Implicit Manifold Reconstruction

Siu-Wing Cheng†  Man-Kwun Chiu‡

Abstract

Let $\mathcal{M} \subset \mathbb{R}^d$ be a compact, smooth and boundaryless manifold with dimension $m$ and unit reach. We show how to construct a function $\varphi : \mathbb{R}^d \to \mathbb{R}^{d-m}$ from a uniform $(\varepsilon, \kappa)$-sample $P$ of $\mathcal{M}$ that offers several guarantees. Let $Z_\varphi$ denote the zero set of $\varphi$. Let $\hat{\mathcal{M}}$ denote the set of points at distance $\varepsilon$ or less from $\mathcal{M}$. There exists $\varepsilon_0 \in (0, 1)$ that decreases as $d$ increases such that if $\varepsilon \leq \varepsilon_0$, the following guarantees hold. First, $Z_\varphi \cap \hat{\mathcal{M}}$ is a faithful approximation of $\mathcal{M}$ in the sense that $Z_\varphi \cap \hat{\mathcal{M}}$ is homeomorphic to $\mathcal{M}$, the Hausdorff distance between $Z_\varphi \cap \hat{\mathcal{M}}$ and $\mathcal{M}$ is $O(m^{5/2} \varepsilon^2)$, and the normal spaces at nearby points in $Z_\varphi \cap \hat{\mathcal{M}}$ and $\mathcal{M}$ make an angle $O(m^2 \sqrt{\kappa \varepsilon})$. Second, $\varphi$ has local support; in particular, the value of $\varphi$ at a point is affected only by sample points in $P$ that lie within a distance of $O(m \varepsilon)$. Third, we give a projection operator that only uses sample points in $P$ at distance $O(m \varepsilon)$ from the initial point. The projection operator maps any initial point near $P$ onto $Z_\varphi \cap \hat{\mathcal{M}}$ in the limit by repeated applications.

* A preliminary version appears in Proceedings of the 25th Annual ACM-SIAM Symposium on Discrete Algorithms, 2014, 161–173.
† Supported by Research Grants Council, Hong Kong, China (project no. 612109). Department of Computer Science and Engineering, HKUST, Hong Kong.
‡ Supported by JST ERATO Grant Number JPMJER1201, Japan and ERC StG 757609. Institut für Informatik, Freie Universität Berlin, Berlin, Germany.
1 Introduction

Sensory devices and numerical experiments may generate numerous data points in $\mathbb{R}^d$ for some large $d$ due to the large number of attributes of the data that are being monitored. It is often believed that the data points are governed by some hidden processes with fewer controlling parameters, and therefore, the data points may lie in some $m$-dimensional manifold $\mathcal{M}$ for some $m \ll d$. This motivates the study of manifold reconstruction.

In computational geometry, there are several known results that offer provably faithful reconstructions in the sense that the reconstruction is topologically equivalent to $\mathcal{M}$, the Hausdorff distance between the reconstruction and $\mathcal{M}$ decreases as the sampling density increases, and the angular error between the tangent spaces at nearby points in the reconstruction and $\mathcal{M}$ decreases as the sampling density increases. These include the weighted cocone complex by Cheng, Dey and Ramos [13], the weighted witness complex by Boissonnat, Guibas and Oudot [9], and the tangential Delaunay complex by Boissonnat and Ghosh [8]. These reconstructions are $m$-dimensional simplicial complexes with the given sample points as vertices. The corresponding reconstruction algorithms have to deal with the challenging issue of “sliver removal” in high dimensions.

Solutions of partial differential equations on manifolds are required in quite a few areas such as biology [33], image processing [41, 43], weathering [18], and fluid dynamics [36, 37]. The underlying manifold is often specified by a point cloud. It has been reported [31] that local reconstructions of a manifold in the form of zero level sets of local functions are preferred for solving partial differential equations on the manifold. Several numerical methods for solving partial differential equations on level sets have been developed [5, 22, 31, 38].

In this paper, we propose an implicit reconstruction for manifolds with arbitrary codimension in $\mathbb{R}^d$. Let $\mathcal{M}$ be a compact, smooth, and boundaryless manifold with unit reach. Let $P$ be a uniform $(\varepsilon, \kappa)$-sample of $\mathcal{M}$, that is, every point in $\mathcal{M}$ is at distance $\varepsilon$ or less from some point in $P$ and the number of sample points inside any $d$-ball of radius $\varepsilon$ is at most some constant $\kappa$. We assume that the following information is specified in the input: (i) the manifold dimension $m$, (ii) a neighborhood radius $\gamma = 4\varepsilon$, and (iii) approximate tangent spaces at points in $P$ such that the true tangent space at each point in $P$ makes an angle at most $m\gamma$ with the given approximate tangent space at that point. There are many algorithms for estimating the manifold dimension (e.g. [12, 14, 25, 30, 40]). When the sample points satisfy some local uniformity condition (e.g., a constant upper bound on the number of sample points inside any ball of radius $\varepsilon$ centered in $\mathcal{M}$), the neighborhood radius $\gamma$ can be set by measuring the maximum distance from a sample point to its $k$th nearest neighbor for some appropriate $k$. If the sample points are drawn from an independent and identical distribution on $\mathcal{M}$, a recently proposed reach estimator can be used to set $\gamma$ [5]. There are many algorithms for estimating tangent spaces (e.g. [4, 11, 23, 32, 39]), which give an $O(\varepsilon)$ angular error.

We use the conditions of $\gamma = 4\varepsilon$ and angular error at most $m\gamma$ in order to keep the number of unknown parameters small. One may worry about satisfying these two conditions simultaneously, but it is not a concern as we explain below. Suppose that the estimation algorithms return an angular error bound of $c\varepsilon$ for some known constant $c \geq 1$ and a value $\ell$ such that $\varepsilon \leq \ell = O(\varepsilon)$. We can set $\gamma = \max\{4\ell, c\ell\}$. Then, the angular error is at most $c\varepsilon \leq c\ell \leq m\gamma$. Moreover, letting $c' = \max\{\frac{\ell}{\varepsilon}, \frac{c\ell}{\varepsilon}\}$, the input sample can be viewed as a uniform $(\varepsilon', \kappa')$-sample, where $\varepsilon' = c'\varepsilon = \gamma/4$ and $\kappa' = (2c' + 1)^d\kappa$, because a packing argument shows that if any $d$-ball of radius $\varepsilon$ contains at most $\kappa$ sample points, then any $d$-ball of radius $c'\varepsilon$ contains at most $(2c' + 1)^d\kappa$ sample points.

Our main result is a formula for a function $\varphi : \mathbb{R}^d \to \mathbb{R}^{d-m}$ using the $(\varepsilon, \kappa)$-sample $P$ and the neighborhood radius $\gamma$ such that the zero set of $\varphi$ near $\mathcal{M}$ forms a reconstruction of $\mathcal{M}$. Let $Z_\varphi$ denote the zero set of $\varphi$. Let $\mathcal{M}$ denote the set of points at distance $\varepsilon$ or less
from $\mathcal{M}$. We prove that there exists $\varepsilon_0 \in (0, 1)$ that decreases as $d$ increases such that if $\varepsilon \leq \varepsilon_0$, the following guarantees hold. First, $Z_\varphi \cap \hat{\mathcal{M}}$ is a faithful approximation of $\mathcal{M}$ in the sense that $Z_\varphi \cap \hat{\mathcal{M}}$ is homeomorphic to $\mathcal{M}$, the Hausdorff distance between $Z_\varphi \cap \hat{\mathcal{M}}$ and $\mathcal{M}$ is $O(m^{5/2} \gamma^2) = O(m^{5/2} \varepsilon^2)$, and the normal spaces at nearby points in $Z_\varphi \cap \hat{\mathcal{M}}$ and $\mathcal{M}$ make an angle $O(m^2 \sqrt{\kappa \gamma}) = O(m^2 \sqrt{\kappa \varepsilon})$. Second, $\varphi$ has local support; in particular, the value of $\varphi$ at a point is affected only by sample points in $P$ that lie within a distance of $m \gamma$. Third, we give a projection operator that only uses sample points in $P$ at a distance $m \gamma$ from the initial point. The projection operator maps any initial point near $P$ onto $Z_\varphi \cap \hat{\mathcal{M}}$ in the limit by repeated applications.

Implicit surfaces in three dimensions have been extensively studied, particularly in computer graphics and solid modeling (e.g. [2, 10, 26, 29]). Two functions have been defined in [17, 28] and shown to give faithful reconstruction of the underlying surface in three dimensions. In $\mathbb{R}^d$, a function is defined in [7] and shown to give faithful reconstruction of $(d-1)$-dimensional manifold. There seems to be no prior work with provable guarantees on implicit reconstructions of manifolds in $\mathbb{R}^d$ with codimension less than $d - 1$. In the computer graphics community, similar functions have been proposed as projection operators by Adamson and Alexa [1] for designing a complex of surface patches connected via vertices and curves in three dimensions. Each surface patch is the set of stationary points under a projection operator. For each surface patch, some input points with prescribed tangent spaces are given for defining the corresponding projection operator, but these input points need not form an $\varepsilon$-sample of the resulting surface patch. It is discussed how to generalize the framework to $\mathbb{R}^d$ for a complex of submanifolds. However, no mathematical guarantee was provided in [1] for $\mathbb{R}^3$ or $\mathbb{R}^d$.

Although the zero set of our function $\varphi$ has a subset near $\mathcal{M}$ that is a faithful reconstruction, $\varphi$ should not be confused to be a smooth implicit function as in the Implicit Function Theorem. If the normal bundle of $\mathcal{M}$ is topologically non-trivial, one cannot define a smooth implicit function whose zero set is a faithful reconstruction of $\mathcal{M}$.

We provide the definition of our function $\varphi$ in the next section. Afterwards, we give the proofs of the theoretical guarantees.

## 2 Function formulation

We use lowercase and uppercase letters in math font to denote column vectors and matrices, respectively. A point is always specified as a column vector. Given a matrix $K$, we use $\text{col}(K)$ to denote the column space of $K$. We call the unit eigenvectors of a square matrix corresponding to the $k$ largest (resp. smallest) eigenvalues the $k$ most dominant (resp. least dominant) unit eigenvectors.

Recall that $\gamma = 4 \varepsilon$ is the input neighborhood radius. We will make use of a weight function $\omega : \mathbb{R}^d \to \mathbb{R}$ defined as

$$\omega(x, p) = \frac{h(||x - p||)}{\sum_{q \in P} h(||x - q||)},$$

where

$$h(s) = \begin{cases} \left(1 - \frac{s}{m \gamma}\right)^{2m} \left(\frac{2s}{\gamma} + 1\right), & \text{if } s \in [0, m \gamma], \\ 0, & \text{if } s > m \gamma. \end{cases}$$

Note that $h$ is differentiable in $(0, \infty)$ and $h'(s) = 0$ for $s \geq m \gamma$. This weight function is inspired by the Wendland functions [12].
Since approximate tangent spaces at the sample points are specified in the input, we can assume that a \( d \times m \) matrix \( T_p \) is given for each \( p \in P \) such that \( T_p \) has orthogonal unit columns and \( \text{col}(T_p) \) is the approximate tangent space at \( p \). Define the following matrix and vector space for each point \( x \in \mathbb{R}^d \):

\[
C_x = \sum_{p \in P} \omega(x, p) \cdot T_p \cdot T_p^t, \\
L_x = \text{space spanned by the } (d-m) \text{ least dominant unit eigenvectors of } C_x.
\]

The \((d-m)\) least dominant unit eigenvectors of \( T_p \cdot T_p^t \) span an approximate normal space of \( \mathcal{M} \) at \( p \). So \( L_x \) is the “weighted average” of the approximate normal spaces at the sample points near \( x \).

Define a class \( \Phi \) of functions \( \varphi : \mathbb{R}^d \to \mathbb{R}^{d-m} \) as follows:

\[
\Phi = \left\{ \varphi : \varphi(x) = \sum_{p \in P} \omega(x, p) \cdot B_{\varphi,x}^t \cdot (x - p) \right\}, \quad \text{where } B_{\varphi,x} \text{ is any } d \times (d-m) \text{ matrix}
\]

with linearly independent columns such that \( \text{col}(B_{\varphi,x}) = L_x \).

Evaluating \( \varphi(x) \) requires only the sample points at distance \( m\gamma \) or less from \( x \), and \( \omega \) gives more weight to sample points nearer \( x \). Different choices of \( B_{\varphi,x} \) at each \( x \in \mathbb{R}^d \) give rise to different functions in \( \Phi \). A natural choice is a \( d \times (d-m) \) matrix consisting of \( d-m \) orthogonal unit vectors that span \( L_x \). We denote the corresponding function in \( \Phi \) by \( \varphi \) and so

\[
\varphi(x) = \sum_{p \in P} \omega(x, p) \cdot B_{\varphi,x}^t \cdot (x - p).
\]

We will show that every function in \( \Phi \) has the same zero set. \( Z_\varphi \) as a whole is not a good reconstruction of \( \mathcal{M} \). Indeed, by definition, \( \varphi(x) = 0 \) for any \( x \in \mathbb{R}^d \) at distance \( m\gamma \) or more from \( \mathcal{M} \). We focus on the subset \( \hat{\mathcal{M}} \) of \( \mathbb{R}^d \) (i.e., the set of points at distance \( \varepsilon \) or less from \( \mathcal{M} \)). We show that \( Z_\varphi \cap \hat{\mathcal{M}} \) is a faithful reconstruction of \( \mathcal{M} \).

3 Preliminaries

3.1 Definitions

Given a matrix or vector, the corresponding italic lowercase letter with subscripts denotes an element. For example, \( k_{ij} \) denotes the \((i,j)\) entry of a matrix \( K \) and \( v_i \) denotes the \(i\)-th coordinate of a vector \( v \). We use \( I_j \) to denote a \( j \times j \) identity matrix and \( 0_{i,j} \) an \( i \times j \) zero matrix. The 2-norms of \( v \) and \( K \) are \( \|v\| = \left(\sum_i v_i^2\right)^{1/2} \) and \( \|K\| = \max \{ \|Kv\| : \|v\| = 1 \} \).

We use \( B(x, r) \) to denote the geometric \( d \) ball centered at \( x \) with radius \( r \). We use \( \angle(v, E) \) to denote the angle between a vector \( v \) and its projection in an affine subspace \( E \). The angle \( \angle(E, F) \) between two affine subspaces \( E \) and \( F \), where \( \dim(E) \leq \dim(F) \), is \( \max \{ \angle(v, F) : \text{vector } v \in E \} \).

The normal space of \( \mathcal{M} \) at a point \( z \), denoted \( N_z \), is the linear subspace of \( \mathbb{R}^d \) that comprises of all vectors normal to \( \mathcal{M} \) at \( z \). Each vector in \( N_z \) has \( d \) coordinates although \( N_z \) has dimension \( d-m \). The tangent space of \( \mathcal{M} \) at \( z \), denoted \( T_z \), is the orthogonal complement of \( N_z \).

The medial axis of \( \mathcal{M} \) is the closure of the set of points in \( \mathbb{R}^d \) that have two or more closest points in \( \mathcal{M} \). The local feature size at a point \( z \in \mathcal{M} \) is the distance from \( z \) to the medial axis. We assume that the reach or minimum local feature size of \( \mathcal{M} \) is 1.

Let \( \nu \) denote the nearest point map. That is, for every point \( x \) that does not belong to the medial axis of \( \mathcal{M} \), \( \nu(x) \) is the point in \( \mathcal{M} \) nearest to \( x \).
3.2 Basic results

We need the following basic results on \( \varepsilon \)-sampling theory, matrices, and linear subspaces.

**Lemma 3.1** \((\text{[13, 23]})\)

(i) For all \( y, z \in M \), if \( \|y - z\| \leq \xi \) for some \( \xi < 1 \), \( y \) is at distance \( \xi^2/2 \) or less from \( z + T_\varepsilon \).

(ii) For all \( y, z \in M \), if \( \|y - z\| \leq \xi \) for a small enough \( \xi \), then \( \angle(N_y, N_z) \leq 4\xi \).

**Lemma 3.2** Let \( P \) be a uniform \((\varepsilon, \kappa)\)-sample of \( M \). For any \( x \in \mathbb{R}^d \) and any \( t \in \left[ 1, \frac{1}{\sqrt{2d}} \right] \), \( |P \cap B(x, t\varepsilon)| \leq (4t + 1)^m \kappa \).

**Proof.** We first show an upper bound on the minimum number of balls with radii \( \varepsilon \) such that their union contains \( M \cap B(x, t\varepsilon) \), which will imply the desired result. We pick a maximal set \( S \) of points in \( M \cap B(x, t\varepsilon) \) such that any two of them are at distance \( \varepsilon \) or more apart. It implies that \( M \cap B(x, t\varepsilon) \subseteq \cup_{z \in S} B(z, \varepsilon) \). Otherwise there exists a point \( z \in M \cap B(x, t\varepsilon) \) such that the distance between \( z \) and \( S \) is larger than \( \varepsilon \), then we can get a larger set by adding \( z \) to \( S \), a contradiction to the definition of \( S \). Let \( S' \) denote the projection of \( S \) onto \( x + T_{\nu(x)} \).

By Lemma 3.1(i), the distance between any two points in \( S' \) is at least \( \varepsilon - (t\varepsilon)^2 \geq \varepsilon/2 \) when \( t \leq \frac{1}{\sqrt{2d}} \). Thus, any two balls centered at points in \( S' \) with radius \( \varepsilon/4 \) are interior-disjoint. Since the projection of \( M \cap B(x, t\varepsilon) \) into \( x + T_{\nu(x)} \) is contained in \( (x + T_{\nu(x)}) \cap B(x, t\varepsilon) \), \( |S'| \) is no larger than the size of a maximal packing of interior-disjoint \( m \)-dimensional balls with radius \( \varepsilon/4 \) in \( (x + T_{\nu(x)}) \cap B(x, t\varepsilon + \varepsilon/4) \), which is at most the volume of \( (x + T_{\nu(x)}) \cap B(x, t\varepsilon + \varepsilon/4) \) divided by \( (\varepsilon/4)^m V_m \), where \( V_m \) is the volume of a unit \( m \)-ball. Thus, \( |S| = |S'| \leq \frac{(t\varepsilon + \varepsilon/4)^m}{(\varepsilon/4)^m} = (4t + 1)^m \).

