\mathbb{M}$ as before, and take $l = u = 1$. Then,*

$$rch \left(\Phi_{l,u} (\mathbb{M}) \right) = rch \left(\mathbb{M} \right).$$

We note that Proposition 6 does not require $\Phi_{1,1} : \mathbb{R}^n \to \mathbb{R}^m$ to be linear, nor does it place any restriction on m (aside from the requirement—assumed throughout this section—that $m \geq k$). The key insight here is that $\Phi_{1,1}(\cdot)$ preserves not only pairwise distances on $\mathbb{M}$, but also the angles between pairs of tangent spaces of $\mathbb{M}$ (see Proposition 4). This in turn guarantees that $\Phi_{1,1}(\mathbb{M})$ is as regular as $\mathbb{M}$ itself (as recorded in the above proposition).

What about the case $l < 1$ or $u > 1$, where $\Phi_{l,u}(\cdot)$ is not an exact isometry on $\mathbb{M}$? In full generality, there is no relation between $rch(\Phi_{l,u}(\mathbb{M}))$ and $rch(\mathbb{M})$. This is demonstrated in Appendix F where, by adding a smooth "cusp" of width 2δ to a line segment in $\mathbb{R}^2$, we explicitly construct a manifold $\mathbb{M}$ and a near-isometric map $\Phi_{l,u}(\cdot)$ on $\mathbb{M}$ such that $rch(\Phi_{l,u}(\mathbb{M})) \leq \frac{\delta}{\sqrt{2}} \ll \infty = rch(\mathbb{M})$.

Proposition 7. (Negative result) *In general, reach is not preserved under a bi-Lipschitz mapping.*

When $\Phi_{l,u}(\cdot)$ is linear, however, we can lower bound $rch(\Phi_{l,u}(\mathbb{M}))$ in terms of $rch(\mathbb{M})$; the following theorem is proved in Appendix G.

Theorem 8. (Reach under linear near-isometry) *Consider the manifold $\mathbb{M}$ specified earlier. Suppose $m \leq n$, and suppose $\Phi_{l,u}(\cdot)$ to be a rank-m linear map from $\mathbb{R}^n$ to $\mathbb{R}^m$ with its m nonzero singular values in the interval $[\sigma_{\min}, \sigma_{\max}] \subset (0, \infty)$. Then, if $m = n$, it holds that*

$$rch \left(\Phi_{l,u}(\mathbb{M}) \right) \geq \frac{\sigma_{\min}^2}{\sigma_{\max}} \cdot rch(\mathbb{M}),$$

whereas, if $m < n$, it holds that

$$rch\left(\Phi_{l,u}(\mathbb{M})\right) \geq \frac{\sigma_{\min}^2 l^2}{\sigma_{\max}^3} \cdot rch(\mathbb{M}).$$

We note that Theorem 8 does require $\Phi_{l,u} : \mathbb{R}^n \to \mathbb{R}^m$ to be linear, and it also requires that $m \leq n$ (in addition to the requirement that $m \geq k$). The proof, without any loss of generality, models the linear map $\Phi_{l,u}(\cdot)$ as an orthogonal projection followed by stretching in $\mathbb{R}^m$ along different coordinates, and then records how the reach changes in each of these two steps. In the first step, we use Definition 1 to directly quantify the rather involved effect of orthogonal projection on the reach. On the other hand, qualitatively speaking, modestly scaling an object along different coordinates will not substantially distort its geometry, and we will solidify this notion in the second step of the proof. Combining the two steps proves Theorem 8.

4 Discussion

Manifold embeddings have a rich history in differential topology and geometry. The following discussion might help put our results in the proper context. First, the (strong) Whitney embedding theorem [23] states that any k-dimensional smooth manifold $\mathbb{M}$ (not necessarily originating as a submanifold of any Euclidean space) can be smoothly embedded into $\mathbb{R}^{2k}$.[8] In contrast, the type of bi-Lipschitz map $\Phi_{l,u} : \mathbb{R}^n \to \mathbb{R}^m$ that we consider in this paper can be interpreted as *stably* embedding $\mathbb{M}$ (which does originate as a submanifold of $\mathbb{R}^n$) into $\mathbb{R}^m$. In particular, stability in this sense refers to the preservation of extrinsic, Euclidean distances as formalized in (1). As noted in [1], this type of embedding can also imply the preservation of intrinsic, geodesic distances. The celebrated Nash embedding theorem [18] also concerns the embedding of manifolds in Euclidean space, but with an exact isometry. However, isometry in this context refers only to the preservation of intrinsic, geodesic distances.

In Section 5, we discuss how certain randomized constructions have been used to construct linear mappings $\Phi_{l,u} : \mathbb{R}^n \to \mathbb{R}^m$ that satisfy (1) with m as small as possible. Deterministic linear constructions, adapted to the structure of the manifold $\mathbb{M}$ have also been proposed [2, 11] to satisfy (1). Other manifold learning algorithms such as ISOMAP [21], however, typically involve nonlinear maps designed to preserve only intrinsic, geodesic distances.

Regarding the implications of bi-Lipschitz mappings that we have derived in this paper, let us begin by discussing the assumptions made in Proposition 3. First, it is worth noting that the dimension of a manifold is actually invariant under any homeomorphism.[9] One might also consider replacing the bi-Lipschitz map in Proposition 3 with a "locally" bi-Lipschitz map, one that barely distorts the angle between any pair of tangent vectors anchored at the same point on $\mathbb{M}$. It is not difficult to verify that the dimension of $\mathbb{M}$ remains invariant under a locally bi-Lipschitz map and that such a map only modestly changes the volume. The diameter and reach, however, might change drastically under a locally bi-Lipschitz map. One example is rolling a thin and very long strip of paper along its length (as one would roll up a carpet); this transformation does not distort the angle between any pair of tangent vectors (at the same point) but can arbitrarily reduce the diameter and reach. On the other hand, it is possible to construct an operator that maps such a manifold to itself while significantly distorting the tangent vectors (one could envision placing one's palm on a rubber sheet, and twisting). In this case, the diameter, volume, and reach would all remain the same. However, depending on the conditioning of the Jacobian at a point on the

---

[8]In this context, a smooth embedding is simply a diffeomorphism.

[9]A homeomorphism is a continuous and bijective map with a continuous inverse.

sheet, a pair of tangent vectors at that point could map to vectors with a significantly different angle between them.

Regarding Proposition 5, we believe that there may be some room for improvement. While this result confirms that the angles between tangent spaces are better preserved as the isometry of the bi-Lipschitz map tightens (i.e., as $\Delta_{l,u} \to 0$), for a fixed p and q, the right hand side of the bound does not go to zero. We do believe that some dependence on $\|p - q\|_2^2$ is natural: for points farther away, the isometry constant must be tighter in order to preserve the angles between the respective tangent spaces. However, it may be possible to derive a bound on the right hand side that scales with the product of $\|p - q\|_2^2$ and $\Delta_{l,u}$, rather than their sum.

Finally, let us explore the tightness of the bounds in Theorem 8 using some toy examples. For instance, consider $\mathbb{M}$ to be the unit circle in $\mathbb{R}^2$, and let $\Phi_{l,u} : \mathbb{R}^2 \to \mathbb{R}^2$ be the linear map specified by the matrix

$$\Phi_{l,u} = \begin{bmatrix} 3 & 0 \\ 0 & \frac{1}{2} \end{bmatrix}.$$

Then $\sigma_{\min} = 1/2$, $\sigma_{max} = 3$, and $l = 1/2$, $u = 3$. Here, $\Phi_{l,u}(\mathbb{M})$ is an ellipse with principal axes $r_1 = 3$ and $r_2 = 1/2$. Thus, by Proposition 2, $\text{rch}(\Phi_{l,u}(\mathbb{M})) = r_2^2/r_1 = 1/12$, which precisely matches the lower bound in Theorem 8.

As another example, suppose that $\mathbb{M}$ is the unit circle along the plane in $\mathbb{R}^3$ that passes through $e_1 = [1, 0, 0]^T$ and makes an angle θ with $e_2 = [0, 1, 0]^T$. Let $\Phi_{l,u} : \mathbb{R}^3 \to \mathbb{R}^2$ be the linear map specified by the matrix

$$\Phi_{l,u} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}.$$

Then, one can verify that $\sigma_{\min} = \sigma_{\max} = 1$, and that $l = \cos\theta$, $u = 1$. In this case, $\Phi_{l,u}(\mathbb{M})$ is an ellipse with principal axes $r_1 = 1$ and $r_2 = \cos\theta = l$. The reach of this ellipse is given by Proposition 2: $\text{rch}(\Phi_{l,u}(\mathbb{M})) = r_2^2/r_1 = l^2$, which exactly matches the lower bound in Theorem 8.[10]

However, it is not difficult to construct an example where the bound in Theorem 8 is *not* tight. Again suppose that $\mathbb{M}$ is the unit circle along the plane in $\mathbb{R}^3$ that passes through $e_1 = [1, 0, 0]^T$ and makes an angle θ with $e_2 = [0, 1, 0]^T$. Now, let $\Phi_{l,u} : \mathbb{R}^3 \to \mathbb{R}^2$ be the linear map specified by the matrix

$$\Phi_{l,u} = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}.$$

Then, one can verify that $\sigma_{\min} = 1$, $\sigma_{\max} = 2$, and that $l = \cos\theta$, $u = 2$. In this case, $\Phi_{l,u}(\mathbb{M})$ is an ellipse with principal axes $r_1 = 2$ and $r_2 = \cos\theta = l$. The reach of this ellipse is given by Proposition 2: $\text{rch}(\Phi_{l,u}(\mathbb{M})) = r_2^2/r_1 = \cos^2\theta/2$, which is larger than what Theorem 8 predicts, namely

$$\frac{\sigma_{\min}^2 l^2}{\sigma_{\max}^3} = \frac{\cos^2\theta}{8}.$$

This appears to be an artifact of the proof technique which, when $m < n$, decomposes $\Phi_{l,u}(\cdot)$ into two bi-Lipschitz maps with isometry constants of $\frac{l}{\sigma_{\max}} \leq \frac{u}{\sigma_{\min}}$ and $\sigma_{\min} \leq \sigma_{\max}$, respectively.

5   Applications in Dimensionality Reduction

We conclude by noting that the Question posed in Section 1 is strongly motivated by recent advances in signal processing and machine learning. The Information Age has carried with it the burden

---

[10]Note the lower bounds in Theorem 8 scale linearly if $\Phi_{l,u}(\cdot)$ were to be scaled. As a result, the same conclusion holds true if we replace the map $\Phi_{l,u}(\cdot)$ in this example with, say, $2\Phi_{l,u}(\cdot)$.

of acquiring, storing, processing, and communicating increasingly higher dimensional signals and data sets. Fortunately, in many cases, the information contained within a high-dimensional signal or data set actually obeys some sort of concise, low-dimensional model [4,9]. Of particular interest to us here is the common scenario where the data lives on a k-dimensional submanifold $\mathbb{M} \subset \mathbb{R}^n$ (and typically $k \ll n$) [6,9].

The low dimension of the manifold model motivates the use of *compressive measurements* for simplifying the data acquisition process. Rather than designing a (possibly very expensive) sensor to measure a signal $x \in \mathbb{R}^n$, for example, it often suffices to design a sensor that can measure a much shorter vector $\chi = \Phi_{l,u} x \in \mathbb{R}^m$, where $\Phi_{l,u} \in \mathbb{R}^{m \times n}$ is a linear measurement operator that acts as a near-isometry on $\mathbb{M}$ (see (1)), and where typically $m \ll n$. Data inference tasks (e.g., classification or parameter estimation) can be performed—both reliably and inexpensively—in the measurement space $\mathbb{R}^m$ (since $m \ll n$) [12]. In fact, it is even possible to reconstruct the original signal x given only $\chi = \Phi_{l,u} x$ (and the knowledge of $\mathbb{M}$ and $\Phi_{l,u}$) [15,20].

In this context, two important cases are worth mentioning in regards to Theorem 8. First, consider the case where the manifold $\mathbb{M}$ is concentrated in or around an m-dimensional subspace of $\mathbb{R}^n$, and suppose $\Phi_{l,u}$ is chosen as an orthogonal projection onto an orthogonal basis for that subspace. Such optimal projections are the topic of classical Principal Component Analysis (PCA). In this case, the m nonzero singular values of Φ will exactly equal 1, and if $l \approx 1$, Theorem 8 guarantees a lower bound on $\mathrm{rch}\left(\Phi_{l,u}(\mathbb{M})\right)$ that is approximately equal to $\mathrm{rch}(\mathbb{M})$.