Then, \( |P \cap B(x, t\varepsilon)| \leq (4t + 1)^m \kappa \) by the definition of uniform \((\varepsilon, \kappa)\)-sampling. \( \square \)

Partition a square matrix \( K \) into blocks:

\[
\begin{pmatrix}
K_{i1} & \cdots & K_{ir} \\
\vdots & \ddots & \vdots \\
K_{r1} & \cdots & K_{rr}
\end{pmatrix}
\]

The matrices \( K_{ii} \) are square, but they may have different dimensions. For \( j \neq i \), \( K_{ij} \) may be square or rectangular. For any \( i, j, k \in [1, r] \), \( K_{ik} \) and \( K_{jk} \) have the same number of columns and \( K_{ij} \) and \( K_{ik} \) have the same number of rows. Each row of blocks \( (K_{i1} \cdots K_{ir}) \) defines a **generalized gershgorin set** \( G_i \) as follows. Let \( n_i \) be the dimension of \( K_{ii} \).

\[
G_i = \left\{ \mu \in \mathbb{R} : \frac{1}{\|(K_{ii} - \mu I_{n_i})^{-1}\|} \leq \sum_{j \neq i} \|K_{ij}\| \right\}
\]

It follows that the numbers in \( G_i \) are at least the smallest eigenvalue of \( K_{ii} \) minus \( \sum_{i \neq j} \|K_{ij}\| \) and at most the maximum eigenvalue of \( K_{ii} \) plus \( \sum_{i \neq j} \|K_{ij}\| \). The eigenvalues of \( K_{ii} \) are defined to be in \( G_i \) using a continuity argument \([20]\).

**Lemma 3.3** \((\text{[20]})\) Consider any partition of a square matrix \( K \) into blocks. Every eigenvalue of \( K \) lies in some generalized gershgorin set \( G_i \) with respect to this partition. Moreover, if a generalized gershgorin set \( G_i \) is disjoint from the union of the other generalized gershgorin sets, then \( G_i \) contains exactly \( n_i \) eigenvalues of \( K \), where \( n_i \) is the dimension of \( K_{ii} \).

**Lemma 3.4** \((\text{[24]})\) Let \( (U, V) \) be a \( d \times d \) orthogonal matrix, where \( U \) is \( d \times r \) and \( V \) is \( d \times (d-r) \). Let \( K \) be a \( d \times r \) matrix with orthogonal unit columns. Then, \( \angle(\text{col}(U), \text{col}(K)) = \arcsin(\|V^T \cdot K\|) \).
Lemma 3.5 ([19] Lemma 1.1) Let $M_1$ be an $s \times s$ real symmetric matrix with eigenvalues $\lambda_1, \ldots, \lambda_s$ in an arbitrary order. Let $v_i$ denote a unit eigenvector of $M_1$ corresponding to $\lambda_i$. If $M_1 + M_2$ is a real symmetric matrix, $\sigma$ is an eigenvalue of $M_1 + M_2$, and $e$ is a unit eigenvector of $M_1 + M_2$ corresponding to $\sigma$, then for every $r \in [1, s-1]$, the angle between $e$ and the space spanned by $\{v_1, \ldots, v_r\}$ is at most $\arcsin\left(\frac{\|M_2\|}{\min_{i \in [r+1, s]} |\lambda_i - \sigma|}\right)$.

Lemma 3.6 Let $V$ and $W$ be two linear subspaces of the same dimension $k$ in $\mathbb{R}^d$ such that $\theta = \angle(V, W) < \pi/2$.

(i) For each orthonormal basis $\{v_1, \ldots, v_k\}$ of $V$, there exists an orthonormal basis $\{w_1, \ldots, w_k\}$ of $W$ such that $\angle(v_i, w_i) \leq \theta$ for $i \in [1, k]$ and $\angle(v_i, w_j - v_j) \in \left[\frac{\pi}{2} - \theta, \frac{\pi}{2} + \theta\right]$ for $i, j \in [1, k]$.

(ii) If $k > d/2$, then there exist orthonormal bases $\{v_1, \ldots, v_k\}$ and $\{w_1, \ldots, w_k\}$ of $V$ and $W$, respectively, such that $v_i = w_i$ for $i \in [1, 2k - d]$, $\angle(v_i, w_i) \leq \theta$ for $i \in [1, k]$, and $\angle(v_i, w_j - v_j) \in \left[\frac{\pi}{2} - \theta, \frac{\pi}{2} + \theta\right]$ for $i, j \in [1, k]$. Hence, for any distinct $i$ and $j$, if $i \in [1, 2k - d]$ or $j \in [1, 2k - d]$, then $v_i \perp w_j$.

Proof. We make use of principal angles and principal vectors [6] [21] [35]. Pick unit vectors $a_1 \in V$ and $b_1 \in W$ that minimizes $\angle(a_1, b_1)$. For $i \in [2, k]$, pick unit vectors $a_i \in V$ and $b_i \in W$ that minimizes $\angle(a_i, b_i)$ subject to $a_i \perp a_j$ and $b_i \perp b_j$ for all $j \in [1, i - 1]$. The angles $\angle(a_1, b_1), \ldots, \angle(a_k, b_k)$ are called the principal angles. The vectors $\{a_1, \ldots, a_k\}$ and $\{b_1, \ldots, b_k\}$ are called principal vectors. Note that $\{a_1, \ldots, a_k\}$ and $\{b_1, \ldots, b_k\}$ are orthonormal bases of $V$ and $W$, respectively. The alternative definition of principal angles in [21] implies that for $i \in [1, k]$, $\theta_i \leq \theta = \angle(V, W)$. It is also known that $a_i \perp b_j$ for $i \neq j$ [6] [21].

Consider (i). Given an orthonormal basis $\{v_1, \ldots, v_k\}$ of $V$, for each $i \in [1, k]$, $v_i = \sum_{r=1}^{k} c_{ir} a_r$ for some real coefficients $c_{ir}$’s. Correspondingly, define $w_i = \sum_{r=1}^{k} c_{ir} b_r$. Note that $\|w_i\| = (\sum_{r=1}^{k} c_{ir}^2)^{1/2} = \|v_i\| = 1$. Also, for $i \neq j$, $w_i w_j = \sum_{r=1}^{k} c_{ir} c_{jr} = v_i v_j = 0$. So $\{w_1, \ldots, w_k\}$ is an orthonormal basis of $W$.

For $i \in [1, k]$, $v_i w_i = \sum_{r=1}^{k} c_{ir}^2 a_r^2 b_r \geq \cos \theta$ because $\angle(a_r, b_r) \leq \theta$ and $\sum_{r=1}^{k} c_{ir}^2 = \|v_i\| = 1$. It follows that $\angle(v_i, w_i) \leq \theta$. Since $v_i$ and $w_i$ are unit vectors and $\angle(v_i, w_i) \leq \theta$, $v_i + w_i$ is an angle bisection between $v_i$ and $w_i$. Hence, $\angle(v_i, v_i + w_i) \leq \theta/2$. It suffices to show that for any $i, j \in [1, k]$, $v_i + w_i \perp w_j - v_j$, which then implies that $\left|\frac{\pi}{2} - \angle(v_i, w_j - v_j)\right| \leq \angle(v_i, v_i + w_i) \leq \theta/2$, completing the proof of (i). To see that $v_i + w_i \perp w_j - v_j$, we check $(v_i + w_i)^T (w_j - v_j) = \sum_{r=1}^{k} (c_{ir} a_r + c_{jr} b_r)^T (c_{jr} b_r - c_{ir} a_r)$. Recall that $a_r$ and $b_r$ are unit vectors and for $r \neq s$, $a_r \perp a_s$, $b_r \perp b_s$, and $a_r \perp b_s$. Therefore, $\sum_{r=1}^{k} (c_{ir} a_r + c_{jr} b_r)^T (c_{jr} b_r - c_{ir} a_r) = 0$.

Consider (ii). Since $k > d/2$, the dimension of $V \cap W$ is at least $2k - d$. Pick an arbitrary subset $\{u_1, \ldots, u_{2k-d}\}$ of the orthonormal basis of $V \cap W$. Set $v_i = w_i = u_i$ for $i \in [1, 2k - d]$. Complete $\{v_1, \ldots, v_{2k-d}\}$ arbitrarily to an orthonormal basis $\{v_1, \ldots, v_k\}$ of $V$. Then, we construct $w_j$ as the same way as in (i) for $j \in [2k - d + 1, k]$.

Lemma 3.7 Let $E_1$ and $E_2$ be two $k$-dimensional linear subspaces. Let $\{u_1, \ldots, u_k\}$ be a basis of $E_1$ consisting of unit vectors such that for any distinct $i, j \in [1, k]$, $\angle(u_i, u_j) \in [\pi/2 - \phi, \pi/2 + \phi]$ for some $\phi \in [0, \arcsin(\frac{1}{k})]$. For any $\theta \in (0, \arcsin\left(\sqrt{\frac{1}{k}} \sin \phi\right))$, if $\angle(u_i, E_2) \leq \theta$ for all $i \in [1, k]$, then $\angle(E_1, E_2) \leq \arctan\left(\sqrt{\frac{\pi\sin \theta}{1 - k \sin^2 \theta - k \sin \phi}}\right)$.

Proof. Orient space such that $E_2$ is spanned by the first $k$ coordinate axes of $\mathbb{R}^d$. Then, for all $i \in [1, k]$, we can write

$$u_i = \begin{pmatrix} v_i \\ w_i \end{pmatrix},$$
where $v_i$ consists of the first $k$ coordinates and $w_i$ consists of the remaining $d - k$ coordinates. Note that

$$
\begin{pmatrix} 0_{k,1} \\ w_i \end{pmatrix} \perp E_2 \quad \text{and} \quad \begin{pmatrix} v_i \\ 0_{d-k,1} \end{pmatrix} \in E_2.
$$

Since $\angle(u_1, E_2) \leq \theta$ by assumption, we have $\|w_i\| \leq \sin \theta$. As a result, $\|v_i\| \in [\cos \theta, 1]$. For any $i \neq j$, we have

$$
(v_i^t \cdot w_i^t) \cdot (v_j^t \cdot w_j^t) \in \left[ \cos \frac{\pi}{2} + \phi, \cos \left( \frac{\pi}{2} - \phi \right) \right]
$$

$$
\Rightarrow v_i^t \cdot v_j + w_i^t \cdot w_j \in [-\sin \phi, \sin \phi]
$$

$$
\Rightarrow |v_i^t \cdot v_j| \leq \|w_i\| \cdot \|w_j\| + \sin \phi \leq \sin^2 \theta + \sin \phi.
$$

Let $n$ be a vector in $E_1$ that makes the angle $\angle(E_1, E_2)$ with $E_2$. By flipping the orientation of any $u_i$’s if necessary, we can ensure that $n$ is a convex combination of \{u_1, \ldots, u_k\}, i.e.,

$$
n = \sum_{i=1}^k \lambda_i \begin{pmatrix} v_i \\ w_i \end{pmatrix}
$$

for some $\lambda_i$’s in $[0, 1]$ such that $\sum_{i=1}^k \lambda_i = 1$. Note that flipping the orientation of any $u_i$ preserves the angle $\angle(u_i, E_2)$ and the fact that for any distinct $i, j \in [1, k], \angle(u_i, u_j) \in [\pi/2 - \phi, \pi/2 + \phi]$. Hence,

$$
\angle(E_1, E_2) = \arctan \left( \frac{\|\sum_{i=1}^k \lambda_i w_i\|}{\|\sum_{i=1}^k \lambda_i v_i\|} \right)
\leq \arctan \left( \frac{\sum_{i=1}^k \lambda_i \|w_i\|}{\sqrt{\sum_{i=1}^k \sum_{j=1}^k \lambda_i \lambda_j \cdot v_i^t \cdot v_j}} \right)
$$

$$
\leq \arctan \left( \frac{\sin \theta}{\sqrt{\cos^2 \theta \sum_{i=1}^k \lambda_i^2 - (\sin^2 \theta + \sin \phi) \sum_{i \neq j} \lambda_i \lambda_j}} \right)
$$

$$
= \arctan \left( \frac{\sin \theta}{\sqrt{\sum_{i=1}^k \lambda_i^2 - (\sin^2 \theta + \sin \phi) \left( \sum_{i=1}^k \lambda_i \right)^2}} \right)
$$

$$
\leq \arctan \left( \frac{\sqrt{k} \sin \theta}{\sqrt{1 - k \sin^2 \theta - k \sin \phi}} \right).
$$

The last step uses the fact that $\sum_{i=1}^k \lambda_i^2$ is minimized when $\lambda_i = 1/k$ for all $i$. \(\blacksquare\)

### 4 Accuracy of $L_\times$

The main result of this section is Lemma 4.2 below: for every point $z \in M$ and every point $x$ near $z$, $N_z$ is approximated by $L_\times$. We need the following technical result. Recall that $\nu$ is the nearest point map.

**Lemma 4.1** Let $x$ be a point at distance $2\varepsilon$ or less from $M$. Assume a coordinate frame such that the columns of $(I_{m} \ 0_{d-m,m})$ form an orthonormal basis of $T_{\nu(x)}$. Partition $C_x$ into

$$
\begin{pmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{pmatrix},
$$

where $C_{11}$ is $m \times m$, $C_{12}$ is $m \times (d-m)$, $C_{21}$ is $(d-m) \times m$, and $C_{22}$ is $(d-m) \times (d-m)$. Then, $\|C_{12}\|$ and $\|C_{21}\|$ are $O(m\gamma)$, $\|C_{22}\|$ is $O(m^2\gamma^2)$, and the smallest eigenvalue of $C_{11}$ is at least $1 - O(m^2\gamma^2)$.
Proof. Consider any sample point \( p \in P \). Partition \( T_p \) into \( \begin{pmatrix} Y_p \\ Z_p \end{pmatrix} \), where \( Y_p \) is \( m \times m \) and \( Z_p \) is \((d - m) \times m \). For all \( p \in P \cap B(x, m\gamma) \),

\[
\|p - \nu(x)\| \leq \|p - x\| + \|x - \nu(x)\| \leq m\gamma + 2\varepsilon < (m + 1)\gamma.
\]

Then, \( \angle(T_p, T_{\nu(x)}) \leq 4(m + 1)\gamma \) by Lemma 3.1(ii).

Since \( \begin{pmatrix} I_m \\ 0_{d-m,m} \end{pmatrix} \) and \( \begin{pmatrix} 0_{m,d-m} \\ I_{d-m} \end{pmatrix} \) form a \( d \times d \) orthogonal matrix, we obtain

\[
\arcsin(\|Z_p\|) = \arcsin(\|(0_{d-m,m} \ I_{d-m}) \cdot T_p\|)
\]

\[
= \angle(T_{\nu(x)}, \text{col}(T_p)) \quad (: \text{Lemma 3.1})
\]

\[
\leq \angle(T_p, T_{\nu(x)}) + \angle(T_p, \text{col}(T_p))
\]

\[
\leq 4(m + 1)\gamma + m\gamma.
\]

(We use the assumption that the input approximate tangent spaces have angular errors at most \( m\gamma \). Although an angular error of \( O(m\gamma) \) also works, an exact bound of \( m\gamma \) makes explicit the input requirement for constructing the formula of \( \varphi \)). Hence, we have

\[
\forall p \in P \cap B(x, m\gamma), \quad \|Z_p\| = O(m\gamma). \tag{1}
\]

Because \( \omega(x, p) \) vanishes for all \( p \notin B(x, m\gamma) \), \( C_{12} = \sum_{p \in P \cap B(x, m\gamma)} \omega(x, p) \cdot Y_p \cdot Z_p^\dagger \). Since the columns in \( T_p \) have unit 2-norm, we get \( \|Y_p\| \leq 1 \). Thus,

\[
\|C_{12}\| = \left\| \sum_{p \in P \cap B(x, m\gamma)} \omega(x, p) \cdot Y_p \cdot Z_p^\dagger \right\| \leq \sum_{p \in P \cap B(x, m\gamma)} \omega(x, p) \cdot \|Z_p\| = O(m\gamma).
\]

Similarly,

\[
\|C_{21}\| = \left\| \sum_{p \in P \cap B(x, m\gamma)} \omega(x, p) \cdot Z_p \cdot Y_p^\dagger \right\| \leq \sum_{p \in P \cap B(x, m\gamma)} \omega(x, p) \cdot \|Z_p\| = O(m\gamma),
\]

\[
\|C_{22}\| = \left\| \sum_{p \in P \cap B(x, m\gamma)} \omega(x, p) \cdot Z_p \cdot Z_p^\dagger \right\| \leq \sum_{p \in P \cap B(x, m\gamma)} \omega(x, p) \cdot \|Z_p\|^2 = O(m^2\gamma^2).
\]

Since \( T_p^\dagger \cdot T_p = Y_p^\dagger \cdot Y_p + Z_p^\dagger \cdot Z_p \), the minimum eigenvalue of \( Y_p^\dagger \cdot Y_p \) is at least the minimum eigenvalue of \( T_p^\dagger \cdot T_p \) minus \( \|Z_p^\dagger \cdot Z_p\| \). Therefore,

\[
\text{minimum eigenvalue of } Y_p^\dagger \cdot Y_p \geq 1 - O(m^2\gamma^2). \tag{2}
\]

\( Y_p \cdot Y_p^\dagger \) has the same eigenvalues as \( Y_p^\dagger \cdot Y_p \). The smallest eigenvalue of a real symmetric matrix \( M \) is \( \min_{\nu \neq 0} (\nu^\dagger \cdot M \cdot \nu) / \|\nu\|^2 \). Then, using the relation \( C_{11} = \sum_{p \in P \cap B(x, m\gamma)} \omega(x, p) \cdot Y_p \cdot Y_p^\dagger \), we conclude that the smallest eigenvalue of \( C_{11} \) is at least the sum of the smallest eigenvalues of \( \omega(x, p) \cdot Y_p \cdot Y_p^\dagger \). This sum is at least \( 1 - O(m^2\gamma^2) \) by (2).

We are ready to show that the angle between \( L_x \) and any nearby normal space of \( \mathcal{M} \) is \( O(m\sqrt{m\gamma}) \).

**Lemma 4.2** For every point \( z \in \mathcal{M} \) and every point \( x \in B(z, 2\varepsilon) \), \( \angle(L_x, N_z) = O(m\sqrt{m\gamma}) \).
Proof. Adopt a coordinate frame such that the columns of \( \begin{pmatrix} I_m \\ 0_{d-m,m} \end{pmatrix} \) form an orthonormal basis of \( T_{\nu(x)} \). Let \( A_x \) be the \( d \times m \) matrix whose columns are the \( m \) most dominant unit eigenvectors of \( C_x \). Thus, \( \text{col}(A_x) \) is the orthogonal complement of \( L_x \). Let \( e = \begin{pmatrix} v \\ w \end{pmatrix} \) be any column vector of \( A_x \), where \( v \) consists of the first \( m \) coordinates and \( w \) consists of the last \( d - m \) coordinates. Then, \( \angle(e, T_{\nu(x)}) = \arctan(||w||/||v||) \).

We show that \( \angle(e, T_{\nu(x)}) = O(m\gamma) \). Partition \( C_x \) into \( \begin{pmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{pmatrix} \), where \( C_{11} \) is \( m \times m \), \( C_{12} \) is \( m \times (d - m) \), \( C_{21} \) is \( (d - m) \times m \), and \( C_{22} \) is \( (d - m) \times (d - m) \). Let \( \sigma \) be the eigenvalue of \( C_x \) corresponding to \( e \). Then,

\[
C_x e = \begin{pmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{pmatrix} \begin{pmatrix} v \\ w \end{pmatrix} = \sigma \begin{pmatrix} v \\ w \end{pmatrix},
\]

which implies that

\[
||w|| = ||(\sigma I_{d-m} - C_{22})^{-1}C_{21}v|| \leq ||(\sigma I_{d-m} - C_{22})^{-1}|| \cdot ||C_{21}||.
\]

Following the definition of generalized gershgorin sets (Section 3), define

\[
G_1 = \left\{ \mu \in \mathbb{R} : \frac{1}{||(C_{11} - \mu I_m)^{-1}||} \leq ||C_{12}|| \right\}, \quad G_2 = \left\{ \mu \in \mathbb{R} : \frac{1}{||(C_{22} - \mu I_{d-m})^{-1}||} \leq ||C_{21}|| \right\}.
\]

The numbers in \( G_1 \) are at least the minimum eigenvalue of \( C_{11} \) minus \( ||C_{12}|| \), which is at least \( 1 - O(m\gamma + m^2\gamma^2) \) by Lemma 4.1. The numbers in \( G_2 \) are at most \( ||C_{22}|| + ||C_{21}|| = O(m\gamma + m^2\gamma^2) \) by Lemma 4.1. Since every number in \( G_1 \) is greater than any number in \( G_2 \), by Lemma 3.3 \( G_1 \) contains the \( m \) largest eigenvalues of \( C_x \). Thus, \( \sigma \) belongs to \( G_1 \) and \( \sigma \geq 1 - O(m\gamma + m^2\gamma^2) \) which is asymptotically greater than \( ||C_{22}|| = O(m^2\gamma^2) \) (Lemma 4.1). Therefore,

\[
||((\sigma I_{d-m} - C_{22})^{-1}|| \leq \frac{1}{1 - O(m\gamma + m^2\gamma^2)}.
\]

By Lemma 4.1 \( ||C_{21}|| = O(m\gamma) \), and therefore,

\[
||w|| \leq ||((\sigma I_{d-m} - C_{22})^{-1}|| \cdot ||C_{21}|| \leq \frac{O(m\gamma)}{1 - O(m\gamma + m^2\gamma^2)} = O(m\gamma).
\]

As a result, \( 1 \geq ||v|| \geq 1 - ||w|| \geq 1 - O(m\gamma) \). Thus, \( \angle(e, T_{\nu(x)}) = \arctan(||w||/||v||) = O(m\gamma) \).

Since \( e \) is any column vector of \( A_x \), the angle bound in the previous paragraph applies to all column vectors of \( A_x \). We can apply Lemma 3.7 with \( E_1 = \text{col}(A_x) \), \( E_2 = T_{\nu(x)} \), \( \{u_1,\ldots,u_m\} \) equal to the columns of \( A_x \), \( \phi = 0 \), \( k = m \), and \( \theta \) equal to the \( O(m\gamma) \) bound on \( \angle(e, T_{\nu(x)}) \). Then,

\[
\angle(\text{col}(A_x), T_{\nu(x)}) \leq \arctan \left( \frac{O(m\sqrt{m\gamma})}{\sqrt{1 - O(m^3\gamma^2)}} \right) = O(m\sqrt{m\gamma}).
\]

Since \( ||\nu(x) - z|| \leq ||x - \nu(x)|| + ||x - z|| \leq 4\varepsilon \), Lemma 3.1(ii) implies that \( \angle(T_{\nu(x)}, T_z) \leq 16\varepsilon \). Hence,

\[
\angle(L_x, N_z) = \angle(\text{col}(A_x), T_z) \leq \angle(\text{col}(A_x), T_{\nu(x)}) + \angle(T_{\nu(x)}, T_z) = O(m\sqrt{m\gamma}).
\]
5 Projection into $L_x$

For every point $z \in \mathcal{M}$ and every unit vector $n \in N_z$, we want to bound the instantaneous change in the normalized projection of $n$ in $L_x$ as $x$ moves. If we view the projection as a map $f$, this is equivalent to analyzing the Jacobian of $f$ which is given in Lemmas 5.6 and 5.7 below.

To this end, some technical results are needed. First, we need to study the variation of $C_x$ as $x$ moves (Lemma 5.1). Second, we need to bound the turn of $L_x$ if $x$ moves slightly (Lemma 5.4).

Let $\delta_k > 0$ denote an arbitrarily small change in the coordinate $x_k$ of $x$. Define

$$\Delta h(||x-p||) = \frac{\partial h(||x-p||)}{\partial x_k} \cdot \delta_k.$$ 

For simplicity, we omit the dependence of $\Delta h(||x-p||)$ on $k$ in the notation.

**Lemma 5.1** Let $x$ be a point at distance $2\varepsilon$ or less from $\mathcal{M}$. Assume a coordinate frame such that the columns of $\begin{pmatrix} 1_m & 0_{d-m,m} \end{pmatrix}$ form an orthonormal basis of $T_x(x)$. Define the $d \times d$ matrix

$$\Delta C_x = \left( \frac{\partial c_{ij}}{\partial x_k} \cdot \delta_k \right),$$

where $c_{ij}$ is the $(i,j)$ entry of $C_x$. The following properties hold when $\delta_k$ is small enough.

1. $\|\Delta C_x\| \leq \frac{O(m\gamma) \cdot \sum_{p \in P} |\Delta h(||x-p||)|}{\sum_{p \in P} h(||x-p||)}$.

2. The $m$ largest eigenvalues of $C_x + \Delta C_x$ are at least $1 - O(m\gamma) - \frac{O(m\gamma) \cdot \sum_{p \in P} |\Delta h(||x-p||)|}{\sum_{p \in P} h(||x-p||)}$.

**Proof.** Using standard calculus, we obtain

$$\Delta C_x = \frac{1}{(\sum_{p \in P} h(||x-p||))^2} \left( \sum_{p,q \in P} h(||x-p||) \cdot \Delta h(||x-q||) \cdot (T_q \cdot T_q^t - T_p \cdot T_p^t) \right).$$

Partition $C_x$ and $\Delta C_x$ as follows:

$$C_x = \begin{pmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{pmatrix}, \quad \Delta C_x = \begin{pmatrix} \Delta C_{11} & \Delta C_{12} \\ \Delta C_{21} & \Delta C_{22} \end{pmatrix},$$

where $C_{11}$ and $\Delta C_{11}$ are $m \times m$, $C_{12}$ and $\Delta C_{12}$ are $m \times (d-m)$, $C_{21}$ and $\Delta C_{21}$ are $(d-m) \times m$, and $C_{22}$ and $\Delta C_{22}$ are $(d-m) \times (d-m)$.

For every sample point $p \in P$, partition $T_p$ into $T_p = \begin{pmatrix} Y_p \\ Z_p \end{pmatrix}$, where $Y_p$ is an $m \times m$ matrix and $Z_p$ is a $(d-m) \times m$ matrix.

By [1] and [2], for every sample point $p \in P \cap B(x, m\gamma)$, $\|Z_p\| = O(m\gamma)$ and the eigenvalues of $Y_p \cdot Y_p^t$ are at least $1 - O(m^2\gamma^2)$. Moreover,

$$\|Y_p \cdot Y_p^t\| = \|Y_p\| = \|T_p \cdot T_p^t - Z_p \cdot Z_p^t\| \leq \|T_p\|^2 + \|Z_p\|^2 = 1 + O(m^2\gamma^2),$$

which also implies that

$$\|Y_p\| = 1 + O(m^2\gamma^2).$$

Because for any real symmetric matrix $M$, $\|M\| = \max_{v \neq 0} (v^t \cdot M \cdot v) / \|v\|^2$, we conclude that $\|Y_q \cdot Y_q^t - Y_p \cdot Y_p^t\|$ is at most the maximum eigenvalue of $Y_q \cdot Y_q^t$ minus the minimum eigenvalue of $Y_p \cdot Y_p^t$. Therefore,

$$\|Y_q \cdot Y_q^t - Y_p \cdot Y_p^t\| \leq 1 + O(m^2\gamma^2) - (1 - O(m^2\gamma^2)) = O(m^2\gamma^2).$$
On the other hand, for every sample point \( p \in B(x, m\gamma) \),

\[ h(||x - p||) = 0, \quad \Delta h(||x - p||) = 0. \]

Consequently,

\[
||\Delta C_{11}|| = \left|\frac{\sum_{p,q \in P} h(||x - p||) \cdot \Delta h(||x - q||) \cdot (Y_q \cdot Y_q' - Y_q \cdot Y_q')}{\left(\sum_{p \in P} h(||x - p||)\right)^2}\right| 
\leq \frac{\sum_{p,q \in P} h(||x - p||) \cdot |\Delta h(||x - q||)| \cdot ||Y_q \cdot Y_q' - Y_q \cdot Y_q'||}{\left(\sum_{p \in P} h(||x - p||)\right)^2} 
= O(m^2\gamma^2) \cdot \frac{\sum_{p \in P} |\Delta h(||x - p||)|}{\sum_{p \in P} h(||x - p||)}. \tag{3}
\]

By symmetry,

\[
||\Delta C_{12}|| = ||\Delta C_{21}|| = \left|\frac{\sum_{p,q \in P} h(||x - p||) \cdot \Delta h(||x - q||) \cdot (Y_q \cdot Z_q - Y_q \cdot Z_q')}{\left(\sum_{p \in P} h(||x - p||)\right)^2}\right| 
\leq \frac{\sum_{p,q \in P} h(||x - p||) \cdot |\Delta h(||x - q||)| \cdot ||Y_q \cdot Z_q - Y_q \cdot Z_q'||}{\left(\sum_{p \in P} h(||x - p||)\right)^2} 
= O(m\gamma) \cdot \frac{\sum_{p \in P} |\Delta h(||x - p||)|}{\sum_{p \in P} h(||x - p||)}. \tag{4}
\]

Similarly,

\[
||\Delta C_{22}|| = \left|\frac{\sum_{p,q \in P} h(||x - p||) \cdot \Delta h(||x - q||) \cdot (Z_q \cdot Z_q' - Z_q \cdot Z_q')}{\left(\sum_{p \in P} h(||x - p||)\right)^2}\right| 
\leq \frac{\sum_{p,q \in P} h(||x - p||) \cdot |\Delta h(||x - q||)| \cdot ||Z_q \cdot Z_q' - Z_q \cdot Z_q'||}{\left(\sum_{p \in P} h(||x - p||)\right)^2} 
= O(m^2\gamma^2) \cdot \frac{\sum_{p \in P} |\Delta h(||x - p||)|}{\sum_{p \in P} h(||x - p||)}. \tag{5}
\]

From the discussion of generalized gershgorin sets (Section 3.2), we have

\[
||\Delta C_k|| \leq \max\{||\Delta C_{11}|| + ||\Delta C_{12}||, ||\Delta C_{21}|| + ||\Delta C_{22}||\}. \tag{6}
\]

The correctness of (i) is then proved by plugging into (6) the inequalities (3), (4), and (5).

Define the following generalized gershgorin sets:

\[
G_1 = \left\{ \mu : \frac{1}{||C_{11} + \Delta C_{11} - \mu I_m||^{-1}} \leq ||C_{12} + \Delta C_{12}|| \right\},
\]

\[
G_2 = \left\{ \mu : \frac{1}{||C_{22} + \Delta C_{22} - \mu I_{d-m}||^{-1}} \leq ||C_{21} + \Delta C_{21}|| \right\}.
\]

We give a lower bound for the values in \( G_1 \) and an upper bound for the values in \( G_2 \).
Consider \( G_1 \). The minimum eigenvalue of \( C_{11} + \Delta C_{11} \) is at least the minimum eigenvalue of \( C_{11} \) minus \( \|\Delta C_{11}\| \). Therefore, by Lemma 4.1 and \(^{\text{(3)}}\),

\[
\text{minimum eigenvalue of } C_{11} + \Delta C_{11} \geq 1 - O(m^2\gamma^2) - \frac{O(m^2\gamma^2) \cdot \sum_{p \in P} |\Delta h(\|x - p\|)|}{\sum_{p \in P} h(\|x - p\|)}.
\]

On the other hand, by Lemma 4.1 and \(^{\text{(4)}}\),

\[
\|C_{12} + \Delta C_{12}\| \leq \|C_{12}\| + \|\Delta C_{12}\| \\
\quad \leq O(m\gamma) + \frac{O(m\gamma) \cdot \sum_{p \in P} |\Delta h(\|x - p\|)|}{\sum_{p \in P} h(\|x - p\|)}.
\] (7)

The values in \( G_1 \) are at least the minimum eigenvalue value of \( C_{11} + \Delta C_{11} \) minus \( \|C_{12} + \Delta C_{12}\| \). Therefore,

\[
\min\{\mu : \mu \in G_1\} \geq 1 - O(m\gamma) - \frac{O(m\gamma) \cdot \sum_{p \in P} |\Delta h(\|x - p\|)|}{\sum_{p \in P} h(\|x - p\|)}.
\] (8)

Consider \( G_2 \). By Lemma 4.1 and \(^{\text{(5)}}\),

\[
\|C_{22} + \Delta C_{22}\| \leq \|C_{22}\| + \|\Delta C_{22}\| \\
\quad \leq O(m^2\gamma^2) + \frac{O(m^2\gamma^2) \cdot \sum_{p \in P} |\Delta h(\|x - p\|)|}{\sum_{p \in P} h(\|x - p\|)}.
\]

By symmetry and \(^{\text{(7)}}\),

\[
\|C_{21} + \Delta C_{21}\| = \|C_{12} + \Delta C_{12}\| \leq O(m\gamma) + \frac{O(m\gamma) \cdot \sum_{p \in P} |\Delta h(\|x - p\|)|}{\sum_{p \in P} h(\|x - p\|)}.
\]

The values in \( G_2 \) are at most \( \|C_{22} + \Delta C_{22}\| + \|C_{21} + \Delta C_{21}\| \). Therefore,

\[
\max\{\mu : \mu \in G_2\} = O(m\gamma) + \frac{O(m\gamma) \cdot \sum_{p \in P} |\Delta h(\|x - p\|)|}{\sum_{p \in P} h(\|x - p\|)}.
\] (9)

It follows from \(^{\text{(8)}}\) and \(^{\text{(9)}}\) that \( G_1 \) and \( G_2 \) are disjoint because every number in \( G_2 \) is much smaller than those in \( G_1 \). Lemma 3.3 implies that \( G_1 \) contains the \( m \) largest eigenvalues of \( C_x + \Delta C_x \). The correctness of (ii) then follows from \(^{\text{(8)}}\).

We need another technical result on bounding \( |\Delta h(\|x - p\|)| \) from above and \( h(\|x - q\|) \) from below, where \( q \) is the nearest sample point to \( \nu(x) \).

**Lemma 5.2** Let \( x \) be any point at distance \( 2\varepsilon \) or less from \( \mathcal{M} \).

(i) For all \( p \in P \), \( |\Delta h(\|x - p\|)| \leq \left( 1 - \frac{\|x - p\|}{m\gamma} \right)^{2m-1} \cdot O\left( \frac{m\delta_k}{\gamma} \right) \).

(ii) \( h(\|x - q\|) > 0.06 \), where \( q \) is the nearest sample point to \( \nu(x) \).

**Proof.** Consider (i). Since \( \Delta h(\|x - p\|) = 0 \) for any \( p \in P \setminus B(x, m\gamma) \), we only need to consider the case of \( \|x - p\| \leq m\gamma \). Taking derivative gives

\[
|\Delta h(\|x - p\|)| \leq 2m \left( 1 - \frac{\|x - p\|}{m\gamma} \right)^{2m-1} \cdot \frac{2\|x - p\|}{\gamma} \cdot \frac{|x_k - p_k|}{m\gamma\|x - p\|} \cdot \delta_k + \frac{1 - \|x - p\|}{m\gamma} \cdot \frac{2|x_k - p_k|}{\gamma\|x - p\|} \cdot \delta_k
\]

\[
\leq \left( 1 - \frac{\|x - p\|}{m\gamma} \right)^{2m-1} \cdot O\left( \frac{m\delta_k}{\gamma} \right),
\]
establishing the correctness of (i).

Consider (ii). As \( P \) is a uniform \((\varepsilon, \kappa)\)-sample, \( \|q - \nu(x)\| \leq \varepsilon \). Therefore, \( \|q - x\| \leq \|x - \nu(x)\| + \|\nu(x) - q\| \leq 3\varepsilon \). Then,

\[
h(||x - q||) = \left(1 - \frac{\|x - q\|}{m\gamma}\right)^{2m} \left(\frac{2\|x - q\|}{\gamma} + 1\right)
\]

\[
\geq \left(1 - \frac{3\varepsilon}{m\gamma}\right)^{2m}
\]

\[
= \left(1 - \frac{3}{4m}\right)^{2m}.
\]

The minimum of \( \left(1 - \frac{3}{4m}\right)^{2m} \) is achieved at \( m = 1 \), and it is equal to 0.0625.