As a second case, suppose Φ is generated randomly as an $m \times n$ matrix with $m < n$. Such matrices are commonly used as tool for dimensionality reduction in the field of Compressive Sensing [3, 4, 8]. Consider, in particular, the situation where Φ is populated with independent and identically distributed Gaussian random variables, each with zero mean and variance $1/m$. In this situation, it is known that assuming

$$ m \gtrsim \frac{k}{\Delta_{l,u}} \log\left[\frac{(\mathrm{vol}_k(\mathbb{M}))^{\frac{1}{k}}}{\mathrm{rch}(\mathbb{M})} \right], \tag{3} $$

then with high probability (1) will hold [9] (see also [6]).[11] In this case, one may indeed achieve a near-isometric embedding of the manifold with $l \approx u \approx 1$ and guarantee that the reach of the manifold does not collapse. However, the singular values of Φ will cluster around $\sqrt{n/m}$. Consequently, the lower bound on $\mathrm{rch}\left(\Phi_{l,u}(\mathbb{M})\right)$ will be lower than in the first case: around $\sqrt{m/n} \cdot \mathrm{rch}(\mathbb{M})$.

References

[1] R. G. Baraniuk and M. B. Wakin. Random projections of smooth manifolds. *Found. Comput. Math.*, 9(1):51–77, 2009.

[2] D. S. Broomhead and M. J. Kirby. The Whitney reduction network: A method for computing autoassociative graphs. *Neural Computation*, 13(11):2595–2616, 2001.

[3] E. J. Candès. The restricted isometry property and its implications for compressed sensing. *Comptes Rendus Mathematique*, 346(9):589–592, 2008.

[4] E. J. Candès and M. B. Wakin. An introduction to compressive sampling. *IEEE Signal Processing Magazine*, 25(2):21–30, 2008.

---

[11]We have somewhat simplified the right hand side above for clarity, hence the sign $\gtrsim$. Exact details may be found in [9].

[5] F. Chazal and R. Soufflet. Stability and finiteness properties of medial axis and skeleton. *Journal of Dynamical and Control Systems*, 10(2):149–170, 2004.

[6] K. L. Clarkson. Tighter bounds for random projections of manifolds. In *Proceedings of the twenty-fourth annual symposium on Computational geometry*, pages 39–48. ACM, 2008.

[7] M. A. Davenport, M. F. Duarte, M. B. Wakin, J. N. Laska, D. Takhar, K. F. Kelly, and R. G. Baraniuk. The smashed filter for compressive classification and target recognition. In *Proc. Comput. Imaging V at SPIE Electronic Imaging*, January 2007.

[8] D. L. Donoho. Compressed sensing. *IEEE Transactions on Information Theory*, 52(4):1289–1306, 2006.

[9] A. Eftekhari and M. B. Wakin. New analysis of manifold embeddings and signal recovery from compressive measurements. *Applied and Computational Harmonic Analysis*, 39(1):67–109, 2015.

[10] H. Federer. Curvature measures. *Trans. Amer. Math. Soc*, 93(3):418–491, 1959.

[11] C. Hegde, A. C. Sankaranarayanan, W. Yin, and R. G. Baraniuk. NuMax: A convex approach for learning near-isometric linear embeddings. *IEEE Transactions on Signal Processing*, 63(22):6109–6121, 2015.

[12] C. Hegde, M. Wakin, and R. Baraniuk. Random projections for manifold learning. In *Advances in neural information processing systems*, pages 641–648, 2008.

[13] R. A. Horn and C. R. Johnson. *Topics in Matrix Analysis*. Cambridge University Press, 1994.

[14] R. Howard. Blaschke's rolling theorem for manifolds with boundary. *Manuscripta Mathematica*, 99(4):471–483, 1999.

[15] M. A. Iwen and M. Maggioni. Approximation of points on low-dimensional manifolds via random linear projections. *Information and Inference*, 2(1):1–31, 2013.

[16] J. M. Lee. *Manifolds and Differential Geometry*. Graduate studies in mathematics. American Mathematical Society, 2009.

[17] F. Morgan. *Riemannian Geometry: A Beginners Guide, Second Edition*. A K Peters Series. Taylor & Francis, 1998.

[18] J. Nash. The imbedding problem for Riemannian manifolds. *Annals of Mathematics*, pages 20–63, 1956.

[19] P. Niyogi, S. Smale, and S. Weinberger. Finding the homology of submanifolds with high confidence from random samples. *Discrete Comput. Geom.*, 39(1):419–441, 2008.

[20] P. Shah and V. Chandrasekaran. Iterative projections for signal identification on manifolds: Global recovery guarantees. In *Allerton Conference on Communication, Control, and Computing*, pages 760–767. IEEE, 2011.

[21] J. B. Tenenbaum, V. De Silva, and J. C. Langford. A global geometric framework for nonlinear dimensionality reduction. *Science*, 290(5500):2319–2323, 2000.

[22] N. Verma. Distance preserving embeddings for general n-dimensional manifolds. *J. Machine Learning Research*, 14(Aug):2415–2448, 2013.

[23] H. Whitney. Differentiable manifolds. *Annals of Mathematics*, pages 645–680, 1936.

[24] H. L. Yap, M. B. Wakin, and C. J. Rozell. Stable manifold embeddings with structured random matrices. *IEEE J. Sel. Topics Signal Proc.*, 7(4):720–730, August 2013.

A Proof of Proposition 2 (Reach of an ellipsoid)

To establish a lower bound on the reach, we rely on Blaschke's rolling theorem, as stated in [5].

Theorem 9. (Blaschke's rolling theorem [5]) *Let $V \subset \mathbb{R}^n$ be a relatively compact convex open set. Assume that the boundary of V, ∂V, is a C^2 manifold. Assume that the maximum K of the principal curvatures at any point of ∂V is finite. Then, for all $0 < \epsilon < 1/K$ the Euclidean ball of radius ϵ can roll freely on ∂V in the interior of V. More precisely, for all $x \in \partial V$, the ball of radius ϵ which is tangent to ∂V at x has only x as an intersection point with ∂V.*

Letting V denote the interior of our ellipsoid (so that $\partial V = \mathbb{M}$), the problem of bounding $\mathrm{rch}(\mathbb{M})$ reduces to that of bounding the maximum K of the principal curvatures at any point of $\mathbb{M}$. Theorem 9 then implies that the open normal bundle of $\mathbb{M}$ of radius r is embedded in $\mathbb{R}^n$ for all $r < 1/K$. Therefore, $\mathrm{rch}(\mathbb{M}) \geq 1/K$.

To bound K, we consider the following argument. Without loss of generality, we suppose the ellipsoid is centered at the origin and aligned with the canonical axes, so that points $x \in \mathbb{M}$ are defined by the equation

$$\sum_{i=1}^{n} \frac{x_i^2}{r_i^2} = 1,$$

for $r_1 \geq \cdots \geq r_n > 0$. Equivalently, $x^T D x = 1$ with $D \in \mathbb{R}^{n \times n}$ the diagonal matrix containing $\{r_i^{-2}\}$.

Next, we consider a rigid transformation (which leaves reach unchanged) that shifts and rotates the ellipsoid as follows: for an arbitrary point $b \in \mathbb{M}$, we shift the point b to the origin and rotate the ellipsoid so that, at the origin, it is orthogonal to e_n, the nth canonical vector in $\mathbb{R}^n$. More precisely, we consider the transformation $x = Ry + b$ for some unitary rotation matrix $R \in \mathbb{R}^{n \times n}$ to be defined below. Under this change of variables,

$$1 = (Ry + b)^T D (Ry + b) = y^T R^T D R y + 2 b^T D R y + b^T D b.$$

Since $b \in \mathbb{M}$, we have

$$b^T D b = 1, \tag{4}$$

and the transformed ellipsoid becomes

$$y^T R^T D R y + 2 b^T D R y = 0,$$

which passes through the origin. Ignoring the quadratic terms, we find that $2 b^T D R y = 0$. Therefore, the normal direction at the origin is $D R^T b$. We choose R such that the normal vector is aligned with the last coordinate e_n, i.e.,

$$D R^T b = -\|D R^T b\|_2 \cdot e_n. \tag{5}$$

With this choice, the ellipsoid is defined by the equation

$$y_n = \frac{y^T R^T D R y}{2\|DR^T b\|_2},$$

(6)

which now passes through the origin and has the normal vector of e_n. Let us decompose $R^T D R$ according to the index n, i.e., we let

$$D' = \begin{bmatrix} D'_{11} & D'_{12} \\ D'^T_{12} & D'_{22} \end{bmatrix} := R^T D R,$$

for short, with

$$D'_{11} \in \mathbb{R}^{(n-1)\times(n-1)}, \quad D'_{12} \in \mathbb{R}^{n-1}, \quad D'_{22} \in \mathbb{R}.$$

Then, we can write (6) as

$$2\|DR^T b\|_2 \cdot y_n = y_{\backslash n}^T D'_{11} y_{\backslash n} + \left(2 y_{\backslash n}^T D'_{12}\right) y_n + D'_{22} y_n^2,$$

where $y_{\backslash n} \in \mathbb{R}^{n-1}$ contains all the entries of y except y_n. We now rotate the tangent plane $y_n = 0$ by making the change of variables $z = L y_{\backslash n} \in \mathbb{R}^{n-1}$ where the orthonormal matrix $L \in \mathbb{R}^{(n-1)\times(n-1)}$ diagonalizes D'_{11}. We arrive at

$$2\|DR^T b\|_2 \cdot y_n = z^T E z + \left(2 z^T F\right) y_n + D'_{22} y_n^2,$$

$$E := L^T D'_{11} L \in \mathbb{R}^{(n-1)\times(n-1)}, \quad F := L^T D'_{12} \in \mathbb{R}^{n-1},$$

where E is diagonal. Note that the first two terms (constant and linear) of the Taylor expansion of y_n in terms of z are both zero. The principal curvatures are the eigenvalues of the Hessian $[\frac{\partial^2 y_n}{\partial z_j \partial z_{j'}}]$ evaluated at $z = 0$. By taking derivatives of both sides, we may verify that

$$\frac{\partial^2 y_n}{\partial z_j \partial z_{j'}}\big|_{z=0} = \frac{E}{\|DR^T b\|_2}.$$

Note that

$$\left\| \frac{\partial^2 y_n}{\partial z_j \partial z_{j'}}\big|_{z=0} \right\| \leq \frac{\|E\|_2}{\|DR^T b\|_2} = \frac{\|L^T D'_{11} L\|}{\|DR^T b\|_2} \leq \frac{\|D\|}{\|DR^T b\|_2} = \frac{r_n^{-2}}{\|DR^T b\|_2},$$

where we recall that $r_n = \min_i r_i$. By design, $b^T D b = \sum b_i^2 r_i^{-2} = 1$ and consequently we note that

$$
\min_b \frac{\|DR^T b\|_2}{\sqrt{b^T D b}} = \min_b \frac{\|DR^T b\|_2}{\|D^{\frac{1}{2}} b\|_2}
$$

$$
= \min_c \frac{\|DR^T D^{-\frac{1}{2}} c\|_2}{\|c\|_2}
$$

$$
= \sqrt{\lambda_{min}\left[\left(D^{-\frac{1}{2}} R D^{\frac{1}{2}}\right) D \left(D^{\frac{1}{2}} R^T D^{-\frac{1}{2}}\right)\right]}
$$

$$
\geq r_1^{-1} \sqrt{\lambda_{min}\left[\left(D^{-\frac{1}{2}} R D^{\frac{1}{2}}\right)\left(D^{\frac{1}{2}} R^T D^{-\frac{1}{2}}\right)\right]}
$$

$$
= r_1^{-1} \sqrt{\lambda_{min}\left[D^{-\frac{1}{2}} R D R^T D^{-\frac{1}{2}}\right]}
$$

$$
= r_1^{-1} \sqrt{\lambda_{min}\left[D^{-1} R D R^T\right]}
$$

$$
= r_1^{-1} \sqrt{\lambda_{min}\left[D^{-1} R D\right]}
$$

$$
= r_1^{-1} \sqrt{\lambda_{min}\left[R\right]}
$$

$$
= r_1^{-1},
$$

where the sixth and eighth lines hold because the spectrum of a matrix is unchanged under a similarity transform. In the fourth line, we used the fact that $D \succcurlyeq r_1^{-2} \cdot I_n$ because $r_1 = \max_i r_i$. Overall, we arrive at

$$
\left\|\frac{\partial^2 y_n}{\partial z_j \partial z_{j'}}\big|_{z=0}\right\| \leq \frac{r_n^{-2}}{r_1^{-1}} = \frac{r_1}{r_n^2}.
$$

That is, the largest principal curvature of an ellipsoid is bounded above by $K = r_1/r_n^2$. In fact, this is achieved at each endpoint of the major axis of the ellipsoid, which means that the largest principal curvature equals r_1/r_n^2 exactly. From the rolling theorem, we conclude that the reach of the ellipsoid is lower bounded by r_n^2/r_1.

In fact, Blaschke's rolling theorem is tight (see [14]) in that—in the parlance of Theorem 9— the Euclidean ball of radius $1/K$ cannot roll freely on ∂V in the interior of V. Since we have established above that $K = r_1/r_n^2$ exactly, it follows that the reach of the ellipsoid is also upper bounded bounded by r_n^2/r_1, and therefore, must equal exactly r_n^2/r_1.

B Proof of Proposition 3 (Basic Properties)

We prove only the claims concerning dimension and volume, as the others are self-evident. We also set $\Phi(\cdot) = \Phi_{l,u}(\cdot)$ for short.

Let $x \in \mathbb{M}$ be arbitrary. Fix a tangent vector $v \in \mathbb{T}_x \mathbb{M}$, and consider a curve $\gamma : [-1, 1] \to \mathbb{M}$ such that

$$
\gamma(0) = x, \qquad \frac{d\gamma}{dt}(0) = v. \tag{7}
$$

Note that $\Phi \circ \gamma : [-1, 1] \to \Phi(\mathbb{M})$ is a curve on $\Phi(\mathbb{M})$ that passes through $\Phi(x) = \Phi(\gamma(0))$ with the tangent direction

$$
\frac{d(\Phi \circ \gamma)}{dt}(0) = D\Phi(\gamma(0)) \cdot \frac{d\gamma}{dt}(0) = D\Phi(x) \cdot v, \tag{8}
$$

where $D\Phi(x) \in \mathbb{R}^{m \times n}$ is the Jacobian of $\Phi(\cdot)$ at $x \in \mathbb{R}^n$. On the other hand,

$$
\begin{aligned}
\left\| \frac{d\,(\Phi \circ \gamma)}{dt}(0) \right\|_2 &= \lim_{t \to 0} \frac{\|\Phi(\gamma(t)) - \Phi(\gamma(0))\|_2}{t} \\
&\geq l \cdot \lim_{t \to 0} \frac{\|\gamma(t) - \gamma(0)\|_2}{t} \qquad (\text{see } (2)) \\
&= l \cdot \left\| \frac{d\gamma}{dt}(0) \right\|_2 \\
&= l \cdot \|v\|_2 > 0, \qquad (\text{see } (7))
\end{aligned}
\tag{9}
$$

where the second to last line holds because the derivative exists. Comparing (8) with (9), we conclude that the Jacobian $D\Phi(x) : \mathbb{T}_x\mathbb{M} \to \mathbb{T}_{\Phi(x)}\Phi(\mathbb{M})$ is an injective map. Since any tangent vector to $\Phi(\mathbb{M})$ at $\Phi(x)$ can be written as in (8) (for some curve $\gamma(\cdot)$), $D\Phi(x)$ is in fact a bijective map. Since the choice of $x \in \mathbb{M}$ was arbitrary, we conclude that $\dim(\Phi(\mathbb{M})) = \dim(\mathbb{M}) = k$.

In fact, a matching upper bound for (9) exists and we have that

$$
l\|v\|_2 \leq \|D\Phi(x) \cdot v\|_2 \leq u\|v\|_2, \qquad \forall v \in \mathbb{T}_x\mathbb{M}, \quad \forall x \in \mathbb{M}.
\tag{10}
$$

That is, $\Phi(\cdot)$ nearly preserves the lengths of tangent vectors of $\mathbb{M}$. To prove the last claim in Proposition 3, note that

$$
\text{vol}_k(\mathbb{M}) = \int_{x \in \mathbb{M}} \text{vol}_k(dx),
$$

where $\text{vol}_k(dx)$ is the volume of the parallelotope formed by columns of

$$
dx \in \overbrace{\mathbb{T}_x\mathbb{M} \times \cdots \times \mathbb{T}_x\mathbb{M}}^{k \text{ times}} \subset \mathbb{R}^{n \times k}.
$$

Similarly,

$$
\text{vol}_k(\Phi(\mathbb{M})) = \int_{y \in \Phi(\mathbb{M})} \text{vol}_k(dy),
$$

where

$$
dy \in \overbrace{\mathbb{T}_{\Phi(x)}\Phi(\mathbb{M}) \times \cdots \times \mathbb{T}_{\Phi(x)}\Phi(\mathbb{M})}^{k \text{ times}} \subset \mathbb{R}^{m \times k}.
$$

From (1), recall that $\Phi(\cdot)$ is injective on $\mathbb{M}$. Then, with the change of variables $y = \Phi(x)$, we find that

$$
\text{vol}_k(\Phi(\mathbb{M})) = \int_{x \in \mathbb{M}} \text{vol}_k(D\Phi(x) \cdot dx).
\tag{11}
$$

Note that

$$
\text{vol}_k(D\Phi(x) \cdot dx) = \sqrt{\det [dx^T \cdot D\Phi(x)^T D\Phi(x) \cdot dx]}.
$$

On the other hand, by (10), it holds that

$$
l^2 \cdot dx^T \cdot dx \preccurlyeq dx^T \cdot D\Phi(x)^T D\Phi(x) \cdot dx \preccurlyeq u^2 \cdot dx^T \cdot dx,
$$

and therefore

$$
l^k \cdot \sqrt{\det [dx^T \cdot dx]} \leq \sqrt{\det [dx^T \cdot D\Phi(x)^T D\Phi(x) \cdot dx]} \leq u^k \cdot \sqrt{\det [dx^T \cdot dx]},
$$

which reduces to

$$l^k \cdot \text{vol}_k(dx) \leq \text{vol}_k\left(D\Phi(x) \cdot dx\right) \leq u^k \cdot \text{vol}_k(dx).$$

Substituting the bound above into (11) yields

$$l^k \cdot \text{vol}_k(\mathbb{M}) \leq \text{vol}_k\left(\Phi\left(\mathbb{M}\right)\right) \leq u^k \cdot \text{vol}_k(\mathbb{M}),$$

which completes the proof of Proposition 3.

C   Proof of Proposition 4 (Tangent Bundle Under Exact Isometry)

Fix $p, q \in \mathbb{M}$ and unit-length vectors $e_p \in \mathbb{T}_p\mathbb{M}$ and $e_q \in \mathbb{T}_q\mathbb{M}$. Let $\theta_p, \theta_q : [-1, 1] \to \mathbb{M}$ be geodesic curves that pass through p and q, respectively, such that

$$\theta_p(0) = p, \qquad e_p = \theta_p'(0) = \lim_{t \to 0} \frac{\theta_p(t) - \theta_p(0)}{t},$$

$$\theta_q(0) = q, \qquad e_q = \theta_q'(0). \tag{12}$$

For fixed $t > 0$, we consider $\theta_p(t) - p$ on one hand and $\theta_q(t) - q$ on the other hand. We set $\Phi(\cdot) = \Phi_{l,u}(\cdot)$ for short. Because $\Phi(\cdot)$ is an exact isometry on $\mathbb{M}$, we have for any $t > 0$ that

$$
\begin{aligned}
&\langle \Phi\left(\theta_p(t)\right) - \Phi(p), \Phi\left(\theta_q(t)\right) - \Phi(q) \rangle \\
&= \langle \Phi\left(\theta_p(t)\right), \Phi\left(\theta_q(t)\right) \rangle - \langle \Phi(p), \Phi\left(\theta_q(t)\right) \rangle - \langle \Phi\left(\theta_p(t)\right), \Phi(q) \rangle + \langle \Phi(p), \Phi(q) \rangle \\
&= \frac{1}{2}\|\Phi\left(\theta_p(t)\right) - \Phi(q)\|^2 + \frac{1}{2}\|\Phi(p) - \Phi\left(\theta_q(t)\right)\|^2 \\
&\quad - \frac{1}{2}\|\Phi\left(\theta_p(t)\right) - \Phi\left(\theta_q(t)\right)\|^2 - \frac{1}{2}\|\Phi(p) - \Phi(q)\|^2 \\
&= \frac{1}{2}\|\theta_p(t) - q\|^2 + \frac{1}{2}\|p - \theta_q(t)\|^2 - \frac{1}{2}\|\theta_p(t) - \theta_q(t)\|^2 - \frac{1}{2}\|p - q\|^2 \\
&= \langle \theta_p(t), \theta_q(t) \rangle - \langle p, \theta_q(t) \rangle - \langle \theta_p(t), q \rangle + \langle p, q \rangle \\
&= \langle \theta_p(t) - p, \theta_q(t) - q \rangle,
\end{aligned}
$$

where the second and fourth equalities follow from the polarization identity. Hence,

$$
\begin{aligned}
\langle D\Phi(p) \cdot e_p, D\Phi(q) \cdot e_q \rangle &= \left\langle \lim_{t \to 0} \frac{\Phi\left(\theta_p(t)\right) - \Phi\left(p\right)}{t}, \lim_{t \to 0} \frac{\Phi\left(\theta_q(t)\right) - \Phi\left(q\right)}{t} \right\rangle \\
&= \lim_{t \to 0} \frac{\langle \Phi\left(\theta_p(t)\right) - \Phi\left(p\right), \Phi\left(\theta_q(t)\right) - \Phi\left(q\right) \rangle}{t^2} \\
&= \lim_{t \to 0} \frac{\langle \theta_p(t) - p, \theta_q(t) - q \rangle}{t^2} \\
&= \left\langle \lim_{t \to 0} \frac{\theta_p(t) - p}{t}, \lim_{t \to 0} \frac{\theta_q(t) - q}{t} \right\rangle \\
&= \langle e_p, e_q \rangle. \tag{13}
\end{aligned}
$$

Recalling the definition of principal angle, we then note that

$$
\begin{aligned}
\cos\left(\angle\left[\mathbb{T}_{\Phi(p)}\Phi(\mathbb{M}),\mathbb{T}_{\Phi(q)}\Phi(\mathbb{M})\right]\right) &= \min_{e_{\Phi,p}\in\mathbb{T}_{\Phi(p)}\Phi(\mathbb{M})}\ \max_{e_{\Phi,q}\in\mathbb{T}_{\Phi(q)}\Phi(\mathbb{M})}\cos\left(\angle\left[e_{\Phi,p},e_{\Phi,q}\right]\right) \\
&= \min_{e_p\in\mathbb{T}_p\mathbb{M}}\ \max_{e_q\in\mathbb{T}_q\mathbb{M}}\cos\left(\angle\left[D\Phi(p)\cdot e_p,D\Phi(q)\cdot e_q\right]\right) \\
&= \min_{e_p\in\mathbb{T}_p\mathbb{M}}\ \max_{e_q\in\mathbb{T}_q\mathbb{M}}\frac{\langle D\Phi(p)\cdot e_p,D\Phi(q)\cdot e_q\rangle}{\|D\Phi(p)\cdot e_p\|_2\cdot\|D\Phi(q)\cdot e_q\|_2} \\
&= \min_{e_p\in\mathbb{T}_p\mathbb{M}}\ \max_{e_q\in\mathbb{T}_q\mathbb{M}}\frac{\langle e_p,e_q\rangle}{\|e_p\|_2\cdot\|e_q\|_2} \qquad \text{(see (10) and (13))} \\
&= \cos\left(\angle\left[\mathbb{T}_p\mathbb{M},\mathbb{T}_q\mathbb{M}\right]\right).
\end{aligned}
$$

The second line above holds because $D\Phi(p):\mathbb{T}_p\mathbb{M}\to\mathbb{T}_{\Phi(p)}\Phi(\mathbb{M})$ is a bijective map (see (10)). This completes the proof of Proposition 4.[12]

D   Proof of Proposition 5 (Tangent Bundle Under Linear Near-Isometry)

For short, we will use the notation Φ instead of $\Phi_{l,u}$, the matrix representation of the linear map $\Phi_{l,u}(\cdot)$. Fix $p,q\in\mathbb{M}$, as well as $e_p\in\mathbb{T}_p\mathbb{M}$ and $e_q\in\mathbb{T}_q\mathbb{M}$. Let $\theta_p:[-1,1]\to\mathbb{M}$ and $\theta_q:[-1,1]\to\mathbb{M}$ be *unit speed*[13] geodesic curves passing through p and q, respectively, such that (12) holds. For some sufficiently small $\delta>0$, it holds that

$$
e_p=\theta_p'(0)\approx\frac{\theta_p(\delta)-\theta_p(0)}{\delta}, \qquad e_q=\theta_q'(0)\approx\frac{\theta_q(\delta)-\theta_q(0)}{\delta}.
$$