The following lemma allows us to ignore the contribution of the points near the boundary of \( B(x, m\gamma) \) in \( \sum_{p \in P \cap B(x, m\gamma)} |\Delta h(||x - p||)| \sum_{p \in P \cap B(x, m\gamma)} h(||x - p||) \).

**Lemma 5.3** Let \( x \) be any point at distance \( 2\varepsilon \) or less from \( M \). Let \( P \) be a uniform \((\varepsilon, \kappa)\)-sample of \( M \). Let \( r = \sqrt{m\varepsilon}/3 \). Then,

\[
\sum_{p \in P \cap B(x, m\gamma)} \left(1 - \frac{\|x - p\|}{m\gamma}\right)^{2m-1} \leq (23\kappa + 1) \cdot \sum_{p \in P \cap B(x, m\gamma - r)} \left(1 - \frac{\|x - p\|}{m\gamma}\right)^{2m-1}.
\]

**Proof.** Observe that

\[
\sum_{p \in P \cap B(x, m\gamma)} \left(1 - \frac{\|x - p\|}{m\gamma}\right)^{2m-1} = \sum_{p \in P \cap B(x, m\gamma - r)} \left(1 - \frac{\|x - p\|}{m\gamma}\right)^{2m-1} + \sum_{p \in P \cap B(x, m\gamma) \setminus B(x, m\gamma - r)} \left(1 - \frac{\|x - p\|}{m\gamma}\right)^{2m-1}.
\]

We prove the lemma by bounding the two terms on the right hand side above.

We show a lower bound for the first term. As \( P \) is a uniform \((\varepsilon, \kappa)\)-sample, there exists some point \( q \in P \) such that \( \|q - \nu(x)\| \leq \varepsilon \). Therefore, \( \|q - x\| \leq \|x - \nu(x)\| + \|\nu(x) - q\| \leq 3\varepsilon \leq m\gamma - r \). Then,

\[
\sum_{p \in P \cap B(x, m\gamma - r)} \left(1 - \frac{\|x - p\|}{m\gamma}\right)^{2m-1} \geq \left(1 - \frac{\|x - q\|}{m\gamma}\right)^{2m-1}
\]

\[
\geq \left(1 - \frac{3\varepsilon}{m\gamma}\right)^{2m-1}
\]

\[
\geq \left(1 - \frac{3}{4m}\right)^{2m}.
\]

The quantity \( \left(1 - \frac{3}{4m}\right)^{2m} \) achieves its minimum of \( 1/16 \) when \( m = 1 \). Hence,

\[
\sum_{p \in P \cap B(x, m\gamma - r)} \left(1 - \frac{\|x - p\|}{m\gamma}\right)^{2m-1} \geq \frac{1}{16}.
\]
We show an upper bound for the second term. For any point \( p \in B(x, m\gamma) \setminus B(x, m\gamma - r) \),
\[
(1 - \frac{\|x-p\|}{m\gamma})^{2m-1} \quad \text{achieves its maximum of} \quad (\frac{r}{m\gamma})^{2m-1} = (\frac{1}{12\sqrt{m}})^{2m-1} \quad \text{when} \quad \|x-p\| = m\gamma - r.
\]
By Lemma 3.2 and Lemma 5.4, 
\[
|P \cap B(x, m\gamma) \setminus B(x, m\gamma - r)| \leq |P \cap B(x, m\gamma)| \leq (4m\gamma/\varepsilon + 1)^m\kappa. \quad \text{Therefore,}
\]
\[
\sum_{p \in P \cap B(x, m\gamma) \setminus B(x, m\gamma - r)} \left( 1 - \frac{\|x-p\|}{m\gamma} \right)^{2m-1} \leq (16m + 1)^m\kappa \left( \frac{1}{12\sqrt{m}} \right)^{2m-1} \leq (17)^m\kappa \sqrt{m}/12^{2m-1} \leq 17\kappa/12.
\]
Therefore, the second term is at most the first term multiplied by \( 23\kappa \).

We bound the turn of \( L_x \) when \( x \) moves slightly in the next result.

**Lemma 5.4** For every point \( x \) at distance at most \( 2\varepsilon \) from \( M \) and for every vector \( \Delta x \in N_{\nu(x)} \cup T_{\nu(x)} \), if \( \|\Delta x\| \) is small enough and \( x + \Delta x \) is at distance \( 2\varepsilon \) or less from \( M \), then
\[
\angle(L_x, L_{x+\Delta x}) = O(\kappa m^2 \|\Delta x\|).
\]

**Proof.** Adopt a coordinate frame such that the columns of \( \begin{pmatrix} \lambda & \delta \end{pmatrix} \) form an orthonormal basis of \( T_{\nu(x)} \), and \( \Delta x \) points in the direction of the \( x_k \)-axis for some \( k \in [1, d] \). Let \( \delta_k = \|\Delta x\| \).

Every entry of \( C_k + \Delta C_k \) is some algebraic function in \( \delta_k \). By Taylor’s Theorem, the \((i,j)\) entry of \( C_k + \Delta C_k \) is equal to the \((i,j)\) entry of \( C_k + \Delta C_k \) plus or minus an \( O(\delta_k^2) \) term. Therefore,
\[
C_k + \Delta C_k = C_k + \Delta C_k + Z,
\]
where \( Z \) is a \( d \times d \) matrix in which every entry is \( \pm O(\delta_k^2) \). It follows that
\[
\|Z\| = O(d\delta_k^2).
\]

Since \( Z = C_k + \Delta C_k - (C_k + \Delta C_k) \), \( Z \) is real symmetric.

Let \( e \) be one of the \( m \) most dominant unit eigenvectors of \( C_k + \Delta C_k \). Let \( \sigma \) be the eigenvalue of \( C_k + \Delta C_k \) corresponding to \( e \). Therefore,
\[
C_k + \Delta C_k \cdot e = (C_k + \Delta C_k + Z) \cdot e = \sigma e.
\]

Let \( A_k \) be the \( d \times m \) matrix consisting of the \( m \) most dominant unit eigenvectors of \( C_k \). So \( \text{col}(A_k) \) is the linear subspace spanned by these eigenvectors. Let \( \Lambda \) be the set of the \( d - m \) smallest eigenvalues of \( C_k \). We apply Lemma 5.3 with \( M_1 = C_k, M_2 = \Delta C_k + Z, \) and \( r = m \):

\[
\angle(\text{col}(A_k), e) \leq \arcsin \left( \frac{\|\Delta C_k + Z\|}{\min_{\lambda \in \Lambda} |\lambda - \sigma|} \right) \leq \arcsin \left( \frac{\|\Delta C_k\| + \|Z\|}{\min_{\lambda \in \Lambda} |\lambda - \sigma|} \right).
\]

We bound \( \angle(\text{col}(A_k), e) \) by showing an upper bound for \( \|\Delta C_k\| \) and a lower bound for \( |\lambda - \sigma| \).

For all \( p \in P \setminus B(x, m\gamma) \), \( h(\|x-p\|) = D h(x-p) = 0 \). Then, Lemmas 5.1(i), 5.2(i) and 5.3 imply that
\[
\|\Delta C_k\| \leq \frac{O(m^2 \delta_k)}{\sum_{p \in P \cap B(x, m\gamma)} \left( 1 - \frac{\|x-p\|}{m\gamma} \right)^{2m-1}} \leq \frac{O(m^3 \delta_k)}{\sum_{p \in P \cap B(x, m\gamma - r)} \left( 1 - \frac{\|x-p\|}{m\gamma} \right)^{2m} \left( \frac{2\|x-p\|}{\gamma} + 1 \right)}.
\]
where \( r = \sqrt{m\varepsilon}/3 \). In the denominator, \( \left( 1 - \frac{\|x-p\|}{m\gamma} \right) \left( \frac{2\|x-p\|}{\gamma} + 1 \right) \) is at its minimum of \( \frac{2\sqrt{m\varepsilon}}{3\gamma} \) when \( \|x-p\| = m\gamma - r \). It follows that

\[
\|\Delta C_x\| = O(\kappa m^{3/2} \delta_k). \tag{12}
\]

Lemmas 3.3 and 4.1 imply that

\[
\max \{ \lambda : \lambda \in \Lambda \} = O(m\gamma). \tag{13}
\]

We write \( C_x + \Delta C_x \) as the sum \( C_{x+\Delta x} + (-Z) \) and apply Weyl’s inequality [27, Theorem 3.3.16] to conclude that the eigenvalue \( \sigma \) is at least the \( m \)-th largest eigenvalue of \( C_x + \Delta C_x \) minus the largest eigenvalue of \( -Z \). Then, by Lemma 5.1(ii) and (10),

\[
\sigma \geq 1 - O(m\gamma) - O(m\gamma) \cdot \frac{\sum_{p \in P} |\Delta h(\|x-p\|)|}{\sum_{p \in P} h(\|x-p\|)} - O(d\delta_k^2).
\]

Together with (13), we obtain

\[
\min_{\lambda \in \Lambda} |\lambda - \sigma| \geq 1 - O(m\gamma) - \frac{O(m\gamma) \cdot \sum_{p \in P} |\Delta h(\|x-p\|)|}{\sum_{p \in P} h(\|x-p\|)} - O(d\delta_k^2).
\]

As \( \delta_k \) approaches zero, both \( \Delta h(\|x-p\|) \) and \( O(d\delta_k^2) \) approach zero. But \( \sum_{p \in P} h(\|x-p\|) > 0.06 \) by Lemma 5.2(ii). Therefore, for a sufficiently small \( \delta_k \),

\[
\exists \text{ a constant } \eta > 0 \text{ such that } \min \{ \lambda \in \Lambda : |\lambda - \sigma| \} \geq \eta. \tag{14}
\]

Plugging (10), (12) and (14) into (11) gives

\[
\angle(\text{col}(A_x), e) \leq \arcsin \left( \frac{O(\kappa m^{3/2} \delta_k) + O(d\delta_k^2)}{\eta} \right) = O(\kappa m^{3/2} \delta_k).
\]

Since \( e \) is any one of the \( m \) most dominant unit eigenvectors of \( C_{x+\Delta x} \), the angle bound \( O(\kappa m^{3/2} \delta_k) \) holds for all the \( m \) most dominant unit eigenvectors of \( C_{x+\Delta x} \). Then, by Lemma 3.7, \( \text{col}(A_x) \) makes an \( O(\kappa m^{3/2} \delta_k) \) angle with the space spanned by the \( m \) most dominant unit eigenvectors of \( C_{x+\Delta x} \). It follows that \( \angle(L_x, L_{x+\Delta x}) = O(\kappa m^2 \delta_k) \).

Next, we need a technical result on the angle between a vector in some linear subspace to its projection in another linear subspace.

**Lemma 5.5** Let \( E_1 \) and \( E_2 \) be two \( (d-m) \)-dimensional linear subspaces that make an angle \( \phi < \pi/2 \). Let \( n \) be a unit vector in \( \mathbb{R}^d \). Let \( u_i \) be the projection of \( n \) in \( E_i \) for \( i \in [1, 2] \). Let \( \{v_1, \ldots, v_{d-m}\} \) and \( \{w_1, \ldots, w_{d-m}\} \) be bases of \( E_1 \) and \( E_2 \), respectively, that satisfy either Lemma 3.6(i) or Lemma 3.6(ii). Let \( \alpha_1 = \sum_{i=1}^{d-m} (n^i v_i)^2 \) and let \( \alpha_2 = \sum_{i=d-m+1}^{d} (w_i - v_i)^2 n^2 \). If \( \alpha_1 > \alpha_2 + (2m^2 \phi^2) / \cos \phi \), then

\[
\frac{u_i^t u_j}{\|u_i\| \|u_j\|} \geq \sqrt{1 - \frac{\alpha_2}{\alpha_1} \cos \phi - \frac{2m^2 \phi^2}{\sqrt{\alpha_1^2 - \alpha_1 \alpha_2}}}. \tag{15}
\]

**Proof.** By Lemma 3.6

\[
\forall i \in [1, d-2m], \quad v_i = w_i, \tag{15}
\]

\[
\forall i \in [1, d-m], \quad \angle(v_i, w_i) \leq \phi, \tag{16}
\]

\[
\forall i, j \in [1, d-m], \quad \angle(v_i, w_j - v_j) \in [(\pi - \phi)/2, (\pi + \phi)/2]. \tag{17}
\]
If \( m \geq d/2 \), then (15) is vacuous because \([1, d - 2m]\) is an empty range. There is no harm done as \( d - m \leq d/2 \) in this case and Lemma 3.6(i) is applicable, leading to (16) and (17) only. If \( m < d/2 \), then Lemma 3.6(ii) is applicable, leading to (15), (16) and (17).

Since \( u_i \) is the projection of \( n \) into \( E_i \), we have

\[
\begin{align*}
    u_1 &= (v_1 \cdots v_{d-m})(v_1 \cdots v_{d-m})^t n, \\
    u_2 &= (w_1 \cdots w_{d-m})(w_1 \cdots w_{d-m})^t n.
\end{align*}
\]

(18) \hspace{1cm} (19)

We first bound \( u_1^t u_2 \) from below. Standard algebra gives

\[
    u_1^t u_2 = \sum_{i \in [1,d-m]} n^t v_i v_i^t w_i w_i^t n + \sum_{i \neq j, i,j \in [1,d-m]} n^t v_i v_j^t w_i w_j^t n.
\]

(20)

We analyze the second term in (20). By (15), if \( i \neq j \) and \( i \) or \( j \) belongs to \([1,d-2m]\), then \( v_i \perp w_j \). It implies that \( v_i^t w_j = 0 \) in the second term in (20) whenever \( i \) or \( j \) belongs to \([1,d-2m]\). The remaining case is that both \( i \) and \( j \) belong to \([d-2m+1,d-m]\).

Define a vector \( h_i \) for \( i \in [1,d-m] \) as follows:

\[
    \forall i \in [1,d-m], \quad h_i = w_i - v_i.
\]

It follows from (16) that

\[
    \|h_i\| = 2 \sin \left( \frac{\angle (v_i, w_i)}{2} \right) \leq \phi.
\]

(21)

We rewrite (20) using \( w_i = v_i + h_i \) for \( i \in [d-2m+1,d-m] \):

\[
    u_1^t u_2 = \sum_{i \in [1,d-m]} n^t v_i v_i^t w_i w_i^t n + \sum_{i \neq j, i,j \in [d-2m+1,d-m]} n^t v_i v_j^t (v_j + h_j)(v_j + h_j)^t n
\]

\[
    \quad = \sum_{i \in [1,d-m]} n^t v_i v_i^t w_i w_i^t n + \sum_{i \neq j, i,j \in [d-2m+1,d-m]} (n^t v_i v_j^t h_j^t n + n^t v_i v_j^t h_j^t n).
\]

(22)

Notice that if \( m \geq d/2 \), then \( d - 2m + 1 \leq 1 \), which implies that \([d-2m+1,d-m]\) acts as the range \([1,d-m]\). In this case, Lemma 3.6(i) is applicable and so (15) is vacuous, meaning that there is no simplification from (20) to (22).

By (16), we get

\[
    \forall i \in [1,d-m], \quad v_i^t w_i \geq \cos \phi.
\]

(23)

Moreover,

\[
    \forall i,j \in [1,d-m], \quad v_i^t h_j = \|v_i\| \|h_j\| \cos(\angle (v_i, h_j)) \geq -\|h_j\| \sin(\phi/2) \geq -\phi \sin(\phi/2).
\]

(24)

By substituting (23) and (24) into the first and second terms in (22), respectively, we obtain

\[
    u_1^t u_2 \geq \cos \phi \left( \sum_{i \in [1,d-m]} n^t v_i v_i^t n \right) - \phi \sin \left( \frac{\phi}{2} \right) \left( \sum_{i \neq j, i,j \in [d-2m+1,d-m]} (n^t v_i v_j^t n + n^t v_i v_j^t n) \right)
\]

\[
    = \cos \phi \left( \sum_{i \in [1,d-m]} n^t v_i v_i^t n \right) - \phi \sin \left( \frac{\phi}{2} \right) \left( \sum_{i \neq j, i,j \in [d-2m+1,d-m]} n^t v_i v_j^t n \right).
\]
By definition,
\[ u'_1 u_2 \geq \cos \phi \left( \sum_{i \in [1,d-m]} n'_i w'_i n \right) - m^2 \phi^2. \]

Recall from the lemma statement that \( \alpha_1 = \sum_{i=1}^{d-m} (n'_i)^2 \) and \( \alpha_2 = \sum_{i=d-2m+1}^{d-m} (h'_i n)^2 \). We define one more quantity:
\[ \alpha_3 = \sum_{i=d-2m+1}^{d-m} n'_i h'_i n. \]

Standard algebraic manipulation shows that \( \alpha_1 + \alpha_3 = \sum_{i \in [1,d-m]} n'_i w'_i n \), and therefore,
\[ u'_1 u_2 \geq (\alpha_1 + \alpha_3) \cos \phi - m^2 \phi^2. \]

By definition,
\[
\begin{align*}
||u_1|| &= \sqrt{\sum_{i \in [1,d-m]} (n'_i)^2} = \sqrt{\alpha_1}, \\
||u_2|| &= \sqrt{\sum_{i \in [1,d-m]} (n'_i w'_i)^2} \\
&\geq \sqrt{\frac{\sum_{i \in [1,d-2m]} (n'_i)^2 + \sum_{i \in [d-2m+1,d-m]} (n'_i w'_i)^2}{\sum_{i \in [1,d-2m]} (n'_i)^2 + \sum_{i \in [d-2m+1,d-m]} (n'_i (v'_i + h'_i))^2}} \\
&= \sqrt{\alpha_1 + 2\alpha_3 + \alpha_2}. \tag{25}
\end{align*}
\]

Consequently,
\[
\frac{u'_1 u_2}{||u_1|| \cdot ||u_2||} \geq \frac{(\alpha_1 + \alpha_3) \cos \phi - m^2 \phi^2}{\sqrt{\alpha_1 + 2\alpha_3 + \alpha_2}}.
\]

Treating \( \alpha_3 \) as a free variable while fixing the other values, we can apply standard calculus to show that the right hand side of (25) is minimized when \( \alpha_3 = -\frac{m^2 \phi^2}{\cos \phi} \) under the condition that \( \alpha_1 > \alpha_2 + \frac{2m^2 \phi^2}{\cos \phi} \). (This condition ensures that the denominator \( \sqrt{\alpha_1^2 + 2\alpha_1 \alpha_3 + \alpha_1 \alpha_2} \) is real and positive.) This condition is assumed to be satisfied in the lemma statement. Substituting \( \alpha_3 = -\alpha_2 - \frac{m^2 \phi^2}{\cos \phi} \) into (25) gives
\[
\frac{u'_1 u_2}{||u_1|| \cdot ||u_2||} \geq \frac{(\alpha_1 - \alpha_2) \cos \phi - \frac{2m^2 \phi^2}{\cos \phi}}{\sqrt{\alpha_1 (\alpha_1 - \alpha_2) - \frac{2m^2 \phi^2}{\cos \phi}}} \geq \frac{(\alpha_1 - \alpha_2) \cos \phi - \frac{2m^2 \phi^2}{\cos \phi}}{\sqrt{\alpha_1^2 - \alpha_1 \alpha_2}} \geq \sqrt{1 - \frac{\alpha_2}{\alpha_1} \cos \phi - \frac{2m^2 \phi^2}{\sqrt{\alpha_1^2 - \alpha_1 \alpha_2}}.}
\]

15
We are ready to bound the instantaneous change in the normalized projection of a normal vector of \( \mathcal{M} \) into \( L_x \) as \( x \) moves, which is the main result of this section.

**Lemma 5.6** Let \( z \) be any point in \( \mathcal{M} \). Let \( n \) be any unit vector in \( N_z \). Define the function \( f : B(z, 2\varepsilon) \to L_x \) such that \( f(x) \) is the normalized projection of \( n \) into \( L_x \), i.e., \( f(x) \) is the unit vector in \( L_x \) parallel to the projection of \( n \) in \( L_x \). For every point \( x \) in the interior of \( B(z, 2\varepsilon) \) and every \( k \in [1, d] \), \( \| \partial f(x)/\partial x_k \| = O(\kappa^3) \).

**Proof.** Let \( x \) be a point in the interior of \( B(z, 2\varepsilon) \). Consider any index \( k \in [1, d] \). Let \( \Delta x \) be a vector parallel to the \( x_k \)-axis such that \( x + \Delta x \in B(z, 2\varepsilon) \) and \( \delta_k = \| \Delta x \| \) is arbitrarily small. Let \( \phi \) denote the angle \( \angle(L_x, L_{x+\Delta x}) \). By Lemma 5.4, \( \phi = O(\kappa m^2 \delta_k) \). Since \( \phi < \pi/2 \), there are orthonormal bases of \( L_x \) and \( L_{x+\Delta x} \) that satisfy either Lemma 5.5(i) or Lemma 5.5(ii). Let \( \{v_1, \ldots, v_{d-m}\} \) and \( \{w_1, \ldots, w_{d-m}\} \) be such orthonormal bases of \( L_x \) and \( L_{x+\Delta x} \), respectively. We want to apply Lemma 5.5, so we need to verify that \( \alpha_1 > \alpha_2 + (2m^2 \phi^2)/\cos \phi \), where \( \alpha_1 = \sum_{i=1}^{d-m} (\alpha_i v_i)^2 \) and \( \alpha_2 = \sum_{i=d-m+1}^{d} ((w_i - v_i)\bar{n})^2 \).

First, \( \alpha_2 \leq \sum_{i=d-m+1}^{d} \|w_i - v_i\|^2 \). Since \( \angle(v_i, w_i) \leq \phi \) for \( i \in [d - 2m + 1, d - m] \) by Lemma 3.6 we obtain \( \|w_i - v_i\| = 2 \sin \frac{\angle(v_i, w_i)}{2} \leq \phi \). It follows that

\[
\alpha_2 \leq m \phi^2 = O(\kappa^2 m^5 \delta_k^2).
\]

Second, observe that \( \alpha_1 = \| (v_1 \cdots v_{d-m})(v_1 \cdots v_{d-m})^t \bar{n} \|^2 \), where \( (v_1 \cdots v_{d-m})(v_1 \cdots v_{d-m})^t \bar{n} \) is the projection of \( \bar{n} \) into \( L_x \). Therefore, \( \alpha_1 \geq \cos^2(\angle(L_x, N_z)) \). Then, Lemma 4.2 implies that

\[
\alpha_1 \geq \cos^2(O(\kappa \sqrt{m} \gamma)) \geq 1 - O(m^2 \gamma^2).
\]

As \( \alpha_2 + \frac{2m^2 \phi^2}{\cos \phi} \) approaches zero as \( \delta_k \to 0 \), we get \( \alpha_1 > \alpha_2 + \frac{2m^2 \phi^2}{\cos \phi} \). Then, by Lemma 5.5

\[
\frac{u_1^t u_2}{\|u_1\| \|u_2\|} \geq \sqrt{1 - \frac{\alpha_2}{\alpha_1} \cos \phi - \frac{2m^2 \phi^2}{\sqrt{\alpha_1} - \alpha_1 \alpha_2}}.
\]

where \( u_1 \) and \( u_2 \) are the projections of \( \bar{n} \) into \( L_x \) and \( L_{x+\Delta x} \), respectively. Finally,

\[
\left\| \frac{\partial f(x)}{\partial x_k} \right\|^2 = \lim_{\delta_k \to 0} \frac{1}{\delta_k^2} \left( \frac{u_2}{\|u_2\|} - \frac{u_1}{\|u_1\|} \right)^t \left( \frac{u_2}{\|u_2\|} - \frac{u_1}{\|u_1\|} \right) = \lim_{\delta_k \to 0} \frac{1}{\delta_k^2} \left( 2 - \frac{2u_1^t u_2}{\|u_1\| \|u_2\|} \right) \leq \lim_{\delta_k \to 0} \frac{1}{\delta_k^2} \left( 2 - 2 \left( 1 - \frac{\alpha_2}{\alpha_1} \right) \cos \phi + \frac{4m^2 \phi^2}{\sqrt{\alpha_1} - \alpha_1 \alpha_2} \right) \leq \lim_{\delta_k \to 0} \frac{1}{\delta_k^2} \left( 2 - 2 \left( 1 - \frac{\alpha_2}{\alpha_1} \right) \cos \phi + \frac{4m^2 \phi^2}{\sqrt{\alpha_1} - \alpha_1 \alpha_2} \right).
\]

We have shown earlier that \( \alpha_2 \leq m \phi^2 \) and \( \alpha_1 \geq 1 - O(m^3 \gamma^2) \). Using these relations and the
facts that \( \cos \phi \geq 1 - \phi^2/2 \) and \( \phi = O(km^2 \delta_k) \), we obtain

\[
\left\| \frac{\partial f(x)}{\partial x_k} \right\|^2 \leq \lim_{\delta x \to 0} \frac{1}{\delta_k^2} \left( 2 - 2 \left( 1 - \frac{m \phi^2}{\alpha_1} \right) \left( 1 - \frac{\phi^2}{2} \right) + \frac{4m^2 \phi^2}{\alpha_1^2 - \alpha_1 m \phi^2} \right)
\]

\[
= \lim_{\delta x \to 0} \frac{1}{\delta_k^2} \left( 2 - 2 \left( 1 - \frac{m \phi^2}{\alpha_1} - \frac{\phi^2}{2} + \frac{m \phi^4}{2 \alpha_1} \right) + \frac{4m^2 \phi^2}{\alpha_1^2 - \alpha_1 m \phi^2} \right)
\]

\[
\leq \lim_{\delta x \to 0} \frac{1}{\delta_k^2} \left( O \left( \frac{\kappa^2 m^6 \delta_k^2}{\alpha_1} \right) + \frac{O(\kappa^2 m^6 \delta_k^2)}{\sqrt{\alpha_1^2 - \alpha_1 m \delta_k^2}} \right)
\]

\[
= O(\kappa^2 m^6).
\]

We use Lemma 5.6 to bound \( \|J_f(x)\| \). Multiplying the bound in Lemma 5.6 by \( \sqrt{d} \) already gives a bound. We give a tighter analysis that yields a bound independent of \( d \).

**Lemma 5.7** Let \( z \) be any point in \( M \). Let \( J_f \) be the Jacobian of the function \( f : B(z, 2\varepsilon) \to L_x \) defined in Lemma 5.6. For any point \( x \) in the interior of \( B(z, 2\varepsilon) \), \( \|J_f(x)\| = O(km^3) \).

**Proof.** Fix a unit vector \( n \in N_z \) as required in the definition of \( f \) in Lemma 5.6. Let \( x \) be a point in the interior of \( B(z, 2\varepsilon) \). Let \( R \) be any \( d \times d \) orthogonal matrix. Apply the orthogonal transformation induced by \( R \) to \( \mathbb{R}^d \). Then define the function \( g : B(z', 2\varepsilon) \to L_{x'} \), where \( z' = R \cdot z \) and \( x' = R \cdot x \), such that \( g(x') \) is the normalized projection of \( R \cdot n \) into \( L_{x'} \).

First, we show that \( f(x) = R^t \cdot g(x') \). Let \( \ell \) be the length of the projection of \( n \) into \( L_x \). Let \( Q \) be any \( d \times (d - m) \) matrix whose columns form an orthonormal basis of \( L_x \). It follows from the definition of \( f \) that \( f(x) = \frac{1}{\ell} \cdot Q \cdot Q^t \cdot n \). Since an orthogonal transformation preserves lengths, \( \ell \) is also the length of the projection of \( R \cdot n \) into \( L_{x'} \). Then, \( g(x') = \frac{1}{\ell} \cdot Q \cdot Q^t \cdot R \cdot n = \frac{1}{\ell} \cdot R \cdot Q \cdot Q^t \cdot n \), which implies that \( f(x) = R^t \cdot g(x') \).

We show that \( J_f(x) = R^t \cdot J_g(x') \cdot R \). Let \( \Delta x \) be an arbitrarily short vector. By Taylor’s Theorem,

\[
f(x + \Delta x) = f(x) + J_f(x) \cdot \Delta x + e_f,
\]

where \( \|e_f/\|\Delta x\| \) converges to the zero vector as \( \|\Delta x\| \to 0 \). Similarly,

\[
g(R \cdot x + R \cdot \Delta x) = g(x') + J_g(x') \cdot R \cdot \Delta x + e_g,
\]

where \( \|e_g/\|R \cdot \Delta x\| \) converges to the zero vector as \( \|R \cdot \Delta x\| \to 0 \). Since \( R \) is fixed, it means that \( e_g/\|\Delta x\| \) tends to the zero vector as \( \|\Delta x\| \to 0 \). We multiply both sides of (27) by \( R^t \) and then subtract the resulting equation from (26). Some terms cancel each other because \( f(x + \Delta x) = R^t \cdot g(R \cdot (x + \Delta x)) \) and \( f(x) = R^t \cdot g(x') = R^t \cdot g(R \cdot x) \). We obtain

\[
(J_f(x) - R^t \cdot J_g(x') \cdot R) \cdot \Delta x = R^t \cdot e_g - e_f.
\]

Therefore,

\[
\left\| (J_f(x) - R^t \cdot J_g(x') \cdot R) \cdot \Delta x \right\| \leq \left\| R^t \cdot e_g \right\| + \left\| e_f \right\|.
\]

We are free to choose the direction of \( \Delta x \). We choose it such that \( \left\| (J_f - R^t \cdot J_g(x') \cdot R) \cdot \Delta x \right\| = \left\| J_f(x) - R^t \cdot J_g(x') \cdot R \right\| \cdot \|\Delta x\| \), i.e., \( \Delta x \) is an eigenvector corresponding to the largest eigenvalue of \( J_f(x) - R^t \cdot J_g(x') \cdot R \). Then,

\[
\left\| J_f(x) - R^t \cdot J_g(x') \cdot R \right\| \leq \frac{\left\| R^t \right\| \left\| e_g \right\|}{\|\Delta x\|} + \frac{\left\| e_f \right\|}{\|\Delta x\|}.
\]
Since the right hand side tends to zero as $\|\Delta x\| \to 0$, we conclude that
\[
\lim_{\|\Delta x\| \to 0} \|J_f(x) - R^t \cdot J_g(x') \cdot R\| = 0,
\]
which implies that $J_f(x) = R^t \cdot J_g(x') \cdot R$.

By definition, $\|J_f(x)\| = \|J_f(x) \cdot v\|$ for some unit vector $v$. We choose $R$ to be the $d \times d$ orthogonal matrix such that $R \cdot v = (1, 0, \ldots, 0)^t$. Then, $\|R \cdot J_f(x) \cdot v\| = \|R^t \cdot J_f(x) \cdot v\| = \|J_g(x') \cdot (1, 0, \ldots, 0)^t\|$, which is the 2-norm of the first column of $J_g(x')$. Lemma 5.6 is independent of the coordinate frame. So we can apply Lemma 5.6 to $g$ and conclude that the 2-norm of the first column of $J_g(x')$ is $O(km^3)$. As a result, $\|R \cdot J_f(x) \cdot v\| = O(km^3)$. Since multiplying any vector with an orthogonal matrix preserves the 2-norm of the vector, we conclude that $\|J_f(x)\| = \|J_f(x) \cdot v\| = \|R \cdot J_f(x) \cdot v\| = O(km^3)$.

\[\square\]

6 Faithful reconstruction

In this section, we prove our main result that $Z_\varphi \cap \tilde{M}$ is a faithful reconstruction of $\mathcal{M}$. Recall the class $\Phi$ of functions $\varrho : \mathbb{R}^d \to \mathbb{R}^{d-m}$:

\[
\Phi = \left\{ \varrho : \varrho(x) = \sum_{p \in P} \omega(x, p) \cdot B_{\varrho,x}^t (x - p) \right\}, \quad \text{where } B_{\varrho,x} \text{ is any } d \times (d - m) \text{ matrix with linearly independent columns such that } \text{col}(B_{\varrho,x}) = L_x.
\]

We claim that the choice of $B_{\varrho,x}$ has no impact on the zero-set $Z_\varrho$ as long as the columns of $B_{\varrho,x}$ are linearly independent. In this section, we will prove some useful properties of functions in $\Phi$. These properties will allow us to show that $Z_\varrho \cap \tilde{M}$ is a faithful approximation of $\mathcal{M}$.

We will study properties of $Z_\varrho \cap \tilde{M}$ by analyzing $Z_\varrho \cap \hat{\mathcal{M}}$ for another function $\varrho \in \Phi$ conveniently chosen for the analysis. Since we will conduct some local analysis, we are only concerned with functions that are defined near some chosen points in $\mathcal{M}$. This motivates us to define for every point $z \in \mathcal{M}$ the following class $\Phi_z$ of functions:

\[
\Phi_z = \left\{ \varrho : \varrho(z) = \sum_{p \in P} \omega(z, p) \cdot B_{\varrho,x}^t (z - p) \right\}, \quad \text{where } B_{\varrho,x} \text{ is any } d \times (d - m) \text{ matrix with linearly independent columns such that } \text{col}(B_{\varrho,x}) = L_x.
\]

$\Phi_z$ is a local version of $\Phi$. The next result shows that functions in $\Phi_z$ with overlapping domains have consistent zero sets.

**Lemma 6.1** Let $y$ and $z$ be two arbitrary points in $\mathcal{M}$ that are not necessarily distinct. For every point $x \in B(y, 2\varepsilon) \cap B(z, 2\varepsilon)$, if there exists $\varrho \in \Phi_y$ such that $\varrho(x) = 0_{d-m,1}$, then for every $\varrho \in \Phi_y \cup \Phi_z$, $\varrho(x) = 0_{d-m,1}$.

**Proof.** Take two functions $\varrho, \tilde{\varrho} \in \Phi_y \cup \Phi_z$. Fix a point $x \in B(y, 2\varepsilon) \cap B(z, 2\varepsilon)$. By definition, $\varrho(x) = \sum_{p \in P} \omega(x, p) \cdot B_{\varrho,x}^t (x - p)$ and $\tilde{\varrho}(x) = \sum_{p \in P} \omega(x, p) \cdot B_{\tilde{\varrho},x}^t (x - p)$. The columns of $B_{\varrho,x}$ and $B_{\tilde{\varrho},x}$ form two bases of $L_x$, which means that there is a $(d - m) \times (d - m)$ invertible matrix $R$ such that $R \cdot B_{\varrho,x}^t = B_{\tilde{\varrho},x}^t$. If $\varrho(x) = 0_{d-m,1}$, then $\tilde{\varrho}(x) = \sum_{p \in P} \omega(x, p) \cdot R \cdot B_{\tilde{\varrho},x}^t (x - p) = R \cdot \varrho(x) = 0_{d-m,1}$.

We define a particular function $\varrho_z \in \Phi_z$ to analyze the properties of $Z_\varrho \cap \tilde{M}$ in a small neighborhood of $z$. 19
Combining the above observations, we obtain $u$ and it is at most 1 as $d$ increases such that for every point $z \in B(z, 2\varepsilon)$, $d$ decreases as $d$ increases such that for every point $z \in \mathcal{M}$, if $\varepsilon \leq \varepsilon_0$, then $\theta_z \in \Phi_z$ and $\theta_z$ is continuous in the interior of $B(z, 2\varepsilon)$.

**Definition 1** Let $z$ be any point in $\mathcal{M}$. Let $\{v_1, \ldots, v_{d-m}\}$ be any set of unit vectors forming a basis of $N_2$. For $i \in [1, d-m]$, let $f_{v_i}$ be the function that maps every $x$ in $B(z, 2\varepsilon)$ to the normalized projection of $v_i$ in $L_x$. Define a canonical function $\theta_z : B(z, 2\varepsilon) \to \mathbb{R}^{d-m}$ with respect to $z$ and $\{v_1, \ldots, v_{d-m}\}$ such that for all $x \in B(z, 2\varepsilon)$, $\theta_z(x) = \sum_{p \in P} \omega(x, p) \cdot [f_{v_1}(x), \ldots, f_{v_{d-m}}(x)]^t \cdot (x - p)$.

We show that whenever $\varepsilon$ is sufficiently small, $\theta_z$ belongs to $\Phi_z$ and $\theta_z$ is continuous in the interior of $B(z, 2\varepsilon)$.