More concretely,

$$
\begin{aligned}
\left\|\theta_p'(0)-\frac{\theta_p(\delta)-\theta_p(0)}{\delta}\right\|_2 &= \frac{1}{\delta}\left\|\int_0^\delta\int_0^\alpha\theta_p''(\beta)\,d\beta d\alpha\right\|_2 \qquad \left(\theta''(\beta)=\frac{d^2\theta}{d\beta^2}(\beta)\right) \\
&\leq \frac{1}{\delta}\cdot\int_0^\delta\int_0^\alpha\left\|\theta_p''(\beta)\right\|_2\,d\beta d\alpha \\
&\leq \frac{1}{\delta\cdot\mathrm{rch}(\mathbb{M})}\cdot\int_0^\delta\int_0^\alpha d\beta d\alpha \qquad \left(\text{Sec. 2: }\left\|\theta''(\beta)\right\|_2\leq\frac{1}{\mathrm{rch}(\mathbb{M})}\right) \\
&= \frac{\delta}{2\cdot\mathrm{rch}(\mathbb{M})},
\end{aligned}
\tag{14}
$$

and similarly

$$
\left\|\theta_q'(0)-\frac{\theta_q(\delta)-\theta_q(0)}{\delta}\right\|_2\leq\frac{\delta}{2\cdot\mathrm{rch}(\mathbb{M})}.
\tag{15}
$$

To prove Proposition 5, we proceed as follows. First, we approximate $\langle e_p,e_q\rangle$ with a simpler quantity by noting that

$$
\begin{aligned}
\langle e_p,e_q\rangle &= \left\langle e_p-\frac{\theta_p(\delta)-\theta_p(0)}{\delta},e_q\right\rangle \\
&+ \left\langle\frac{\theta_p(\delta)-\theta_p(0)}{\delta},e_q-\frac{\theta_q(\delta)-\theta_q(0)}{\delta}\right\rangle+\left\langle\frac{\theta_p(\delta)-\theta_p(0)}{\delta},\frac{\theta_q(\delta)-\theta_q(0)}{\delta}\right\rangle.
\end{aligned}
\tag{16}
$$

---

[12]For subspaces $\mathbb{A},\mathbb{B}$ and their orthogonal complements $\mathbb{A}^\perp,\mathbb{B}^\perp$, it holds that $\angle[\mathbb{A},\mathbb{B}]=\angle[\mathbb{A}^\perp,\mathbb{B}^\perp]$. Therefore, a similar claim to that in Proposition 4 holds for normal spaces under an exact isometry.

[13]$\theta(\cdot)$ is a unit-speed curve if $\|\theta'(\cdot)\|_2=1$.

The first two terms on the right hand side above are small thanks to (14) and (15). Specifically,

$$
\left| \langle e_p, e_q \rangle - \left\langle \frac{\theta_p(\delta) - \theta_p(0)}{\delta}, \frac{\theta_q(\delta) - \theta_q(0)}{\delta} \right\rangle \right|
$$

$$
\leq \left| \left\langle e_p - \frac{\theta_p(\delta) - \theta_p(0)}{\delta}, e_q \right\rangle \right| + \left| \left\langle \frac{\theta_p(\delta) - \theta_p(0)}{\delta}, e_q - \frac{\theta_q(\delta) - \theta_q(0)}{\delta} \right\rangle \right|
$$

$$
\leq \left\| e_p - \frac{\theta_p(\delta) - \theta_p(0)}{\delta} \right\|_2 \cdot \| e_q \|_2
$$

$$
+ \left(\| e_p \|_2 + \left\| e_p - \frac{\theta_p(\delta) - \theta_p(0)}{\delta} \right\|_2 \right) \cdot \left\| e_q - \frac{\theta_q(\delta) - \theta_q(0)}{\delta} \right\|_2
$$

$$
\leq \frac{\delta}{2 \cdot \mathrm{rch}(\mathbb{M})} + \left(1 + \frac{\delta}{2 \cdot \mathrm{rch}(\mathbb{M})} \right) \cdot \frac{\delta}{2 \cdot \mathrm{rch}(\mathbb{M})} \qquad (\| e_p \|_2 = \| e_q \|_2 = 1, \text{ see (14) and (15)})
$$

$$
\leq \frac{5\delta}{4 \cdot \mathrm{rch}(\mathbb{M})}. \qquad (\text{if } \delta \leq \mathrm{rch}(\mathbb{M})) \tag{17}
$$

Similarly, we approximate $\langle \Phi e_p, \Phi e_q \rangle$ with a more technically amenable quantity:

$$
\left| \langle \Phi e_p, \Phi e_q \rangle - \left\langle \Phi \cdot \frac{\theta_p(\delta) - \theta_p(0)}{\delta}, \Phi \cdot \frac{\theta_q(\delta) - \theta_q(0)}{\delta} \right\rangle \right| \leq \frac{5 \| \Phi \|^2 \delta}{4 \cdot \mathrm{rch}(\mathbb{M})}. \quad (\text{if } \delta \leq \mathrm{rch}(\mathbb{M})) \tag{18}
$$

Above, $\| \Phi \|$ is the spectral norm of the matrix Φ. We now use the triangle inequality, (17), and (18) to simplify the difference $\langle \Phi e_p, \Phi e_q \rangle - \langle e_p, e_q \rangle$ as follows:

$$
| \langle \Phi e_p, \Phi e_q \rangle - \langle e_p, e_q \rangle |
$$

$$
\leq \left| \left\langle \Phi \cdot \frac{\theta_p(\delta) - \theta_p(0)}{\delta}, \Phi \cdot \frac{\theta_q(\delta) - \theta_q(0)}{\delta} \right\rangle - \left\langle \frac{\theta_p(\delta) - \theta_p(0)}{\delta}, \frac{\theta_q(\delta) - \theta_q(0)}{\delta} \right\rangle \right|
$$

$$
+ \frac{5 \left(\| \Phi \|^2 + 1 \right) \cdot \delta}{4 \cdot \mathrm{rch}(\mathbb{M})}. \tag{19}
$$

To control the difference on the right hand side above, we write that

$$
\left\langle \Phi \cdot \frac{\theta_p(\delta) - \theta_p(0)}{\delta}, \Phi \cdot \frac{\theta_q(\delta) - \theta_q(0)}{\delta} \right\rangle
$$

$$
= \frac{1}{2\delta^2} \left(\| \Phi \left(\theta_p(\delta) - \theta_q(0) \right) \|_2^2 + \| \Phi \left(\theta_p(0) - \theta_q(\delta) \right) \|_2^2 - \| \Phi \left(\theta_p(\delta) - \theta_q(\delta) \right) \|_2^2 - \| \Phi \left(\theta_p(0) - \theta_q(0) \right) \|_2^2 \right)
$$

$$
\leq \frac{1}{2\delta^2} \left(u^2 \| \theta_p(\delta) - \theta_q(0) \|_2^2 + u^2 \| \theta_p(0) - \theta_q(\delta) \|_2^2 - l^2 \| \theta_p(\delta) - \theta_q(\delta) \|_2^2 - l^2 \| \theta_p(0) - \theta_q(0) \|_2^2 \right) \qquad (\text{see (2)})
$$

$$
= \frac{1}{2\delta^2} \left(\| \theta_p(\delta) - \theta_q(0) \|_2^2 + \| \theta_p(0) - \theta_q(\delta) \|_2^2 - \| \theta_p(\delta) - \theta_q(\delta) \|_2^2 - \| \theta_p(0) - \theta_q(0) \|_2^2 \right)
$$

$$
+ \frac{1}{2\delta^2} \left((u^2 - 1) \| \theta_p(\delta) - \theta_q(0) \|_2^2 + (u^2 - 1) \| \theta_p(0) - \theta_q(\delta) \|_2^2 \right.
$$

$$
\left. + (1 - l^2) \| \theta_p(\delta) - \theta_q(\delta) \|_2^2 + (1 - l^2) \| \theta_p(0) - \theta_q(0) \|_2^2 \right)
$$

$$
\leq \frac{1}{2\delta^2} \left(\| \theta_p(\delta) - \theta_q(0) \|_2^2 + \| \theta_p(0) - \theta_q(\delta) \|_2^2 - \| \theta_p(\delta) - \theta_q(\delta) \|_2^2 - \| \theta_p(0) - \theta_q(0) \|_2^2 \right)
$$

$$
+ \frac{4\Delta_{l,u}}{2\delta^2} \left(\| \theta_p(0) - \theta_q(0) \|_2 + 2\delta \right)^2 \qquad \left(\Delta_{l,u} := \max \left[1 - l^2, u^2 - 1 \right], \text{ triangle ineq.} \right)
$$

$$
= \frac{1}{\delta^2} \langle \theta_p(\delta) - \theta_p(0), \theta_q(\delta) - \theta_q(0) \rangle + \frac{2\Delta_{l,u}}{\delta^2} \left(\| \theta_p(0) - \theta_q(0) \|_2 + 2\delta \right)^2 \qquad (\text{polarization id.})
$$

$$
= \left\langle \frac{\theta_p(\delta) - \theta_p(0)}{\delta}, \frac{\theta_q(\delta) - \theta_q(0)}{\delta} \right\rangle + \frac{2\Delta_{l,u}}{\delta^2} \left(\| \theta_p(0) - \theta_q(0) \|_2 + 2\delta \right)^2.
$$

16

Above, the first equality follows from the polarization identity. A matching lower bound is obtained similarly:

$$\left\langle \Phi \cdot \frac{\theta_p(\delta) - \theta_p(0)}{\delta}, \Phi \cdot \frac{\theta_q(\delta) - \theta_q(0)}{\delta} \right\rangle$$

$$\geq \left\langle \frac{\theta_p(\delta) - \theta_p(0)}{\delta}, \frac{\theta_q(\delta) - \theta_q(0)}{\delta} \right\rangle - \frac{2\Delta_{l,u}}{\delta^2} \left(\|\theta_p(0) - \theta_q(0)\|_2 + 2\delta \right)^2.$$

Substituting the upper and lower bounds above back into (19), we find that

$$|\langle \Phi e_p, \Phi e_q \rangle - \langle e_p, e_q \rangle|$$

$$\leq \frac{2\Delta_{l,u}}{\delta^2} \left(\|\theta_p(0) - \theta_q(0)\|_2 + 2\delta \right)^2 + \frac{5 \left(\|\Phi\|^2 + 1 \right) \cdot \delta}{4 \cdot \text{rch}(\mathbb{M})}$$

$$\leq \frac{4\Delta_{l,u}}{\delta^2} \cdot \|p - q\|_2^2 + 16\Delta_{l,u} + \frac{5 \left(\|\Phi\|^2 + 1 \right) \cdot \delta}{4 \cdot \text{rch}(\mathbb{M})} \qquad \left((a+b)^2 \leq 2a^2 + 2b^2, \ \forall a, b \in \mathbb{R} \right) \qquad (20)$$

$$\leq \frac{4\Delta_{l,u}}{\delta^2} \cdot \|p - q\|_2^2 + 16\Delta_{l,u} + \frac{5 \left(\max(\|\Phi\|^2, u^2) + 1 \right) \cdot \delta}{4 \cdot \text{rch}(\mathbb{M})}. \qquad (21)$$

In the last line above we introduce the term $\max(\|\Phi\|^2, u^2)$. If the upper bound in (2) is achieved for any pair of points, then it is guaranteed that $\|\Phi\| \geq u$; however, if (2) is loose, then u^2 may dominate this maximum. Now, we may choose

$$\delta = \frac{4 \cdot \text{rch}(\mathbb{M}) \sqrt{\Delta_{l,u}}}{5 \left(\max(\|\Phi\|^2, u^2) + 1 \right)},$$

which is guaranteed to be less than or equal to $\text{rch}(\mathbb{M})$, as required, and we find that

$$|\langle \Phi e_p, \Phi e_q \rangle - \langle e_p, e_q \rangle| \leq \frac{25}{4} \cdot \frac{\|p - q\|_2^2}{\text{rch}(\mathbb{M})^2} \cdot \left(\max(\|\Phi\|^2, u^2) + 1 \right)^2 + 17 \max\left(\Delta_{l,u}, \sqrt{\Delta_{l,u}} \right). \qquad (22)$$