**Lemma 6.2** Let $\theta_z$ be the canonical function with respect to a point $z \in \mathcal{M}$ and some set of unit vectors $\{v_1, \ldots, v_{d-m}\}$ forming a basis of $N_2$ for which there exists some $\phi \in \left[0, \arcsin \left(\frac{1}{3d-m}\right)\right]$ such that for any distinct $i, j \in [1, d-m]$, $\angle(v_i, v_j) \leq \left|\frac{\pi}{2} - \phi, \frac{\pi}{2} + \phi\right]$. There exists $\varepsilon_0 \in (0, 1)$ that decreases as $d$ increases such that for every point $z \in \mathcal{M}$, if $\varepsilon \leq \varepsilon_0$, then $\theta_z \in \Phi_z$ and $\theta_z$ is continuous in the interior of $B(z, 2\varepsilon)$.

**Proof.** To show that $\theta_z \in \Phi_z$, it suffices to prove that $\{f_{v_1}(x), \ldots, f_{v_{d-m}}(x)\}$ form a basis of $L_x$, which boils down to showing that $\{f_{v_1}(x), \ldots, f_{v_{d-m}}(x)\}$ are linearly independent.

Since $\angle(L_x, N_2) = O(m\sqrt{m}\gamma)$ by Lemma 4.2, we get $\angle(f_{v_i}(x), v_j) = O(m\sqrt{m}\gamma)$. Assume to the contrary that $f_{v_i}(x), \ldots, f_{v_{d-m}}(x)$ are linearly dependent. Then,

$$\langle f_{v_i}(x), \text{col}((f_{v_2}(x) \cdots f_{v_{d-m}}(x))) \rangle = 0.$$ 

Since $\langle v_i, \text{col}((f_{v_2}(x) \cdots f_{v_{d-m}}(x))) \rangle = O(m\sqrt{m}\gamma)$ for all $i \in [2, d - m]$, Lemma 3.7 implies that

$$\langle \text{col}((v_2 \cdots v_{d-m})), \text{col}((f_{v_2}(x) \cdots f_{v_{d-m}}(x))) \rangle = O\left(m\sqrt{dm - m^2}\gamma\right).$$

By triangle inequality, $\langle v_1, \text{col}((v_2 \cdots v_{d-m})) \rangle \leq \langle v_1, f_{v_1}(x) \rangle + \langle f_{v_1}(x), \text{col}((v_2 \cdots v_{d-m})) \rangle$. The dimension of $\text{col}((v_2 \cdots v_{d-m}))$ is at least the dimension of $\text{col}((f_{v_2}(x) \cdots f_{v_{d-m}}(x)))$. Thus,

$$\langle f_{v_1}(x), \text{col}((v_2 \cdots v_{d-m})) \rangle \leq \langle f_{v_1}(x), \text{col}((f_{v_2}(x) \cdots f_{v_{d-m}}(x))) \rangle + \langle \text{col}((v_2 \cdots v_{d-m})), \text{col}((f_{v_2}(x) \cdots f_{v_{d-m}}(x))) \rangle.$$

Combining the above observations, we obtain

$$\langle v_1, \text{col}((v_2 \cdots v_{d-m})) \rangle \leq \langle v_1, f_{v_1}(x) \rangle + \langle f_{v_1}(x), \text{col}((v_2 \cdots v_{d-m})) \rangle \leq \langle v_1, f_{v_1}(x) \rangle + \langle f_{v_1}(x), \text{col}((f_{v_2}(x) \cdots f_{v_{d-m}}(x))) \rangle + \langle \text{col}((v_2 \cdots v_{d-m})), \text{col}((f_{v_2}(x) \cdots f_{v_{d-m}}(x))) \rangle = O(m\sqrt{dm - m^2}\gamma).$$

Recall that $\gamma = 4\varepsilon \leq 4\varepsilon_0$. Assume that $\varepsilon_0 < \frac{1}{C\sqrt{dm - m^2}}$ for some appropriate constant $C \geq 1$. Then $\langle v_1, \text{col}((v_2 \cdots v_{d-m})) \rangle < \pi/6$. Note that $\varepsilon_0$ decreases as $d$ increases. Let $u$ be the normalized projection of $v_1$ in $\text{col}(v_2 \cdots v_{d-m})$. It means that

$$v_1 \cdot u > \cos(\pi/6) = \sqrt{3}/2.$$

We can write $u = \sum_{i=2}^{d-m} \lambda_i v_i$ for some $\lambda_i$. Let $k = \arg\max_{i \in [2, d-m]} |\lambda_i|$. We take the dot product of $u$ and $\text{sign}(\lambda_k)v_k$. This dot product is equal to $|\lambda_k| |v_k|^2 + \text{sign}(\lambda_k) \sum_{i \neq k} \lambda_i v_i \cdot v_k$ and it is at most 1 as $u$ and $v_k$ are unit vectors. Since $\langle v_i, v_j \rangle \in \left[\frac{\pi}{2} - \phi, \frac{\pi}{2} + \phi\right]$, the projection of $v_j$ in the direction of $v_i$ has magnitude at most $\sin\phi$. It follows that

$$1 \geq |\lambda_k| - \sum_{i \neq k} |\lambda_i| v_i \cdot v_k \geq |\lambda_k| - (d - m - 2)|\lambda_k| \sin \phi.$$
We get $|\lambda_k| \leq 1/(1 - (d - m - 2)\sin \phi) < 1.5$ because $\sin \phi < \frac{1}{3d-3m}$ by assumption of the lemma. Thus,

$$v_1^t \cdot u = \sum_{i=2}^{d-m} \lambda_i v_i^t \cdot v_i \leq \sin \phi \cdot \sum_{i=2}^{d-m} |\lambda_i| < 1.5(d - m) \sin \phi < 0.5.$$ 

This is a contradiction because we have derived earlier that $v_1^t \cdot u > \sqrt{3}/2$. We conclude that \{f_{v_1(x), \ldots, f_{d-m}}(x)\} are linearly independent, and therefore, $g_x \in \Phi_2$.

By Lemma 5.6 for $i \in [1, d - m]$, $f_{v_i}$ is differentiable and hence continuous in the interior of $B(z, 2\varepsilon)$. Because $g_x$ is a sum of products of continuous functions, $g_x$ is also continuous in the interior of $B(z, 2\varepsilon)$ [34, Ch 2: Corollary 3.7].

Next, we show that the gradient of $g_x$ varies monotonically.

**Lemma 6.3** Let $z$ be any point in $M$. Let $v_i$ be any unit vector in $N_z$. For any $x \in B(z, 2\varepsilon)$, let $g_{x,i}(x) = \sum_{p \in P} \omega(x, p) \cdot f_{v_i}(x)^t \cdot (x - p)$. Let $\tau$ be any value greater than 1. For every $t \geq 1$ and every point $x \in B(z, t\varepsilon)$,

- $\|\nabla g_{x,i}(x)\| \in [1 - O(t^2(1/\sqrt{m}\varepsilon^{-1} + km^4\gamma)), 1 + O(t^2(1/\sqrt{m}\varepsilon^{-1} + km^4\gamma))]$ and
- $v_1^t \cdot \nabla g_{x,i}(x) \geq 1 - O(t^2(1/\sqrt{m}\varepsilon^{-1} + km^4\gamma))$.

**Proof.** From the definition of $g_{x,i}(x) = \sum_{p \in P} \omega(x, p) \cdot f_{v_i}(x)^t \cdot (x - p)$, we obtain

$$\|\nabla g_{x,i}(x)\| \leq \sum_{p \in P} \left( \omega(x, p) + \omega(x, p) \cdot \|J_{f_{v_i}}(x)\| \cdot \|x - p\| \right) + \sum_{p \in P} \nabla \omega(x, p) \cdot f_{v_i}(x)^t \cdot (x - p).$$

Consider the first term in (28). By Lemma 5.7, $\|J_{f_{v_i}}(x)\| = O(\kappa m^3)$. For any $p \notin B(x, m\gamma)$, $\omega(x, p)$ vanishes. If $p \in B(x, m\gamma)$, then

$$\|J_{f_{v_i}}(x)\| \cdot \|x - p\| = O(\kappa m^4\gamma).$$

Therefore,

$$\sum_{p \in P} \left( \omega(x, p) + \omega(x, p) \cdot \|J_{f_{v_i}}(x)\| \cdot \|x - p\| \right) \leq 1 + O(\kappa m^4\gamma).$$

Consider the second term in (28). For any point $p \notin B(x, m\gamma)$, $\nabla \omega(x, p)$ is a zero vector. If $p \in B(x, m\gamma)$, then $\|p - \nu(x)\| \leq \|p - x\| + \|x - \nu(x)\| \leq m\gamma + t\varepsilon = O(m\gamma)$. By Lemma 3.1(i), $p - \nu(x)$ makes an angle $\pi/2 - O(m\gamma)$ with $N_{\nu(x)}$. It follows from Lemma 4.2 that $p - \nu(x)$ makes an angle $\pi/2 - O(m\sqrt{m}\gamma)$ with $L_x$. Therefore, the projection of $p - \nu(x)$ onto $L_x$ has length less than $O(m\sqrt{m}\gamma) \cdot O(m\gamma) = O(m^{5/2}\gamma^2)$. Since $f_{v_i}(x)$ is a unit vector in $L_x$, the projection $p - \nu(x)$ in $L_x$ has length at least $|f_{v_i}(x)^t \cdot (p - \nu(x))| \geq |f_{v_i}(x)^t \cdot (p - x)| - \|x - \nu(x)\|$. Therefore,

$$|f_{v_i}(x)^t \cdot (x - p)| \leq \|x - \nu(x)\| + O(m^{5/2}\gamma^2) \leq t\varepsilon + O(m^{5/2}\gamma^2).$$

We conclude that

$$\left\| \sum_{p \in P} \nabla \omega(x, p) \cdot f_{v_i}(x)^t \cdot (x - p) \right\| \leq O(t\varepsilon + m^{5/2}\gamma^2) \cdot \sum_{p \in P} \|\nabla \omega(x, p)\|. $$

21
Let us establish the upper range limit for \( \|\nabla \omega(x, p)\| \).

By substituting (30) and (34) into (28), we have

\[
\|\nabla \omega(x, p)\| \leq \frac{2 \sum_{p \in P} \frac{dh(||x-p||)}{d||x-p||} \cdot ||p-x|| \cdot \sum_{p \in P} \frac{dh(||x-p||)}{d||x-p||} \cdot ||p-x||}{\sum_{p \in P} h(||x-p||)^2}.
\]

By Lemma 5.2(i), differentiating \( h(||x-p||) \) with respect to \( ||x-p|| \) gives

\[
\left| \frac{dh(||x-p||)}{d||x-p||} \right| \leq O\left(\frac{m}{\gamma}\right) \cdot \left(1 - \frac{||x-p||}{m\gamma}\right)^{2m-1}.
\]

On the other hand,

\[
\sum_{p \in P} h(||x-p||) = \sum_{p \in P} \left(1 - \frac{||x-p||}{m\gamma}\right)^{2m} \left(\frac{2||x-p||}{\gamma} + 1\right).
\]

For all \( p \in P \setminus B(x, m\gamma) \), \( h(||x-p||) = 0 \) and \( \left| \frac{dh(||x-p||)}{d||x-p||} \right| = 0 \). Then,

\[
\sum_{p \in P} \|\nabla \omega(x, p)\| \leq \frac{O\left(\frac{m}{\gamma}\right) \cdot \sum_{p \in P \setminus B(x, m\gamma)} \left(1 - \frac{||x-p||}{m\gamma}\right)^{2m} \left(\frac{2||x-p||}{\gamma} + 1\right)}{\sum_{p \in P \setminus B(x, m\gamma-r)} \left(1 - \frac{||x-p||}{m\gamma}\right)^{2m} \left(\frac{2||x-p||}{\gamma} + 1\right)}.
\]

Let \( r = \sqrt{m\varepsilon}/3 \). By Lemma 5.3

\[
\sum_{p \in P} \|\nabla \omega(x, p)\| \leq \frac{O\left(\frac{m}{\gamma}\right) \cdot \sum_{p \in P \setminus B(x, m\gamma-r)} \left(1 - \frac{||x-p||}{m\gamma}\right)^{2m} \left(\frac{2||x-p||}{\gamma} + 1\right)}{\sum_{p \in P \setminus B(x, m\gamma-r)} \left(1 - \frac{||x-p||}{m\gamma}\right)^{2m} \left(\frac{2||x-p||}{\gamma} + 1\right)}.
\]

In the denominator, the term \( \left(1 - \frac{||x-p||}{m\gamma}\right) \left(\frac{2||x-p||}{\gamma} + 1\right) \) achieves its minimum \( \frac{2\sqrt{m\varepsilon}}{3\gamma} - \frac{2\varepsilon^2}{9\gamma} + \frac{\varepsilon}{3\sqrt{m\gamma}} = \Omega(\sqrt{m}) \) when \( ||x-p|| = m\gamma - r \). It follows that

\[
\sum_{p \in P} \|\nabla \omega(x, p)\| = O(\kappa\sqrt{m}/\gamma).
\]

Substituting (33) into (32) gives

\[
\left\| \sum_{p \in P} \nabla \omega(x, p) \cdot f_{\omega}(x)^t \cdot (x-p) \right\| = O(\kappa \sqrt{m \varepsilon^{t-1} + \kappa m^3 \gamma}).
\]

By substituting (30) and (34) into (28), we have

\[
\|\nabla \theta_t(x)\| \leq 1 + O(\kappa \sqrt{m \varepsilon^{t-1} + \kappa m^3 \gamma}),
\]

establishing the upper range limit for \( \|\nabla \theta_t(x)\| \). Symmetrically,

\[
\|\nabla \theta_t(x)\| \geq \sum_{p \in P} \left( \omega(x, p) - \omega(x, p) \cdot ||J_{f_{\omega}}(x)\| \cdot ||x-p|| \right) - \left| \left\| \sum_{p \in P} \nabla \omega(x, p) \cdot f_{\omega}(x)^t \cdot (x-p) \right\| \right|.
\]
By (29) and (34), we have
\[ \|\nabla \varrho_{m,i}(x)\| \geq 1 - O(km^4\gamma) - O(t\sqrt{m}c^{t-1} + km^3\gamma) = 1 - O(t\sqrt{m}c^{t-1} + km^4\gamma), \]
establishing the lower range limit for \( \|\nabla \varrho_{m,i}(x)\| \).

Observe that
\[ v^i \cdot \nabla \varrho_{m,i}(x) = \sum_{p \in P} \omega(x, p) \cdot v^i \cdot f_{v,i}(x) + \sum_{p \in P} \omega(x, p) \cdot v^i \cdot J_{f_{v,i}}(x^i) \cdot (x - p) + \sum_{p \in P} v^i \cdot \nabla \omega(x, p) \cdot f_{v,i}(x^i) \cdot (x - p). \]

Therefore,
\[ v^i \cdot \nabla \varrho_{m,i}(x) \geq \sum_{p \in P} \omega(x, p) \cdot v^i \cdot f_{v,i}(x) - \left| \sum_{p \in P} \omega(x, p) \cdot v^i \cdot J_{f_{v,i}}(x^i) \cdot (x - p) \right| - \left| \sum_{p \in P} v^i \cdot \nabla \omega(x, p) \cdot f_{v,i}(x^i) \cdot (x - p) \right|. \]

Since \( \langle f_{v,i}(x), v^i \rangle \) is \( O(m\sqrt{m}) \) by Lemma 6.2, we get \( v^i \cdot f_{v,i}(x) \geq 1 - O(m^3\gamma^2) \), which implies that \( \sum_{p \in P} \omega(x, p) \cdot v^i \cdot f_{v,i}(x) \geq 1 - O(m^3\gamma^2) \). The second term is at most \( \sum_{p \in P} \omega(x, p) \cdot \|J_{f_{v,i}}(x^i)\| \cdot \|x - p\| \leq O(km^4\gamma^2) \) by (29). The third term is at most \( \sum_{p \in P} \|\nabla \omega(x, p)\| \cdot |f_{v,i}(x^i) \cdot (x - p)| \), which is \( O(t\sqrt{m}c^{t-1} + km^3\gamma) \) by (31) and (33). As a result, \( v^i \cdot \nabla \varrho_{m,i}(x) \geq 1 - O(t\sqrt{m}c^{t-1} + km^4\gamma) \).

The next result shows that every point \( z \) in \( M \) is near \( Z_{\varrho_{m}} \).

**Lemma 6.4** Let \( \varrho_{m} \) be the canonical function with respect to a point \( z \in M \) and an orthonormal basis \( \{v_1, \ldots, v_{d-m}\} \) of \( N_z \). There exists \( \varepsilon_0 \in (0, 1) \) and \( c_m \geq 1 \) such that if \( \varepsilon \leq \varepsilon_0 \), then \( Z_{\varrho_{m}} \cap B(z, c_m\gamma^2) \cap (z + N_z) \neq \emptyset \) and \( Z_{\varrho_{m}} \cap (B(z, 2\varepsilon) \setminus B(z, c_m\gamma^2)) \cap (z + N_z) = \emptyset \). The value \( \varepsilon_0 \) decreases as \( d \) increases, and \( c_m \) is linear in \( m^{5/2} \).

**Proof.** We first show that \( Z_{\varrho_{m}} \cap (B(z, 2\varepsilon) \setminus B(z, c_m\gamma^2)) \cap (z + N_z) \) is empty. For all \( i \in [1, d-m] \) and all point \( x \in B(z, 2\varepsilon) \), let \( \varrho_{m,i} = \sum_{p \in P} \omega(x, p) : f_{v,i}(x^i) \cdot (x - p) \).