Equipped with this estimate, we argue that

$$\cos\left(\angle\left[\mathbb{T}_{\Phi p}\Phi\mathbb{M}, \mathbb{T}_{\Phi q}\Phi\mathbb{M}\right]\right])$$

$$= \min_{e_{\Phi,p}\in\mathbb{T}_{\Phi p}\Phi\mathbb{M}} \max_{e_{\Phi,q}\in\mathbb{T}_{\Phi q}\Phi\mathbb{M}} \cos\left(\angle\left[e_{\Phi,p}, e_{\Phi,q}\right]\right)$$

$$= \min_{e_p\in\mathbb{T}_p\mathbb{M}} \max_{e_q\in\mathbb{T}_q\mathbb{M}} \cos\left(\angle\left[\Phi e_p, \Phi e_q\right]\right)$$

$$= \min_{e_p\in\mathbb{T}_p\mathbb{M}} \max_{e_q\in\mathbb{T}_q\mathbb{M}} \frac{\langle\Phi e_p, \Phi e_q\rangle}{\|\Phi e_p\|_2 \cdot \|\Phi e_q\|_2}$$

$$\leq \min_{e_p\in\mathbb{T}_p\mathbb{M}} \max_{e_q\in\mathbb{T}_q\mathbb{M}} \frac{\langle e_p, e_q\rangle + |\langle\Phi e_p, \Phi e_q\rangle - \langle e_p, e_q\rangle|}{l \cdot \|e_p\|_2 \cdot l \cdot \|e_q\|_2} \qquad \text{(triangle inequality and (10))}$$

$$= \min_{e_p\in\mathbb{T}_p\mathbb{M}} \max_{e_q\in\mathbb{T}_q\mathbb{M}} \frac{\langle e_p, e_q\rangle + |\langle\Phi e_p, \Phi e_q\rangle - \langle e_p, e_q\rangle|}{l^2} \qquad (\|e_p\|_2 = \|e_q\|_2 = 1)$$

$$\leq \left(\frac{1}{l^2} \cdot \min_{e_p\in\mathbb{T}_p\mathbb{M}} \max_{e_q\in\mathbb{T}_q\mathbb{M}} \langle e_p, e_q\rangle\right) + \frac{25}{4l^2} \cdot \frac{\|p-q\|_2^2}{\text{rch}(\mathbb{M})^2} \cdot \left(\max(\|\Phi\|^2, u^2) + 1\right)^2$$

$$\qquad + \frac{17\max\left(\Delta_{l,u}, \sqrt{\Delta_{l,u}}\right)}{l^2} \qquad \text{(see (22))}$$

$$= l^{-2} \cdot \cos\left(\angle\left[\mathbb{T}_p\mathbb{M}, \mathbb{T}_q\mathbb{M}\right]\right) + \frac{25}{4l^2} \cdot \frac{\|p-q\|_2^2}{\text{rch}(\mathbb{M})^2} \cdot \left(\max(\|\Phi\|^2, u^2) + 1\right)^2 + \frac{17\max\left(\Delta_{l,u}, \sqrt{\Delta_{l,u}}\right)}{l^2}$$

$$= \cos\left(\angle\left[\mathbb{T}_p\mathbb{M}, \mathbb{T}_q\mathbb{M}\right]\right) + \frac{1-l^2}{l^2} \cdot \cos\left(\angle\left[\mathbb{T}_p\mathbb{M}, \mathbb{T}_q\mathbb{M}\right]\right)$$

$$\qquad + \frac{25}{4l^2} \cdot \frac{\|p-q\|_2^2}{\text{rch}(\mathbb{M})^2} \cdot \left(\max(\|\Phi\|^2, u^2) + 1\right)^2 + \frac{17\max\left(\Delta_{l,u}, \sqrt{\Delta_{l,u}}\right)}{l^2}$$

$$\leq \cos\left(\angle\left[\mathbb{T}_p\mathbb{M}, \mathbb{T}_q\mathbb{M}\right]\right) + \frac{1-l^2}{l^2} + \frac{25}{4l^2} \cdot \frac{\|p-q\|_2^2}{\text{rch}(\mathbb{M})^2} \cdot \left(\max(\|\Phi\|^2, u^2) + 1\right)^2 + \frac{17\max\left(\Delta_{l,u}, \sqrt{\Delta_{l,u}}\right)}{l^2}$$

$$\leq \cos\left(\angle\left[\mathbb{T}_p\mathbb{M}, \mathbb{T}_q\mathbb{M}\right]\right) + \frac{25}{4l^2} \cdot \frac{\|p-q\|_2^2}{\Delta_{l,u} \cdot \text{rch}(\mathbb{M})^2} \cdot \left(\max(\|\Phi\|^2, u^2) + 1\right)^2 + \frac{18\max\left(\Delta_{l,u}, \sqrt{\Delta_{l,u}}\right)}{l^2}.$$

A lower bound is obtained similarly:

$$\cos\left(\angle\left[\mathbb{T}_{\Phi p}\Phi\mathbb{M}, \mathbb{T}_{\Phi q}\Phi\mathbb{M}\right]\right)$$

$$\geq u^{-2} \cdot \cos\left(\angle\left[\mathbb{T}_p\mathbb{M}, \mathbb{T}_q\mathbb{M}\right]\right) - \frac{25}{4u^2} \cdot \frac{\|p-q\|_2^2}{\text{rch}(\mathbb{M})^2} \cdot \left(\max(\|\Phi\|^2, u^2) + 1\right)^2 - \frac{17\max\left(\Delta_{l,u}, \sqrt{\Delta_{l,u}}\right)}{u^2}$$

$$= \cos\left(\angle\left[\mathbb{T}_p\mathbb{M}, \mathbb{T}_q\mathbb{M}\right]\right) - \frac{u^2-1}{u^2} \cdot \cos\left(\angle\left[\mathbb{T}_p\mathbb{M}, \mathbb{T}_q\mathbb{M}\right]\right)$$

$$\qquad - \frac{25}{4u^2} \cdot \frac{\|p-q\|_2^2}{\text{rch}(\mathbb{M})^2} \cdot \left(\max(\|\Phi\|^2, u^2) + 1\right)^2 - \frac{17\max\left(\Delta_{l,u}, \sqrt{\Delta_{l,u}}\right)}{u^2}$$

$$\geq \cos\left(\angle\left[\mathbb{T}_p\mathbb{M}, \mathbb{T}_q\mathbb{M}\right]\right) - \frac{25}{4u^2} \cdot \frac{\|p-q\|_2^2}{\text{rch}(\mathbb{M})^2} \cdot \left(\max(\|\Phi\|^2, u^2) + 1\right)^2 - \frac{18\max\left(\Delta_{l,u}, \sqrt{\Delta_{l,u}}\right)}{u^2}$$

$$\geq \cos\left(\angle\left[\mathbb{T}_p\mathbb{M}, \mathbb{T}_q\mathbb{M}\right]\right) - \frac{25}{4l^2} \cdot \frac{\|p-q\|_2^2}{\text{rch}(\mathbb{M})^2} \cdot \left(\max(\|\Phi\|^2, u^2) + 1\right)^2 - \frac{18\max\left(\Delta_{l,u}, \sqrt{\Delta_{l,u}}\right)}{l^2},$$

where the last line uses the fact that $l \leq u$. This completes the proof of Proposition 5.

E Proof of Proposition 6 (Reach Under Exact Isometry)

By our assumption,

$$\|\Phi_{1,1}(x) - \Phi_{1,1}(y)\|_2 = \|x - y\|, \qquad \forall x, y \in \mathbb{M},$$

and we wish to determine $\text{rch}(\Phi_{1,1}(\mathbb{M}))$ in relation to $\text{rch}(\mathbb{M})$. We set $\Phi(\cdot) = \Phi_{1,1}(\cdot)$ for short throughout the proof.

Fix an arbitrary point $x \in \mathbb{M}$. Take another arbitrary point $y \in \mathbb{M}$ and let $\gamma(\cdot)$ be a unit-speed geodesic curve that passes through x and y so that

$$\gamma(0) = x, \quad \gamma(d) = y, \tag{23}$$

for some $d > 0$. Next, we consider $\Phi(x), \Phi(y) \in \Phi(\mathbb{M})$ and fix a normal vector $w_\Phi \in \mathbb{N}_{\Phi(x)}\Phi(\mathbb{M})$ with $\|w_\Phi\|_2 = 1$. For $r_\Phi > 0$ to be assigned later, we set

$$z_\Phi = \Phi(x) + r_\Phi \cdot w_\Phi. \tag{24}$$

Our objective is to show that, for sufficiently small r_Φ, z_Φ is closer to $\Phi(x)$ than any other point $\Phi(y) \in \Phi(\mathbb{M})$. To that end, we write that

$$
\begin{aligned}
&\|\Phi(y) - z_\Phi\|_2^2 - r_\Phi^2 \\
&= \|\Phi(y) - z_\Phi\|_2^2 - \|\Phi(x) - z_\Phi\|_2^2 \qquad \text{(see (24))} \\
&= \|\Phi(\gamma(d)) - z_\Phi\|_2^2 - \|\Phi(\gamma(0)) - z_\Phi\|_2^2 \qquad \text{(see (23))} \\
&= \int_0^d \frac{d}{dt} \|\Phi(\gamma(t)) - z_\Phi\|_2^2 \, dt \\
&= 2 \int_0^d \langle \Phi(\gamma(t)) - z_\Phi, D\Phi(\gamma(t)) \cdot \gamma'(t) \rangle \, dt \qquad \left(\gamma'(t) := \frac{d\gamma}{dt}(t) \right) \\
&= 2 \int_0^d \langle \Phi(\gamma(t)) - \Phi(\gamma(0)) - r_\Phi \cdot w_\Phi, D\Phi(\gamma(t)) \cdot \gamma'(t) \rangle \, dt \qquad \text{(see (24))} \\
&= 2 \int_0^d \langle \Phi(\gamma(t)) - \Phi(\gamma(0)), D\Phi(\gamma(t)) \cdot \gamma'(t) \rangle \, dt - 2r_\Phi \int_0^d \langle w_\Phi, D\Phi(\gamma(t)) \cdot \gamma'(t) \rangle \, dt \\
&= \left(2 \int_0^d \langle \Phi(\gamma(t)) - \Phi(\gamma(0)), D\Phi(\gamma(t)) \cdot \gamma'(t) \rangle \, dt \right) - 2r_\Phi \langle w_\Phi, \Phi(\gamma(d)) - \Phi(\gamma(0)) \rangle \\
&= \left(2 \int_0^d \langle \Phi(\gamma(t)) - \Phi(\gamma(0)), D\Phi(\gamma(t)) \cdot \gamma'(t) \rangle \, dt \right) - 2r_\Phi \langle w_\Phi, \Phi(y) - \Phi(x) \rangle. \tag{25}
\end{aligned}
$$

Above, we used the fact that

$$\frac{d}{dt} \|\Phi(\gamma(t)) - z_\Phi\|_2^2 = 2 \langle \Phi(\gamma(t)) - z_\Phi, D\Phi(\gamma(t)) \cdot \gamma'(t) \rangle,$$

where $D\Phi(\cdot) \in \mathbb{R}^{m \times n}$ is the Jacobian of $\Phi(\cdot)$. In order to deal with the two terms in the last line of (25), we invoke a variation of Proposition 4 (which we state without proof).

Lemma 10. *Take $p, q \in \mathbb{M}$ and fix $e \in \mathbb{T}_p\mathbb{M}$ with unit length. Then*

$$\langle \Phi(q) - \Phi(p), D\Phi(p) \cdot e \rangle = \langle q - p, e \rangle.$$

Moreover,

$$\angle\left[\Phi(q) - \Phi(p), \mathbb{T}_{\Phi(p)}\mathbb{M}\right] = \angle\left[q - p, \mathbb{T}_p\mathbb{M}\right],$$

$$\angle\left[\Phi(q) - \Phi(p), \mathbb{N}_{\Phi(p)}\mathbb{M}\right] = \angle\left[q - p, \mathbb{N}_p\mathbb{M}\right],$$

where $\angle[q - p, \mathbb{T}_p\mathbb{M}]$ is the smallest angle between $q - p$ and all directions in the subspace $\mathbb{T}_p\mathbb{M}$. The other angles are defined similarly.

Thanks to the above result, there exists a unit length normal vector $w \in \mathbb{N}_x\mathbb{M}$ for which we can write the following concerning the two inner products in the last line of (25):[14]

$$\left\langle \Phi\left(\gamma(t)\right) - \Phi\left(\gamma(0)\right), D\Phi(\gamma(t)) \cdot \gamma'(t) \right\rangle = \left\langle \gamma(t) - \gamma(0), \gamma'(t) \right\rangle,$$

$$\left\langle w_\Phi, \Phi(y) - \Phi(x) \right\rangle = \left\langle w, y - x \right\rangle.$$

We are now in position to rewrite the last line of (25) and find that

$$\|\Phi(y) - z_\Phi\|_2^2 - r_\Phi^2 = \left(2 \int_0^d \left\langle \gamma(t) - \gamma(0), \gamma'(t) \right\rangle \, dt \right) - 2r_\Phi \left\langle w, y - x \right\rangle.$$

We set $z = x + r_\Phi \cdot w$ and trace back our steps in (25) to arrive at

$$\|\Phi(y) - z_\Phi\|_2^2 - r_\Phi^2 = \left(2 \int_0^d \left\langle \gamma(t) - \gamma(0), \gamma'(t) \right\rangle \, dt \right) - 2r_\Phi \left\langle w, y - x \right\rangle$$

$$= \|y - z\|_2^2 - r_\Phi^2,$$

or

$$\|\Phi(y) - z_\Phi\|_2^2 = \|y - z\|_2^2. \tag{26}$$

By definition, as long as $r_\Phi < \mathrm{rch}(\mathbb{M})$, z is closer to x than it is to y, i.e.,

$$r_\Phi < \mathrm{rch}(\mathbb{M}) \implies \|z - y\|_2^2 > \|z - x\|_2^2 = r_\Phi^2.$$

Now, thanks to (26), as long as $r_\Phi < \mathrm{rch}(\mathbb{M})$, z_Φ is closer to $\Phi(x)$ than it is to $\Phi(y)$, i.e.,

$$r_\Phi < \mathrm{rch}(\mathbb{M}) \implies \|z_\Phi - \Phi(y)\|_2^2 > \|z_\Phi - \Phi(x)\|_2^2 = r_\Phi^2.$$

Our choice of $x, y \in \mathbb{M}$ and $w_\Phi \in \mathbb{N}_{\Phi(x)}\Phi(\mathbb{M})$ were arbitrary and therefore the reach of $\Phi(\mathbb{M})$ is at least $\mathrm{rch}(\mathbb{M})$, i.e., $\mathrm{rch}(\Phi(\mathbb{M})) \geq \mathrm{rch}(\mathbb{M})$. In the opposite direction, we find that $\mathrm{rch}(\mathbb{M}) \geq \mathrm{rch}(\Phi(\mathbb{M}))$, as the (inverse) map that takes $\Phi(\mathbb{M})$ back to $\mathbb{M}$ too is obviously an isometry. This completes the proof of Proposition 6.

F Proof of Proposition 7 (Negative Result)

Here, we establish Proposition 7 by constructing a suitable example. Suppose that $\mathbb{M} \subset \mathbb{R}^2$ is a part of the horizontal axis in $\mathbb{R}^2$. More specifically, we set

$$\mathbb{M} = \left\{ \begin{bmatrix} x \\ y_\mathbb{M}(x) \end{bmatrix} : y_\mathbb{M}(x) = 0, \quad |x| < 1 \right\}. \tag{27}$$

Note that $\mathrm{rch}(\mathbb{M}) = \infty$. We next design a smooth manifold $\mathbb{M}_* \subset \mathbb{R}^2$ such that

---

[14]For the second identity, the existence of normal vector w is guaranteed for the following reason. Set $\alpha = \angle[y - x, \mathbb{N}_x\mathbb{M}] = \angle[\Phi(y) - \Phi(x), \mathbb{N}_{\Phi(x)}\Phi(\mathbb{M})]$. Then, by definition of the angle between subspaces, for any angle $\beta \geq \alpha$, there must exist some unit length vector in $\mathbb{N}_x\mathbb{M}$ (say w) such that $\beta = \angle[y - x, w]$.
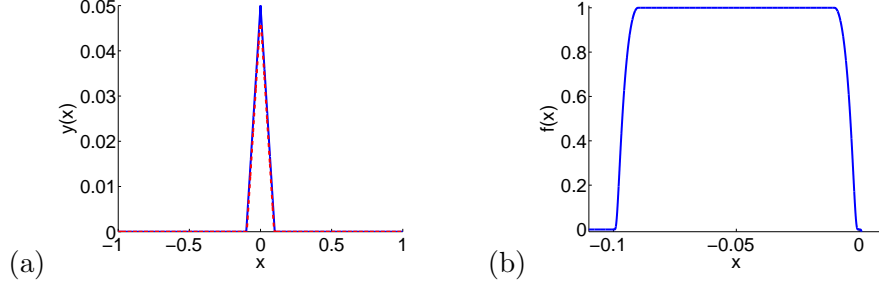
Figure 2: *Setup in Appendix F. (a) The solid blue curve shows a plot of $y_{\mathbb{M}_\circ}(x)$ as a function of x, while the dashed red curve shows a plot of $y_{\mathbb{M}_*}(x)$. (b) Plot of $f(x)$ for a narrow range of x. In all plots, $\delta = 0.1$, $c = 0.05$, and $\rho = 0.01$.*

- $\mathbb{M}_*$ is nearly-isometric to $\mathbb{M}$, and yet

- $\mathrm{rch}(\mathbb{M}_*) \approx 0 \ll \infty = \mathrm{rch}(\mathbb{M})$.

To that end, we proceed in two steps. First, we place a triangle wave function at the origin. More specifically, given $0 < c < \delta < 1$, let

$$
y_{\mathbb{M}_\circ}(x) = c \cdot \begin{cases} 1 - \frac{|x|}{\delta} & |x| \leq \delta, \\ 0 & \delta \leq |x| < 1, \end{cases} \tag{28}
$$

be the triangle wave. This specifies a manifold in $\mathbb{R}^2$ which we denote by $\mathbb{M}_\circ$:

$$
\mathbb{M}_\circ = \left\{ \begin{bmatrix} x \\ y_{\mathbb{M}_\circ}(x) \end{bmatrix} : |x| < 1 \right\} \subset \mathbb{R}^2.
$$

The solid blue curve in Figure 2(a) shows a plot of $y_{\mathbb{M}_\circ}(x)$. Note that $\mathbb{M}_\circ$ is not smooth because of the corners (cusps) at $[\pm\delta, 0]^T$ and at $[0, c]^T$. This manifold is therefore not suited for our example because the map between $\mathbb{M}$ and $\mathbb{M}_*$ must be, by our assumption, smooth. To work around this issue, we simply smooth out the corners of $\mathbb{M}_\circ$ as follows. Given $0 < \rho \ll \delta$ and when $x \in (0, 1)$, let

$$
f(x) = \begin{cases} -\exp\left(1 - \frac{1}{1 - \left(\frac{x - \rho}{\rho}\right)^2}\right) & 0 < x \leq \rho, \\ -1 & \rho \leq x \leq \delta - \rho, \\ -\exp\left(1 - \frac{1}{1 - \left(\frac{x - (\delta - \rho)}{\rho}\right)^2}\right) & \delta - \rho \leq x < \delta, \\ 0 & \delta \leq x, \end{cases} \tag{29}
$$

and, when $x \in (-1, 0)$, set

$$
f(x) = -f(-x), \qquad x < 0.
$$

We also set $f(0) = 0$. Closely related to the so-called *bump function*, $f(\cdot)$ is an odd function compactly supported on $[-\delta, \delta]$ which is also smooth (and in fact infinitely differentiable). We now specify $\mathbb{M}_*$ as follows:

$$
\mathbb{M}_* = \left\{ \begin{bmatrix} x \\ y_{\mathbb{M}_*}(x) \end{bmatrix} : y_{\mathbb{M}_*}(x) = \frac{c}{\delta} \cdot \int_{-\infty}^{x} f(z)\, dz, \quad |x| < 1 \right\} \subset \mathbb{R}^2. \tag{30}
$$

21

The dashed red curve in Figure 2(a) shows a plot of $y_{\mathbb{M}_*}(x)$. Note that $\mathbb{M}_*$ closely resembles $\mathbb{M}_\circ$ but, in contrast, is smooth. Furthermore, it is not difficult to verify that

$$\text{rch}\left(\mathbb{M}_\circ\right), \text{rch}\left(\mathbb{M}_*\right) \leq \frac{\delta}{\sqrt{2}}. \tag{31}$$

Indeed, for $\mathbb{M}_\circ$, the origin has a distance of at most $\delta/\sqrt{2}$ from both of the line segments starting at $[0, c]^T$ and ending at $[\pm\delta, 0]^T$. Therefore, $\text{rch}(\mathbb{M}_\circ) \leq \delta/\sqrt{2}$. The same statement holds true for $\mathbb{M}_*$ since the graph of $\mathbb{M}_*$ is also symmetric about the vertical axis and lies beneath that of $\mathbb{M}_\circ$ (i.e., $0 \leq y_{\mathbb{M}_*}(\cdot) \leq y_{\mathbb{M}_\circ}(\cdot)$), and lastly because the origin does not belong to $\mathbb{M}_*$.

Calculating the Isometry Constant of $\mathbb{M}_*$. We next show that $\mathbb{M}_*$ is nearly-isometric to $\mathbb{M}$. First, from (28) and (30), note that

$$\frac{c}{\delta} |f(x)| = \left| \frac{dy_{\mathbb{M}_*}}{dx}(x) \right| \leq \left| \frac{dy_{\mathbb{M}_\circ}}{dx}(x) \right|, \qquad \forall x \notin \{0, \pm\delta\}. \tag{32}$$

Then, from (32) and the fact that both $f(\cdot)$ and the derivative of $y_{\mathbb{M}_\circ}(\cdot)$ only change sign at the origin, we write the following whenever both $x_1, x_2 \in (-1, 1)$ have the same sign:

$$\begin{aligned}
|y_{\mathbb{M}_*}(x_1) - y_{\mathbb{M}_*}(x_2)| &= \frac{c}{\delta} \left| \int_{x_2}^{x_1} f(x)\, dx \right| \\
&= \frac{c}{\delta} \int_{x_2}^{x_1} |f(x)|\, dx \\
&\leq \int_{x_2}^{x_1} \left| \frac{dy_{\mathbb{M}_\circ}}{dx}(x) \right| dx \qquad \text{(see (32))} \\
&= \left| \int_{x_2}^{x_1} \frac{dy_{\mathbb{M}_\circ}}{dx}(x)\, dx \right| \\
&= |y_{\mathbb{M}_\circ}(x_1) - y_{\mathbb{M}_\circ}(x_2)|.
\end{aligned}$$

That is, on each side of the origin, $\mathbb{M}_*$ is a *contraction* of $\mathbb{M}_\circ$. This conclusion is in fact true everywhere. Indeed, for arbitrary $x_1, x_2 \in (-1, 1)$, we use the symmetry of both $\mathbb{M}_\circ$ and $\mathbb{M}_*$ to argue that

$$\begin{aligned}
|y_{\mathbb{M}_*}(x_1) - y_{\mathbb{M}_*}(x_2)| &= |y_{\mathbb{M}_*}(|x_1|) - y_{\mathbb{M}_*}(|x_2|)| \qquad \text{(by symmetry)} \\
&\leq |y_{\mathbb{M}_\circ}(x_1) - y_{\mathbb{M}_\circ}(x_2)|, \qquad \forall x_1, x_2 \in (-1, 1),
\end{aligned}$$

and, consequently,

$$1 \leq \frac{\left\| \begin{bmatrix} x_1 \\ y_{\mathbb{M}_*}(x_1) \end{bmatrix} - \begin{bmatrix} x_2 \\ y_{\mathbb{M}_*}(x_2) \end{bmatrix} \right\|_2}{\left\| \begin{bmatrix} x_1 \\ y_{\mathbb{M}}(x_1) \end{bmatrix} - \begin{bmatrix} x_2 \\ y_{\mathbb{M}}(x_2) \end{bmatrix} \right\|_2} \leq \frac{\left\| \begin{bmatrix} x_1 \\ y_{\mathbb{M}_\circ}(x_1) \end{bmatrix} - \begin{bmatrix} x_2 \\ y_{\mathbb{M}_\circ}(x_2) \end{bmatrix} \right\|_2}{\left\| \begin{bmatrix} x_1 \\ y_{\mathbb{M}}(x_1) \end{bmatrix} - \begin{bmatrix} x_2 \\ y_{\mathbb{M}}(x_2) \end{bmatrix} \right\|_2}, \quad \forall x_1, x_2 \in (-1, 1). \tag{33}$$

(Above, we invoked the fact that $y_{\mathbb{M}}(x) = 0$ for all x, so that the denominators above all equal $|x_1 - x_2|$.) In words, the isometry constant of $\mathbb{M}_*$ (as a mapping from $\mathbb{M}$) is always upper-bounded by the isometry constant of $\mathbb{M}_\circ$. We therefore focus on computing the latter quantity, as it is more convenient.