We claim that there exists a value \( c_m \geq 1 \) that is linear in \( m^{5/2} \) such that for every \( x \in B(z, 2\varepsilon) \cap (z + N_z) \) and every \( i \in [1, d-m] \), if \( v^i \cdot (x - z) \geq c_m\gamma^2 \), then \( \varrho_{m,i}(x) > 0 \). We ignore all \( p \in P \setminus B(x, m\gamma) \) because \( \omega(x, p) = 0 \) in this case, so such points have no influence over \( \varrho_{m,i}(x) \). \( P \cap B(x, m\gamma) \) is non-empty because, by uniform \((\varepsilon, \kappa)\)-sampling, there is a point \( p \in P \) such that \( \|q - z\| \leq \varepsilon \) which implies that \( \|q - x\| \leq \|x - z\| + \|q - z\| \leq 3\varepsilon \leq m\gamma \). For every \( p \in P \cap B(x, m\gamma) \),
\[ v^i \cdot (x - p) \geq v^i \cdot (x - z) - |v^i \cdot (z - p)|. \]
The first term is bounded from below as \( v^i \cdot (x - z) \geq c_m\gamma^2 \) by assumption. Consider the second term. Since \( \|p - z\| \leq \|p - x\| + \|x - z\| \leq m\gamma + 2\varepsilon < (m + 1)\gamma \), Lemma 3.4(i) implies that the second term \( |v^i \cdot (z - p)| \) is at most \( (m + 1)^2\gamma^2/2 \). It follows that
\[ v^i \cdot (x - p) \geq c_m\gamma^2 - (m + 1)^2\gamma^2/2. \]

For \( i \in [1, d-m] \), define \( h_i(x) = f_{v,i}(x) - v^i \). Lemma 4.2 implies that
\[ \|h_i(x)\| \leq 2\sin \frac{\angle(x, N_z)}{2} = O(m\sqrt{m} \gamma). \]
Assume without loss of generality that $y \in \{a, b\}$ and has side length 2. Our claim in the previous paragraph ensures the existence and uniqueness of $\omega_{x, i}(x)$. Any number greater than 1 will do.

Observe that

$$f_{v_i}(x) \cdot (x - p) = v_i \cdot (x - p) + h_i(x) \cdot (x - p)$$

where $v_i \cdot (x - p) \geq c_m \gamma^2 - (m + 1)^2 \gamma^2/2 - \|h_i(x)\| \cdot \|x - p\| \geq c_m \gamma^2 - (m + 1)^2 \gamma^2/2 - O(m^{5/2} \gamma^2)$

> 0,

whenever $c_m$ is a large enough value that is linear in $m^{5/2}$. As a result, $\omega_{x, i}(x) > 0$. This proves our claim.

We can symmetrically show that if $v_i \cdot (x - z) \leq -c_m \gamma^2$, then $\omega_{x, i}(x) < 0$. Thus, $\omega_{x, i}^{-1}(0) \cap B(z, 2\varepsilon) \cap (z + N_z)$ lies in a $(d - m)$-dimensional slab $S_{v_i} \subset z + N_z$ that is bounded by two $(d - m - 1)$-dimensional flats orthogonal to $v_i$ and at distance $c_m \gamma^2$ from $z$. It follows that $(Z_{y_i} \cap (B(z, 2\varepsilon) \cap (z + N_z))) \setminus S_{v_i} = \emptyset$. By Lemma 6.3, $Z_{y_i}$ is identical for any choice of the orthonormal basis $\{v_1, \ldots, v_{d-m}\}$ of $N_z$. It means that we can set $v_i$ to be any unit vector $v \in N_z$ and the proof above still works. Observe that $\bigcap_{v \in N_z} S_v = B(z, c_m \gamma^2) \cap (z + N_z)$. Hence, $Z_{y_i} \cap (B(z, 2\varepsilon) \setminus B(z, c_m \gamma^2)) \cap (z + N_z) = \emptyset$.

To establish that $Z_{y_i} \cap (B(z, c_m \gamma^2)) \cap (z + N_z) \neq \emptyset$, it suffices to show that $\bigcap_{i=1}^{d-m} Z_{v_i}(0)$ contains a point in $\bigcap_{i=1}^{d-m} S_{v_i}$. This is because $\bigcap_{i=1}^{d-m} S_{v_i}$ is contained in $B(z, c_m \sqrt{d - m} \gamma^2)$, and for $\varepsilon_0 \leq 1/(16c_m \sqrt{d - m})$, we have $B(z, c_m \sqrt{d - m} \gamma^2) \subseteq B(z, \varepsilon_0)$ as $c_m \sqrt{d - m} \varepsilon^2 \leq 16c_m \sqrt{d - m} \varepsilon^2 \leq 16c_m \sqrt{d - m} \varepsilon_0^2$.

In fact, we choose an even smaller $\varepsilon_0$ such that $\sqrt{\varepsilon_0} \leq 1/(16c_m \sqrt{d - m})$, which gives $c_m \sqrt{d - m} \gamma^2 \leq \varepsilon_0^{3/2}$. This will allow us apply Lemma 6.3 later. The exponent $3/2$ is an arbitrary choice. Any number greater than 1 will do.

Let $C = \bigcap_{i=1}^{d-m} S_{v_i}$. It is a $(d - m)$-dimensional cube that lies in $z + N_z$, has $z$ as its center, and has side length $2c_m \gamma^2$. The facets of $C$ are orthogonal to the directions $v_1, \ldots, v_{d-m}$.

Adopt a coordinate frame such that $v_1, \ldots, v_{d-m}$ are the first $d - m$ coordinate axes of $\mathbb{R}^d$. For $i \in [1, d - m]$, define $H_i$ to be the set of maximal line segments that lie inside $C$ and are parallel to the direction $v_i$.

First, we claim that every line segment $l \in H_i$ intersects $\omega_{x, i}^{-1}(0)$ at exactly one point. We have shown earlier that $\omega_{x, i}$ has opposite signs at the endpoints of $l$. So $l \cap \omega_{x, i}^{-1}(0) \neq \emptyset$. Suppose to the contrary that $l \cap \omega_{x, i}^{-1}(0)$ contains two distinct points $y_1$ and $y_2$. So $y_1 - y_2$ is parallel to $v_i$. Assume without loss of generality that $y_1 - y_2$ has the same orientation as $v_i$. By Lemma 6.3, $y_1 - y_2 \cdot \omega_{x, i}(x) > 0$ for every $x \in B(z, c_m \sqrt{d - m} \gamma^2) \subseteq B(z, \varepsilon_0^{3/2})$. But then $\omega_{x, i}(x)$ increases strictly monotonically from $y_2$ to $y_1$, which implies that $\omega_{x, i}(y_1) > 0$. This is a contradiction because $y_1 \in \omega_{x, i}^{-1}(0)$, thereby establishing our claim.

Define a function $g_i : C \to [-c_m \gamma^2, c_m \gamma^2]$ such that $g_i(x) = b_{i,x}$, where

- $(x_1, x_{i-1}, k_{i,x}, x_{i+1}, \ldots, x_d) \in C$ and
- $\omega_{x, i}(x_1, x_{i-1}, k_{i,x}, x_{i+1}, \ldots, x_d) = 0$.

Our claim in the previous paragraph ensures the existence and uniqueness of $b_{i,x}$. We show that $g_i$ is continuous. Since $\omega_{x, i}$ is continuous, $\omega_{x, i}^{-1}(0)$ is compact Ch 3: Theorem 5.4, Ch 5: Theorem 2.11, which implies that for any interval $[a, b] \subset \mathbb{R}$, $\omega_{x, i}^{-1}(0) \cap \{x \in C : x_i \in [a, b]\}$ is compact. Let $\pi_i$ be the function that projects points in $C$ onto the linear subspace spanned by $\{v_1, \ldots, v_{i-1}, v_{i+1}, \ldots, v_{d-m}\}$. Since $\pi_i$ is continuous, its image is compact and so is the following product Ch 5: Theorem 2.9 & Theorem 4.2):

$$\pi_i \left(\omega_{x, i}^{-1}(0) \cap \{x \in C : x_i \in [a, b]\}\right) \times [-c_m \gamma^2, c_m \gamma^2].$$
Observe that this product is homeomorphic to \( g_t^{-1}([a, b]) \). Therefore, \( g_t^{-1}([a, b]) \) is compact for any interval \([a, b] \subset \mathbb{R}\), which implies that \( g_t \) is continuous \cite[Ch 2: Theorem 6.10]{34}.

Define a function \( g : C \rightarrow C \) such that

\[
g(x) = (g_1(x), \ldots, g_{d-m}(x))^t.
\]

The function \( g \) is continuous as each \( g_i \) is continuous. Notice that \( \rho_{Z_i}^{-1}(0) \cap C \) is the subset of \( C \) that satisfy the equation \( g_i(x_1, \ldots, x_i, \ldots, x_d) = x_i \). Since \( \rho_i(x) = (\rho_1(x), \ldots, \rho_{d-m}(x))^t \), we conclude that \( Z_{\theta_i} \cap C \) is the subset of \( C \) that satisfy the equation \( g(x) = x \). By the Brouwer fixed-point theorem \cite[Ch 4: Theorem 4.6]{34}, there is indeed such a point in \( C \).

Recall that \( \nu \) is the map that sends every point in \( \mathbb{R}^d \) to its nearest point in \( \mathcal{M} \). We need to show that \( Z_\varphi \cap \tilde{\mathcal{M}} \) is compact in order to prove that \( Z_\varphi \cap \tilde{\mathcal{M}} \) and \( \mathcal{M} \) are homeomorphic.

**Lemma 6.5** \( Z_\varphi \cap \tilde{\mathcal{M}} \) is compact.

**Proof.** By Lemmas \ref{lem:6.1} and \ref{lem:6.2} for any point \( z \in \mathcal{M} \), \( Z_\varphi \) agrees locally with \( Z_{\varphi_z} \) where \( \varphi_z \) is the canonical function with respect to \( z \) and any orthonormal basis of \( N_z \). Our strategy is to construct a finite number of such \( Z_{\varphi_z} \)'s and prove that each is compact. The lemma then follows as a finite union of compact sets is compact.

Take a maximal set \( Y \) of points in \( \tilde{\mathcal{M}} \) such that any two of them are at distance \( \varepsilon^* \) or more apart. It implies that any two balls centered at points in \( \varepsilon^*/2 \) are interior-disjoint. Since \( \tilde{\mathcal{M}} \) is the product of \( \mathcal{M} \) and a ball of radius \( \varepsilon \), \( \tilde{\mathcal{M}} \) is compact \cite[Ch 5: Theorem 4.2]{34}. It follows that \( |Y| \) is finite. The maximality also implies that \( \tilde{\mathcal{M}} \subseteq \bigcup_{y \in Y} B(y, \varepsilon^*) \). The intersection \( Z_\varphi \cap \bigcup_{y \in Y} B(y, \varepsilon^*) \) is equal to \( \bigcup_{y \in Y} Z_\varphi \cap B(y, \varepsilon^*) \) which is a union of \( Z_{\varphi_{\nu(y)}} \cap B(\nu(y), \varepsilon^* + \varepsilon) \) because \( \|y - \nu(y)\| \leq \varepsilon \). By Lemmas \ref{lem:6.1} and \ref{lem:6.2}, \( Z_\varphi \cap B(\nu(y), \varepsilon^* + \varepsilon) = Z_{\varphi_{\nu(y)}} \cap B(\nu(y), \varepsilon^* + \varepsilon) \). Therefore,

\[
Z_\varphi \cap \tilde{\mathcal{M}} \subseteq \bigcup_{y \in Y} Z_{\varphi_{\nu(y)}} \cap B(\nu(y), \varepsilon^* + \varepsilon) \subseteq \bigcup_{y \in Y} Z_{\varphi_{\nu(y)}} \cap B(\nu(y), \varepsilon^* + \varepsilon).
\]

As \( \varphi_{\nu(y)} \) is continuous in the interior of \( B(\nu(y), 2\varepsilon) \) by Lemma \ref{lem:6.2}, \( Z_{\varphi_{\nu(y)}} \cap B(\nu(y), \varepsilon^* + \varepsilon) \) is compact \cite[Ch 3: Theorem 5.4, Ch 5: Theorem 2.11]{34}. It implies that the finite union \( \bigcup_{y \in Y} Z_{\varphi_{\nu(y)}} \cap B(\nu(y), \varepsilon^* + \varepsilon) \) is also compact. Finally, observe that

\[
Z_\varphi \cap \tilde{\mathcal{M}} = \left( \bigcup_{y \in Y} Z_{\varphi_{\nu(y)}} \cap B(\nu(y), \varepsilon^* + \varepsilon) \right) \cap \tilde{\mathcal{M}},
\]

which is compact because it is the intersection of two compact subsets in \( \mathbb{R}^d \).

We are ready to prove the faithful approximation of \( \mathcal{M} \) by \( Z_\varphi \cap \tilde{\mathcal{M}} \).

**Theorem 6.1** Let \( \mathcal{M} \) be an \( m \)-dimensional compact smooth manifold in \( \mathbb{R}^d \). Let \( P \) be a uniform \((\varepsilon, \kappa)\)-sample of \( \mathcal{M} \) for some constant \( \kappa \geq 1 \). We assume that \( \mathcal{M} \) has unit reach, \( m \) is known, a neighborhood radius \( \gamma = 4\varepsilon \), and approximate tangent spaces with angular errors at most \( m\gamma \) are specified at the points in \( P \). Let \( \mathcal{M} \) be the set of points within a distance \( \varepsilon \) from \( \mathcal{M} \). We can construct a function \( \varphi : \mathbb{R}^d \rightarrow \mathbb{R}^{d-m} \) for which there exists \( \varepsilon_0 \in (0, 1) \) that decreases as \( d \) increases such that the following properties hold whenever \( \varepsilon \leq \varepsilon_0 \).

(i) The restriction of the nearest point map to \( Z_\varphi \cap \tilde{\mathcal{M}} \) is a homeomorphism between \( Z_\varphi \cap \tilde{\mathcal{M}} \) and \( \mathcal{M} \).

(ii) The Hausdorff distance between \( Z_\varphi \cap \tilde{\mathcal{M}} \) and \( \mathcal{M} \) is \( O(m^{5/2}\gamma^2) = O(m^{5/2}\varepsilon^2) \).
(iii) For all \( x \in Z_\varphi \cap \hat{M} \), \( N_{\nu(x)} \) makes an \( O(m^2 \sqrt{\kappa \tau}) = O(m^2 \sqrt{\kappa \varepsilon}) \) angle with the normal space of \( Z_\varphi \) at \( x \).

Proof. Consider (i). Let \( \mu \) denote the restriction of \( \nu \) to \( Z_\varphi \cap \hat{M} \). First, we show that \( \mu \) is injective. Suppose to the contrary that there are two points \( y_1, y_2 \in Z_\varphi \cap \hat{M} \) such that \( \mu(y_1) \) and \( \mu(y_2) \) are the same point \( z \in M \). Then, \( y_1 \) and \( y_2 \) belong to \( z + N_z \), which implies that \( y_1 - y_2 \in N_z \). Note that \( y_1 \) and \( y_2 \) lie in \( B(z, \varepsilon) \). By Lemmas 6.1 and 6.2, \( Z_\varphi \cap B(z, \varepsilon) = Z_\varphi \cap B(z, \varepsilon) \). Then, Lemma 6.4 implies that \( y_1 \) and \( y_2 \) belong to \( B(z, t \gamma^2) \) for some large enough \( t \) that is linear in \( m^{5/2} \). By Lemma 6.3, we can define \( v_1 = y_1 - y_2 \) and get \( (y_1 - y_2)^\top \nabla_{B_{\ell,2}}(x) > 0 \) for all \( x \in B(z, t \gamma^2) \) when \( \varepsilon_0 \) is sufficiently small. But then \( \theta_{x,1}(x) \) increases strictly monotonically from \( y_2 \) to \( y_1 \), which implies that \( \theta_{x,1}(y_1) > 0 \). This is a contradiction because \( y_1 \) belongs to \( Z_\varphi \) and hence \( Z_{\theta_x} \) by Lemmas 6.1 and 6.2. This proves that \( \mu \) is injective.

Next, we show that \( \mu \) is surjective. Let \( z \) be any point in \( M \). It follows from Lemmas 6.1, 6.2 and 6.4 that there exists a point \( y \in Z_\varphi \cap \hat{M} \cap (z + N_z) \). We show that \( \mu \) must map \( y \) to \( z \). Suppose that \( \mu \) maps \( y \) to another point \( z_2 \in \hat{M} \), i.e. \( \|y - z_2\| < \|y - z\| \). We grow a ball \( B \) tangent to \( M \) at \( z \) by moving its center linearly from \( z \) towards \( y \). When \( B \) is tiny, it touches \( \hat{M} \) only at \( z \). When the center of \( B \) reaches \( y \), \( B \) contains both \( z \) and \( z_2 \). Thus, the radius of the growing \( B \) must become the local feature size of \( M \) at \( z \) before or when its center reaches \( y \). Recall that the reach of \( M \) is assumed to be \( 1 \). Thus, \( \|y - z\| \geq 1 > \varepsilon \). This contradicts the fact that \( y \in \hat{M} \cap (z + N_z) \), thereby proving that \( \mu \) is surjective.

Since \( Z_\varphi \cap \hat{M} \) avoids the medial axis, the restriction \( \mu \) is continuous. Therefore, \( \mu \) is a continuous bijection from \( Z_\varphi \cap \hat{M} \) to \( M \). The spaces \( M \) and \( Z_\varphi \cap \hat{M} \) are compact by assumption and Lemma 6.5, respectively, so we conclude from the existence of \( \mu \) that \( M \) and \( Z_\varphi \cap \hat{M} \) are homeomorphic [34] Ch 5: Theorem 2.14. This proves the correctness of (i).

Consider (ii). By Lemmas 6.1, 6.2 and 6.4, for any point \( z \in \hat{M} \), there exists a point \( x \in Z_\varphi \) within a distance of \( c_m \gamma^2 \), where \( c_m \geq 1 \) is some value linear in \( m^{5/2} \). Therefore, \( c_m \gamma^2 = O(m^{5/2} \varepsilon_0 \varepsilon) < \varepsilon \) for a small enough \( \varepsilon_0 \). So \( x \in Z_\varphi \cap \hat{M} \). It follows that the directed Hausdorff distance from \( M \) to \( Z_\varphi \cap \hat{M} \) is \( O(m^{5/2} \gamma^2) \). Conversely, for any point \( x \in Z_\varphi \cap \hat{M} \), \( \|\nu(x) - x\| \leq \varepsilon \) and \( x \in \nu(x) + N_{\nu(x)} \). By Lemmas 6.1, 6.2 and 6.4, \( Z_\varphi \cap (B(\nu(x), 2\varepsilon) \setminus B(\nu(x), c_m \gamma^2)) \cap (\nu(x) + N_{\nu(x)}) \) is empty. So \( \|\nu(x) - x\| \leq c_m \gamma^2 = O(m^{5/2} \gamma^2) \). It follows that the directed Hausdorff distance from \( Z_\varphi \cap \hat{M} \) to \( M \) is \( O(m^{5/2} \gamma^2) \).