To that end, we use (27) and (28) as follows. There are a few possibilities which we next study one by one. For a pair of points x_1, x_2 far from the origin, we have that

$$\frac{\left\|\begin{bmatrix} x_1 \\ y_{\mathbb{M}_\circ}(x_1) \end{bmatrix} - \begin{bmatrix} x_2 \\ y_{\mathbb{M}_\circ}(x_2) \end{bmatrix}\right\|_2}{\left\|\begin{bmatrix} x_1 \\ y_{\mathbb{M}}(x_1) \end{bmatrix} - \begin{bmatrix} x_2 \\ y_{\mathbb{M}}(x_2) \end{bmatrix}\right\|_2} = \frac{\left\|\begin{bmatrix} x_1 - x_2 \\ 0 \end{bmatrix}\right\|_2}{\left\|\begin{bmatrix} x_1 - x_2 \\ 0 \end{bmatrix}\right\|_2} = 1, \qquad \forall |x_1|, |x_2| \in [\delta, 1).$$

When $|x_1| \in [0, \delta]$ and $|x_2| \in [\delta, 1)$, we have that

$$y_{\mathbb{M}_\circ}(x_1) - y_{\mathbb{M}_\circ}(x_2) = c\left(1 - \frac{|x_1|}{\delta}\right),$$

from which it follows that

$$1 \le \frac{\left\|\begin{bmatrix} x_1 \\ y_{\mathbb{M}_\circ}(x_1) \end{bmatrix} - \begin{bmatrix} x_2 \\ y_{\mathbb{M}_\circ}(x_2) \end{bmatrix}\right\|_2}{\left\|\begin{bmatrix} x_1 \\ y_{\mathbb{M}}(x_1) \end{bmatrix} - \begin{bmatrix} x_2 \\ y_{\mathbb{M}}(x_2) \end{bmatrix}\right\|_2} = \frac{\left\|\begin{bmatrix} x_1 - x_2 \\ c(1 - |x_1|/\delta) \end{bmatrix}\right\|_2}{\left\|\begin{bmatrix} x_1 - x_2 \\ 0 \end{bmatrix}\right\|_2} \le \sqrt{1 + \frac{c^2}{\delta^2}}$$

for all $|x_1| \in [0, \delta]$ and $|x_2| \in [\delta, 1)$. (Above, the ratio is maximized when $x_2 = \delta \cdot \text{sign}(x_1)$.) Finally, when both points are near the origin, we may again confirm that

$$1 \le \frac{\left\|\begin{bmatrix} x_1 \\ y_{\mathbb{M}_\circ}(x_1) \end{bmatrix} - \begin{bmatrix} x_2 \\ y_{\mathbb{M}_\circ}(x_2) \end{bmatrix}\right\|_2}{\left\|\begin{bmatrix} x_1 \\ y_{\mathbb{M}}(x_1) \end{bmatrix} - \begin{bmatrix} x_2 \\ y_{\mathbb{M}}(x_2) \end{bmatrix}\right\|_2} = \frac{\left\|\begin{bmatrix} x_1 - x_2 \\ -\frac{c}{\delta} \cdot (|x_1| - |x_2|) \end{bmatrix}\right\|_2}{\left\|\begin{bmatrix} x_1 - x_2 \\ 0 \end{bmatrix}\right\|_2} \le \sqrt{1 + \frac{c^2}{\delta^2}}$$

for all $|x_1|, |x_2| \le \delta$. (Above, we again invoked the symmetry of $\mathbb{M}_\circ$ about the vertical axis.) Overall, since $\sqrt{1 + \frac{c^2}{\delta^2}} \le 1 + \frac{c}{\delta}$, we arrive at

$$\left| \frac{\left\|\begin{bmatrix} x_1 \\ y_{\mathbb{M}_\circ}(x_1) \end{bmatrix} - \begin{bmatrix} x_2 \\ y_{\mathbb{M}_\circ}(x_2) \end{bmatrix}\right\|_2}{\left\|\begin{bmatrix} x_1 \\ y_{\mathbb{M}}(x_1) \end{bmatrix} - \begin{bmatrix} x_2 \\ y_{\mathbb{M}}(x_2) \end{bmatrix}\right\|_2} - 1 \right| \le \frac{c}{\delta}, \qquad \forall x_1, x_2 \in (-1, 1).$$

Owing to (33), we conclude that

$$\left| \frac{\left\|\begin{bmatrix} x_1 \\ y_{\mathbb{M}_*}(x_1) \end{bmatrix} - \begin{bmatrix} x_2 \\ y_{\mathbb{M}_*}(x_2) \end{bmatrix}\right\|_2}{\left\|\begin{bmatrix} x_1 \\ y_{\mathbb{M}}(x_1) \end{bmatrix} - \begin{bmatrix} x_2 \\ y_{\mathbb{M}}(x_2) \end{bmatrix}\right\|_2} - 1 \right| \le \frac{c}{\delta}, \qquad \forall x_1, x_2 \in (-1, 1). \tag{34}$$

Completing the Construction. Putting (31) and (34) together, we deduce the following. There exists a smooth map $\Phi : \mathbb{R}^2 \to \mathbb{R}^2$ such that $\mathbb{M}_* = \Phi(\mathbb{M})$. Moreover, $\Phi(\cdot)$ is a near-isometry on $\mathbb{M}$ with isometry constants $u = 1 + \frac{c}{\delta}$ and $l = 1$. In particular, since $c < \delta$ by design, $u < 2$ can be made arbitrarily close to one. Nevertheless,

$$\text{rch}\left(\Phi(\mathbb{M})\right) = \text{rch}\left(\mathbb{M}_*\right) \le \frac{\delta}{\sqrt{2}} \ll \infty = \text{rch}(\mathbb{M}).$$

This argument establishes that, in general, it is not possible to lower bound $\text{rch}\left(\Phi(\mathbb{M})\right)$ in terms of $\text{rch}(\mathbb{M})$. Since the mapping $\Phi(\cdot)$ is invertible on $\mathbb{M}$, we note that the entire argument can be reversed, and so in general it is also not possible to upper bound $\text{rch}\left(\Phi(\mathbb{M})\right)$ in terms of $\text{rch}(\mathbb{M})$.

G Proof of Theorem 8
(Reach Under Linear Near-Isometry)

For the sake of brevity, we will use the notation Φ instead of $\Phi_{l,u}$, the matrix representation of the linear map $\Phi_{l,u}(\cdot)$. Let $\Phi = U\Sigma V^* \in \mathbb{R}^{m \times n}$ be the singular value decomposition of Φ, with orthobasis $U \in \mathbb{R}^{m \times m}$, $\Sigma \in \mathbb{R}^{m \times n}$ defined as

$$\Sigma = \begin{bmatrix} \sigma_1 & & & & \\ & \sigma_2 & & & 0_{m \times (n-m)} \\ & & \ddots & & \\ & & & \sigma_m & \end{bmatrix} \in \mathbb{R}^{m \times n}, \tag{35}$$

and orthobasis $V \in \mathbb{R}^{n \times n}$. In particular, when $m = n$, Σ is a square and diagonal matrix. Note that $\Phi\mathbb{M} = U\Sigma V^*\mathbb{M}$. Since V^* acts as an isometry on all of $\mathbb{R}^n$, it follows from Proposition 6 that

$$\mathrm{rch}(\mathbb{M}') := \mathrm{rch}(V^*\mathbb{M}) = \mathrm{rch}(\mathbb{M}), \tag{36}$$

where we set $\mathbb{M}' := V^*\mathbb{M}$ to keep the notation compact. Thus, it suffices to consider the action of $U\Sigma$ on the rotated manifold $\mathbb{M}'$. Moreover, since U acts as an isometry on all of $\mathbb{R}^m$, we can apply Proposition 6 again and conclude that

$$\mathrm{rch}(\Phi\mathbb{M}) = \mathrm{rch}(U\Sigma V^*\mathbb{M}) = \mathrm{rch}(U\Sigma\mathbb{M}') = \mathrm{rch}(\Sigma\mathbb{M}').$$

Thus, it suffices to merely consider the action of Σ on $\mathbb{M}'$. For any pair of points $x, y \in M'$, note that $Vx, Vy \in \mathbb{M}$, that

$$\|x - y\|_2 = \|Vx - Vy\|_2,$$

and that

$$\|\Sigma x - \Sigma y\|_2 = \|U\Sigma V^*Vx - U\Sigma V^*Vy\|_2 = \|\Phi Vx - \Phi Vy\|_2.$$

Therefore, it follows from (2) that

$$l \cdot \|x - y\|_2 \le \|\Sigma x - \Sigma y\|_2 \le u \cdot \|x - y\|_2, \qquad \forall x, y \in \mathbb{M}'. \tag{37}$$

To summarize, (37) states that Σ acts as a near-isometry on $\mathbb{M}'$, and we wish to bound $\mathrm{rch}(\Sigma\mathbb{M}')$ in terms of $\mathrm{rch}(\mathbb{M}')$.

Using (35), we can factor Σ as follows:

$$\Sigma = \Sigma' \begin{bmatrix} I_m & 0_{m \times (n-m)} \end{bmatrix} =: \Sigma' \cdot I_{m,n} \in \mathbb{R}^{m \times n}, \tag{38}$$

where $\Sigma' \in \mathbb{R}^{m \times m}$ is a square diagonal matrix containing the nonzero singular values of Φ, and $I_{m,n} \in \mathbb{R}^{m \times n}$ is a rectangular matrix comprised of the $m \times m$ identity and an $m \times (n-m)$ matrix of zeros. In particular, when $m = n$, then $I_{m,n} = I_n$ is the identity matrix (and hence an exact isometry). By assumption, the diagonal entries of Σ' belong to the interval $[\sigma_{\min}, \sigma_{\max}]$, so that

$$\sigma_{\min}\|g\|_2 \le \|\Sigma'g\|_2 \le \sigma_{\max}\|g\|_2, \qquad \forall g \in \mathbb{R}^m. \tag{39}$$

In words, Σ' is a bi-Lipschitz map on $\mathbb{R}^m$. If $m < n$, the fact that Σ and Σ' are bi-Lipschitz maps on $\mathbb{M}'$ and $\mathbb{R}^m$, respectively, implies that $I_{m,n} \in \mathbb{R}^{m \times n}$ is a bi-Lipschitz map on $\mathbb{M}'$. Indeed, it follows from (37) and (39) that

$$\sigma_{\max}\|I_{m,n}(x - y)\|_2 \ge \|\Sigma' \cdot I_{m,n}(x - y)\|_2 = \|\Sigma(x - y)\|_2 \ge l\|x - y\|_2, \qquad \forall x, y \in \mathbb{M}',$$

and consequently,

$$\|I_{m,n}(x - y)\|_2 \ge \frac{l}{\sigma_{\max}}\|x - y\|_2, \qquad \forall x, y \in \mathbb{M}'. \tag{40}$$

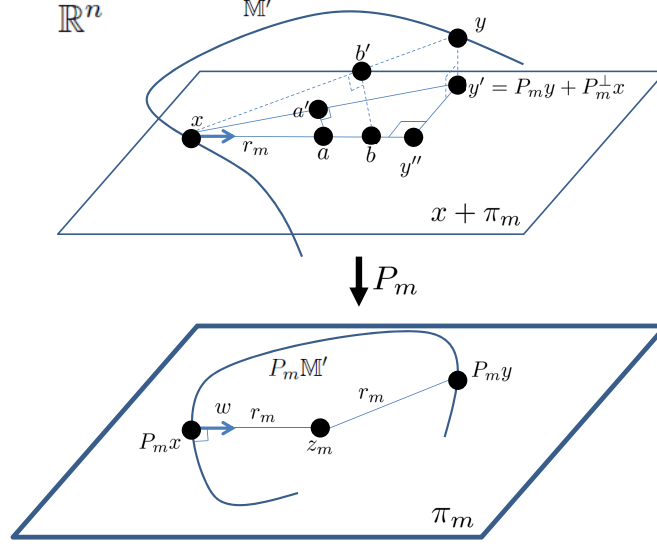We now proceed in two steps to prove the claim in Theorem 8.

Figure 3: *Proof construction in Appendix G.1.*

G.1   First Step

The first step is to compare $\mathrm{rch}(I_{m,n}\mathbb{M}')$ with $\mathrm{rch}(\mathbb{M}')$. When $m = n$, $I_{m,n} = I_n$ is the identity matrix and $I_{m,n}\mathbb{M}' = \mathbb{M}'$. Consequently, $\mathrm{rch}(I_{m,n}\mathbb{M}') = \mathrm{rch}(\mathbb{M}')$.

We therefore focus on the case where $m < n$. Let π_m denote the m-dimensional subspace spanning the first m coordinates in $\mathbb{R}^n$, and let $P_m x \in \mathbb{R}^n$ be the projection of x onto the subspace π_m. Note that $P_m\mathbb{M}'$ (the projection of $\mathbb{M}'$ onto the subspace π_m) is isometric to $I_{m,n}\mathbb{M}'$. By Proposition 6, we conclude that $\mathrm{rch}(P_m\mathbb{M}') = \mathrm{rch}(I_{m,n}\mathbb{M}')$. Therefore, we shift our attention to calculating $\mathrm{rch}(P_m\mathbb{M}')$.

For an arbitrary point $x \in \mathbb{M}'$, consider $P_m x$ and an arbitrary unit-length normal vector $w \in \mathbb{N}_{P_m x} P_m\mathbb{M}' \cap \pi_m$, i.e., w is a normal vector at $P_m x$ but along π_m. (The restriction of w to the subspace π_m is due to our interest in computing $\mathrm{rch}(P_m\mathbb{M}') \subset \pi_m$.) The constructions mentioned throughout this proof are illustrated in Figure 3. For $r_m > 0$ to be set later, consider the point $z_m = P_m x + r_m \cdot w \in \pi_m$. For small enough r_m, we wish to show that $P_m x$ is the unique nearest point to z_m on $P_m\mathbb{M}'$. To that end, if it at all exists, we consider another point $y \in \mathbb{M}'$ such that

$$r_m = \|z_m - P_m x\|_2 = \|z_m - P_m y\|_2. \tag{41}$$

Because $w = P_m w$ is supported only on the first m coordinates, w is also normal to $\mathbb{M}'$, i.e.,

$$w \in \mathbb{N}_{P_m x} P_m\mathbb{M}' \implies w \in \mathbb{N}_x \mathbb{M}'.$$

Consider the points

$$a := x + r_m \cdot w, \qquad b := x + \mathrm{rch}(\mathbb{M}') \cdot w,$$

and note that $a, b \in x + \pi_m$. (Here, $x + \pi_m$ is the affine subspace parallel to π_m and passing through x.) We keep in mind that, by Definition 1, $\|b - y\|_2 > \mathrm{rch}(\mathbb{M}') = \|b - x\|_2$.

Additionally, let

$$y' := P_m y + P_m^{\perp} x \in x + \pi_m,$$

be the projection of y onto the affine plane $x + \pi_m$. Let also y'' be the projection of y' onto the line $\overline{xa}$. Note that y' is the nearest point on the subspace $x + \pi_m$ to y and, in turn, y'' is the nearest

25

point on the line $\overline{xa} \subset x + \pi_m$ to y'. We conclude that y'' is the nearest point to y on the line $\overline{xa}$. In particular,

$$\overline{yy'} \perp (x + \pi_m), \qquad \overline{y'y''} \perp \overline{xa}, \qquad \overline{yy''} \perp \overline{xa}. \tag{42}$$

Also, let $a' \in x + \pi_m$ be the projection of a onto the line $\overline{xy'}$. Finally, let $b' \in \overline{xy}$ be the projection of b onto the line $\overline{xy}$. From (41), it follows that $\|x - a'\|_2 = \|y' - a'\|_2$, and that

$$r_m = \frac{\|x - a'\|_2}{\sin\left(\widehat{a'ax}\right)} = \frac{\frac{1}{2}\|x - y'\|_2}{\sin\left(\widehat{a'ax}\right)} = \frac{\|P_m(x - y)\|_2}{2\sin\left(\widehat{a'ax}\right)} = \frac{\|I_{m,n}(x - y)\|_2}{2\sin\left(\widehat{a'ax}\right)} \geq \frac{\frac{l}{\sigma_{\max}}\|x - y\|_2}{2\sin\left(\widehat{a'ax}\right)}, \tag{43}$$

where $\widehat{a'ax}$ is, of course, the angle formed by the lines $\overline{a'a}$ and $\overline{xa}$ (facing $\overline{xy'}$), and where the inequality in (43) follows from (40). In order to control r_m, we next calculate $\sin(\widehat{a'ax})$.

The geometry of the problem forces that

$$\widehat{a'ax} = \widehat{xy'y''}, \tag{44}$$

and that

$$\overline{y'y''} \perp \overline{xa} \implies \cos\left(\widehat{xy'y''}\right) = \frac{\|y' - y''\|_2}{\|y' - x\|_2}, \tag{45}$$

by (42). Therefore, instead of calculating $\sin(\widehat{a'ax})$, our plan is to find $\cos(\widehat{xy'y''})$ first. To that end, note that

$$\overline{yy'} \perp (x + \pi_m) \implies \|y' - y''\|_2^2 = \|y'' - y\|_2^2 - \|y' - y\|_2^2, \tag{46}$$

by (42) again. After recalling that $\|x - b\|_2 = \text{rch}(\mathbb{M}') < \|y - b\|_2$, we write that

$$\|y'' - y\|_2^2 = \|x - y\|_2^2 \cdot \sin^2\left(\widehat{yxa}\right) \qquad \left(\overline{yy''} \perp \overline{xa} \text{ from (42)}\right)$$
$$= \|x - y\|_2^2 \cdot \cos^2\left(\widehat{xbb'}\right) \qquad \left(\widehat{yxa} + \widehat{xbb'} = \frac{\pi}{2}\right)$$
$$= \|x - y\|_2^2 \cdot \left(1 - \sin^2\left(\widehat{xbb'}\right)\right)$$
$$= \|x - y\|_2^2 \cdot \left(1 - \frac{\|x - b'\|_2^2}{\text{rch}(\mathbb{M}')^2}\right) \qquad \left(\|x - b\|_2 = \text{rch}(\mathbb{M}')\right)$$
$$\geq \|x - y\|_2^2 \cdot \left[1 - \left(\frac{\|x - y\|_2}{2 \cdot \text{rch}(\mathbb{M}')}\right)^2\right]. \qquad \left(\|x - b\|_2 \leq \|y - b\|_2 \Rightarrow \|x - b'\|_2 \leq \|y - b'\|_2\right)$$

Substituting the expression for $\|y'' - y\|_2$ above back in (46), we arrive at

$$\|y' - y''\|_2^2 \geq \|x - y\|_2^2 \cdot \left[1 - \left(\frac{\|x - y\|_2}{2 \cdot \text{rch}(\mathbb{M}')}\right)^2\right] - \|y' - y\|_2^2.$$

In turn, plugging the expression for $\|y' - y''\|_2$ above back into (45) yields

$$\cos^2\left(\widehat{xy'y''}\right) \geq \frac{\|x - y\|_2^2 \cdot \left[1 - \left(\frac{\|x - y\|_2}{2 \cdot \text{rch}(\mathbb{M}')}\right)^2\right] - \|y' - y\|_2^2}{\|y' - x\|_2^2}.$$

By (44), then, we obtain that

$$
\begin{aligned}
\sin^2\left(\widehat{a'ax}\right) &= \sin^2\left(\widehat{xy'y''}\right) \\
&= 1 - \cos^2\left(\widehat{xy'y''}\right) \\
&\leq 1 - \frac{\|x-y\|_2^2 \cdot \left[1 - \left(\frac{\|x-y\|_2}{2 \cdot \mathrm{rch}(\mathbb{M}')}\right)^2\right] - \|y'-y\|_2^2}{\|y'-x\|_2^2} \\
&= \frac{\|y'-x\|_2^2 + \|y'-y\|_2^2 - \|x-y\|_2^2 \cdot \left[1 - \left(\frac{\|x-y\|_2}{2 \cdot \mathrm{rch}(\mathbb{M}')}\right)^2\right]}{\|y'-x\|_2^2} \\
&= \frac{\|x-y\|_2^2 - \|x-y\|_2^2 \cdot \left[1 - \left(\frac{\|x-y\|_2}{2 \cdot \mathrm{rch}(\mathbb{M}')}\right)^2\right]}{\|y'-x\|_2^2} \qquad \left(\overline{yy'} \perp (x + \pi_m)\right) \\
&= \frac{\|x-y\|_2^2 \cdot \left(\frac{\|x-y\|_2}{2 \cdot \mathrm{rch}(\mathbb{M}')}\right)^2}{\|y'-x\|_2^2}.
\end{aligned}
$$

Plugging the above expression for $\sin(\widehat{a'ax})$ back into (43), we obtain that

$$
\begin{aligned}
r_m &\geq \frac{\frac{l}{\sigma_{\max}}\|x-y\|_2}{2\sin\left(\widehat{a'ax}\right)} \\
&\geq \mathrm{rch}(\mathbb{M}') \cdot \frac{l}{\sigma_{\max}} \cdot \frac{\|y'-x\|_2}{\|x-y\|_2} \\
&= \mathrm{rch}(\mathbb{M}') \cdot \left(\frac{l}{\sigma_{\max}}\right)^2, \qquad \text{(see (40))}
\end{aligned}
$$

where the last line uses a similar argument to that used in (43). Since our choices of x and w were otherwise arbitrary, we conclude that

$$
\mathrm{rch}\left(I_{m,n}\mathbb{M}'\right) = \mathrm{rch}\left(P_m\mathbb{M}'\right) \geq \left(\frac{l}{\sigma_{\max}}\right)^2 \cdot \mathrm{rch}(\mathbb{M}'). \tag{47}
$$

G.2 Second Step

So far, we have computed $\mathrm{rch}(I_{m,n}\mathbb{M}')$. In this section, we complete the argument by computing $\mathrm{rch}(\Sigma' \cdot I_{m,n}\mathbb{M}') = \mathrm{rch}(\Sigma\mathbb{M}') = \mathrm{rch}(\Phi\mathbb{M})$. We do so by studying the linear map Σ'.

As indicated in (39), Σ' is a bi-Lipschitz map on $\mathbb{R}^m$ with constants $\sigma_{\min} \leq \sigma_{\max}$. Therefore, intuitively, it should be clear that Σ' does not substantially distort the geometry of $I_{m,n}\mathbb{M}' \subset \mathbb{R}^m$. Next, we make this notion concrete.

Consider an arbitrary point $I_{m,n}x \in I_{m,n}\mathbb{M}'$. By Definition 1, one can place an m-dimensional (Euclidean) ball of radius $\mathrm{rch}(I_{m,n}\mathbb{M}')$ tangent to $I_{m,n}\mathbb{M}'$ at $I_{m,n}x$ that never intersects $I_{m,n}\mathbb{M}'$. Under Σ', this ball turns into an ellipsoid. The longest principal axis of this ellipsoid does not exceed $\sigma_{\max} \cdot \mathrm{rch}(I_{m,n}\mathbb{M}')$ and its smallest principal axis is no less than $\sigma_{\min} \cdot \mathrm{rch}(I_{m,n}\mathbb{M}')$. The reach of this ellipsoid is never less than $\frac{\sigma_{\min}^2}{\sigma_{\max}} \cdot \mathrm{rch}(I_{m,n}\mathbb{M}')$ as guaranteed by Proposition 2. Therefore,

this ellipsoid itself contains an m-dimensional ball of radius $\frac{\sigma_{\min}^2}{\sigma_{\max}} \cdot \text{rch}(I_{m,n}\mathbb{M}')$ that is tangent at $\Sigma' \cdot I_{m,n}x = \Sigma x$ and never intersects the ellipsoid (and hence the manifold $\Sigma \mathbb{M}'$).

It is possible to choose the center of the original Euclidean ball in any direction normal to $I_{m,n}\mathbb{M}'$ at $I_{m,n}x$; the resulting Euclidean balls (which are tangent to Σx) trace out all normal directions at x. Since the choice of x was otherwise arbitrary, it follows that

$$\text{rch}(\Phi\mathbb{M}) = \text{rch}(\Sigma\mathbb{M}') = \text{rch}\left(\Sigma' \cdot I_{m,n}\mathbb{M}'\right) \geq \frac{\sigma_{\min}^2}{\sigma_{\max}} \cdot \text{rch}(I_{m,n}\mathbb{M}'). \tag{48}$$

When $m = n$, in particular, we continue by noting that

$$\begin{aligned}
\text{rch}(\Phi\mathbb{M}) &\geq \frac{\sigma_{\min}^2}{\sigma_{\max}} \cdot \text{rch}\left(I_{m,n}\mathbb{M}'\right) && \text{(see (48))} \\
&= \frac{\sigma_{\min}^2}{\sigma_{\max}} \cdot \text{rch}(\mathbb{M}') && (I_{m,n} = I_n \text{ is the identity matrix}) \\
&= \frac{\sigma_{\min}^2}{\sigma_{\max}} \cdot \text{rch}(\mathbb{M}) && \text{(see (36))}.
\end{aligned}$$

On the other hand, when $m < n$, we write that

$$\begin{aligned}
\text{rch}(\Phi\mathbb{M}) &\geq \frac{\sigma_{\min}^2}{\sigma_{\max}} \cdot \text{rch}\left(I_{m,n}\mathbb{M}'\right) && \text{(see (48))} \\
&\geq \frac{\sigma_{\min}^2}{\sigma_{\max}} \cdot \left(\frac{l}{\sigma_{\max}}\right)^2 \cdot \text{rch}(\mathbb{M}') && \text{(see (47))} \\
&= \frac{\sigma_{\min}^2}{\sigma_{\max}} \cdot \left(\frac{l}{\sigma_{\max}}\right)^2 \cdot \text{rch}(\mathbb{M}).
\end{aligned}$$

This completes the proof of Theorem 8.