Consider (iii). By Lemma 6.3, for every point \( x \in Z_\varphi \cap \hat{M} \) and every unit vector \( v_1 \in N_{\nu(x)} \), \( \|\nabla_{\theta_{\nu(x),1}}(x)\| \leq 1 - O(km^4 \gamma) \) and \( v_1 \cdot \nabla_{\theta_{\nu(x),1}}(x) \geq 1 - O(km^4 \gamma) \). Thus,

\[
\angle(v_1, \nabla_{\theta_{\nu(x),1}}(x)) \leq \arccos \left( \frac{v_1 \cdot \nabla_{\theta_{\nu(x),1}}(x)}{\|\nabla_{\theta_{\nu(x),1}}(x)\|} \right) \leq \arccos \left( \frac{1 - O(km^4 \gamma)}{1 + O(km^4 \gamma)} \right) = O(m^2 \sqrt{\kappa \gamma}).
\]

The vector \( \nabla_{\theta_{\nu(x),1}}(x) \) belongs to the normal space of \( Z_\varphi \) at \( x \). (Recall that \( Z_\varphi \) agrees with \( Z_{\theta_{\nu(x)}} \) locally.) Thus, the angle between \( N_{\nu(x)} \) and the normal space of \( Z_\varphi \) at \( x \) is \( O(m^2 \sqrt{\kappa \gamma}) \).

7 Projection operator

Our proof of convergence will make use of the property that \( B_{\varphi,x} \) is a \( d \times (d - m) \) matrix with orthogonal unit columns such that \( \text{col}(B_{\varphi,x}) = L_x \). Such a matrix can be obtained by an eigen-decomposition of \( C_x \).

We rewrite \( \varphi(x) = \sum_{p \in F} \omega(x, p) \cdot B_{\varphi,x}^\top \cdot (x - p) = B_{\varphi,x}^\top \cdot (x - a_x) \), where \( a_x = \sum_{p \in F} \omega(x, p) \cdot p \). Intuitively, as \( \varphi(a_x) = 0 \), we want to move the current point \( x \) closer to \( a_x \). We also want to move directly onto \( Z_\varphi \) without much drifting. Therefore, it is desirable to move \( x \) within the
affine subspace $x_i + L_{x_i}$ which is roughly normal to $Z_\varphi$.

The projection follows an iterative scheme:

$$x_{i+1} = x_i + B_{\varphi, x_i} \cdot B_{\varphi, x_i}^t \cdot (a_{x_i} - x_i).$$

Note that $B_{\varphi, x_i} \cdot B_{\varphi, x_i}^t \cdot (a_{x_i} - x_i)$ is the projection of the vector $a_{x_i} - x_i$ into $L_{x_i}$. The iterative scheme moves the current point $x_i$ by this projected vector to the new point $x_{i+1}$. In other words, $x_{i+1}$ is the projection of $a_{x_i}$ onto the affine subspace $x_i + L_{x_i}$.

We prove two technical results in order to establish the proof of convergence. The first one shows that any initial point near $M$ is moved to within an $O(m^{7/2}\gamma^2)$ distance from $M$ after a single iteration. Let $\hat{x}_i$ denote the nearest point in $Z_\varphi$ to $x_i$. The second result shows that $\|x_{i+1} - \hat{x}_i\| < \|x_i - \hat{x}_i\|$, which implies that $\|x_{i+1} - \hat{x}_i\| < \|x_i - \hat{x}_i\|$.\[Lemma 7.1\] Let $P$ be a uniform $(\varepsilon, \kappa)$-sample of $M$. For every point $x$ within a distance $m\gamma$ from $P$ and every $d \times (d - m)$ matrix $B_{\varphi, x}$ that satisfies $\text{col}(B_{\varphi, x}) = L_x$, we have $\|\nu - (x)\| = O(m^{7/2}\gamma^2)$, where $y = x + B_{\varphi, x} \cdot B_{\varphi, x}^t \cdot (a_x - x)$.

Proof. For every sample point $p \in B(x, m\gamma)$, $\|p - (x)\| \leq \|p - x\| + \|x - (x)\| = O(m\gamma)$. By Lemma 5.1(ii), the distance between $p$ and $(x) + T_{(x)}$ is $O(m^2\gamma^2)$. As $a_x$ is convex combination of all $p \in B(x, m\gamma)$, the distance between $a_x$ and $(x) + T_{(x)}$ is also $O(m^2\gamma^2)$.

Let $\hat{a}_x$ be the projection of $a_x$. The vector $\hat{a}_x - a_x$ is parallel to $T_{(x)}$, so $\hat{a}_x$ is also at distance $O(m^2\gamma^2)$ from $(x) + T_{(x)}$. As $\hat{a}_x \in (x) + N_{(x)}$, the vector $\hat{a}_x - (x)$ is orthogonal to $T_{(x)}$, which implies that $\|\hat{a}_x - (x)\| = O(m^2\gamma^2)$. Therefore, it suffices to prove that $\|\hat{a}_x - y\| = O(m^{7/2}\gamma^2)$ as $\|y - (x)\| \leq \|\hat{a}_x - y\| + \|\hat{a}_x - (x)\| = \|\hat{a}_x - y\| + O(m^2\gamma^2)$.

Refer to Figure 1(a). By construction, $\hat{a}_x \in (x) + N_{(x)}$. Also, $x - (x) \in N_{(x)}$, implying that $x \in (x) + N_{(x)}$. Therefore, $\angle x \hat{a}_x a_x = \pi/2$. From the previous discussion, $y$ is the projection of $a_x$ onto $x + L_x$. So $\angle xy a_x = \pi/2$. As a result, $x, y, \hat{a}_x$, and $a_x$ lie on a $(d - 1)$-dimensional sphere $S$ that has $a_x$ as a diameter. Since $a_x$ is a convex combination of all $p \in P \cap B(x, m\gamma)$, we have $\|a_x - x\| \leq m\gamma$. Thus, radius($S$) = $O(m\gamma)$.

Since $\angle x \hat{a}_x a_x = \pi/2$, we have $\|\hat{a}_x - x\|^2 = \|\hat{a}_x - a_x\|^2 = \|a_x - x\|^2$. It follows that $\|\hat{a}_x - x\| \geq \|a_x - x\|/2$ or $\|\hat{a}_x - a_x\| \geq \|a_x - x\|/2$. We prove that $\angle a_x x y = O(m^{7/2}\gamma)$ if $\|\hat{a}_x - x\| \geq \|a_x - x\|/2$.

Let $\{v_1, \ldots, v_{d-m}\}$ and $\{w_1, \ldots, w_{d-m}\}$ be orthonormal bases of $N_{(x)}$ and $L_x$, respectively, that satisfy Lemma 3.6. Note that $\hat{a}_x - x \in N_{(x)}$ and $y - x \in L_x$. Refer to Lemma 5.5. Let $(a_x - x)/\|a_x - x\|$ be the unit vector $n$, let $\hat{a}_x - x$ be the vector $u_1$, let $y - x$ be the vector $u_2$ as specified in Lemma 5.5 and let $\phi = \angle(L, N_{(x)}) = O(m^{3/2}\gamma^2)$ by Lemma 4.2. We need to show that the values $\alpha_1$ and $\alpha_2$ defined in Lemma 5.5 satisfy the assumption that $\alpha_1 > \alpha_2 + (2m^2\phi^2)/\cos^2 \phi$.\[Figure 1: (a) The points $x, y, \hat{a}_x,$ and $a_x$ lie on a $(d - 1)$-dimensional sphere with $x a_x$ as a diameter. (b) The circle with center $y$ circumscribes $y a_x a_x$. Also, $\angle a_x o y = 2\angle a_x a_x y$.\]
By Lemma 3.6, \( \angle (v_i, w_i) \leq \phi \) for \( i \in [1, d - m] \), which implies that \( \|v_i - w_i\| \leq 2 \sin(\phi/2) \leq \phi \).

By definition, \( \alpha_2 = \sum_{i=d-m}^{d-m+1} (w_i - v_i)^2 \), and therefore, \( \alpha_2 \leq \sum_{i=d-m}^{d-m+1} |w_i - v_i|^2 \leq m \delta^2 = O(m^4 \gamma^2) \). By definition, \( \alpha_1 \) is the squared norm of the projection of \( n = (\tilde{a}_x - x)/\|\tilde{a}_x - x\| \) onto \( N_{\nu}(x) \). Since \( \tilde{a}_x - x \) is the projection of \( a_x \) onto \( N_{\nu}(x) \), we get \( \alpha_1 = \|\tilde{a}_x - x\|^2/\|\tilde{a}_x - x\|^2 \geq 1/4 \) because \( \|\tilde{a}_x - x\| / \|a_x - x\| \leq 1/2 \) by assumption. This shows that \( \alpha_1 > \alpha_2 + (2m^2 \phi^2) / \cos \phi \).

Then,Lemma 5.5 implies that \( \tilde{a}_x y \leq \angle (u_1, u_2) \leq \arccos \left( \sqrt{1 - \frac{\alpha_2}{\alpha_1}} \cos \phi - \frac{2m^2 \phi^2}{\alpha_1 \cos \phi + \alpha_2} \right) \). One can verify that the right hand side is \( \arccos(1 - O(m^4 \gamma^2)) \) and so \( \tilde{a}_x y = O(m^5 / \gamma^2) \).

Similarly, we can prove that \( \angle (a_x, y) = O(m^5 / \gamma^2) \) if \( \|\tilde{a}_x - a_x\| \geq \|a_x - x\| / 2 \). We conclude that \( \angle (\tilde{a}_x, a_x, y) = O(m^5 / \gamma^2) \) or \( \angle (\tilde{a}_x, a_x, y) = O(m^5 / \gamma^2) \).

Without loss of generality, assume that \( \angle (\tilde{a}_x, a_x, y) = O(m^5 / \gamma^2) \). Consider the circumsphere of \( \tilde{a}_x a_x y \). Let \( o \) be its center. Refer to Figure 1(b). The angle \( \tilde{a}_x o y = 2 \tilde{a}_x a_x y \). Then, \( \|a_x - y\|^2 = 2\|o - y\|^2 \sin(\angle (\tilde{a}_x, o y) / 2) \leq \text{radius}(S) \cdot O(m^5 / \gamma^2) = O(m^5 / \gamma^2) \).

Next, we prove that \( x_{i+1} \) is much closer to \( \rho \) than \( x_i \).

**Lemma 7.2** Let \( P \) be a uniform \((\varepsilon, \kappa)\)-sample of \( M \). There exists \( \varepsilon_0 \in (0, 1) \) that decreases as \( d \) and \( \kappa \) increase such that if \( \varepsilon \leq \varepsilon_0 \), then for any point \( y \) at distance \( m^{7/2} \gamma^2 \) or less from \( M \), we have \( \|y - \tilde{y}\| \leq \kappa^{1/4} \cdot \|y - \hat{y}\| \), where \( \hat{y} \) is the nearest point in \( \rho \cap M \) to \( y \) and \( y' = y + B_{\rho,y} \cdot B_{\rho,y} \cdot (a_y - y) \).

**Proof.** Let \( z = v(y) \). For \( i \in [1, d - m] \), let \( v_i \) be the unit vector in \( N_z \) such that \( B_{\rho,y} = (f_{v_1}(y), \ldots, f_{v_{d-m}}(y)) \) consists of orthogonal unit column vectors. By Lemma 4.2, \( \angle (L_y, N_z) = O(m \sqrt{m} \gamma) \), so for any distinct \( i, j \in [1, d - m] \), \( \angle (v_i, v_j) = \pi/2 \pm O(m \sqrt{m} \gamma) \). This allows us to prove as in the proof of Lemma 6.2 that \( \{v_1, \ldots, v_{d-m}\} \) are linearly independent and form a basis of \( N_z \).

Let \( g_x \) be the canonical function with respect to \( z \) and the basis \( \{v_1, \ldots, v_{d-m}\} \) of \( N_z \). Since \( \|y - \tilde{y}\| \) is at most \( \|y - z\| \) plus the distance from \( z \) to \( \rho \), by Theorem 6.1, we have \( \|y - \tilde{y}\| \leq O(m^{7/2} \gamma^2) + O(m^{5/2} \gamma^2) = O(m^{7/2} \gamma^2) \). So \( \|y - z\| \leq \|y - \tilde{y}\| + \|y - z\| \leq O(m^{7/2} \gamma^2) \).

Therefore, segment \( y \tilde{y} \) is contained in \( B(z, m^{7/2} \gamma^2) \) for some constant \( t \), implying that \( g_x(x) \) is defined for any point \( x \) in the segment \( y \tilde{y} \) as long as \( \varepsilon_0 < 1 / (8t m^{7/2}) \) so that \( t m^{7/2} \gamma^2 \leq 16t m^{7/2} \varepsilon_0 < 2 \varepsilon \). By Lemmas 6.1 and 6.2, \( g_x^{-1}(0) \) agrees with \( \rho \) within \( B(z, t m^{7/2} \gamma^2) \). Then, the following relations follow from Lemma 4.2, Lemma 6.3, Theorem 6.1 and the facts that \( \angle (v_i, f_{v_i}(y)) = O(m \sqrt{m} \gamma) \) for any \( i \in [1, d - m] \), and \( \angle (v_i, f_{v_i}(y)) = \pi/2 \pm O(m \sqrt{m} \gamma) \) for any distinct \( i, j \in [1, d - m] \).

- For all \( i \in [1, d - m] \) and all \( x \in B(z, t m^{7/2} \gamma^2) \), \( \|\nabla \theta_{x,i}(x)\| \leq \|1 - O(k m^4 \gamma), 1 + O(k m^4 \gamma)\| \).
- For all distinct indices \( i, j \in [d - m] \) and for all pair of points \( x, x' \in B(z, t m^{7/2} \gamma^2) \), \( \nabla \theta_{x,i}(x)^T \cdot \nabla \theta_{x,j}(x') = \pm O(k m^4 \gamma) \).
- For all \( i \in [d - m] \), \( f_{v_i}(y)^T \cdot \nabla \theta_{x,i}(y) \in \{1 - O(k m^4 \gamma), 1 + O(k m^4 \gamma)\} \).
- For all distinct \( i, j \in [d - m] \), \( f_{v_i}(y)^T \cdot \nabla \theta_{x,j}(y) = \pm O(k m^4 \gamma) \).

We first prove lower and upper bounds on \( \|\theta(z)\| \). Since \( \tilde{y} \) is the nearest point in \( \tilde{z} \cap \tilde{M} \) to \( y \), the vector \( y - \tilde{y} \) belongs to the normal space of \( \rho \) at \( \tilde{y} \). Recall that \( \theta_{x,i} \) agrees with \( \rho \) locally, so the normal space of \( \rho \) at \( \tilde{y} \) is spanned by \( \{\nabla \theta_{x,1}(\tilde{y}), \ldots, \nabla \theta_{x,d-m}(\tilde{y})\} \). Let
\( u = \sum_{i=1}^{d-m} \lambda_i \cdot \nabla \theta_{z,i}(\tilde{y}) \) denote the unit vector \((y - \tilde{y})/\|y - \tilde{y}\|\). Standard vector calculus gives

\[
\theta_{z}(y) = \left( \int_{0}^{1} (\nabla \theta_{z,1}(\tilde{y} + ru), \ldots, \nabla \theta_{z,d-m}(\tilde{y} + ru))^t \cdot (y - \tilde{y}) \, dr \right)
\]
\[
= \|y - \tilde{y}\| \cdot \int_{0}^{1} (\nabla \theta_{z,1}(\tilde{y} + ru), \ldots, \nabla \theta_{z,d-m}(\tilde{y} + ru))^t \cdot \left( \sum_{i=1}^{d-m} \lambda_i \cdot \nabla \theta_{z,i}(\tilde{y}) \right) \, dr
\]
\[
= \|y - \tilde{y}\| \cdot \left( \lambda_1 + \sum_{i=1}^{d-m} (\pm \lambda_i) \cdot O(\kappa m^4 \gamma) \right)
\]
\[
= \lambda_{d-m} + \sum_{i=1}^{d-m} (\pm \lambda_i) \cdot O(\kappa m^4 \gamma) \right).
\]

Hence,

\[
\sum_{i=1}^{d-m} \lambda_i^2 = \sum_{i=1}^{d-m} |\lambda_i|^2 \leq \frac{\|\theta_{z}(y)\|^2}{\|y - \tilde{y}\|^2} \leq \sum_{i=1}^{d-m} \lambda_i^2 + O(\kappa m^4 \gamma) \cdot \sum_{i=1}^{d-m} \lambda_i^2.
\]

We claim that if \( \varepsilon_0 \) is small enough, then

\[
\forall i \in [1, d - m], \quad |\lambda_i| \leq 1 + O((d - m)\kappa m^4 \gamma).
\]

Let \( k = \arg\max_{i \in [1, d - m]} |\lambda_i| \). We take the dot product of \( \sum_{i=1}^{d-m} \lambda_i \cdot \nabla \theta_{z,i}(\tilde{y}) \) and \( \nabla \theta_{z,k}(\tilde{y}) \) or \(-\nabla \theta_{z,k}(\tilde{y})\) depending on whether \( \lambda_k \) is non-negative or negative, respectively. This dot product is at most \( 1 + O(\kappa m^4 \gamma) \) as \( \|\nabla \theta_{z,k}(\tilde{y})\| = 1 + O(\kappa m^4 \gamma) \). On the other hand, for each \( i \neq k \),

\[
\lambda_i \cdot \nabla \theta_{z,i}(\tilde{y})^t \cdot \nabla \theta_{z,k}(\tilde{y}) \) contributes \( \pm |\lambda_i| \cdot O(\kappa m^4 \gamma) \). It follows that

\[
|\lambda_k| \cdot (1 - O(\kappa m^4 \gamma)) - O(\kappa m^4 \gamma) \sum_{i \neq k} |\lambda_i| \leq 1 + O(\kappa m^4 \gamma)
\]
\[
\Rightarrow (1 - O((d - m)\kappa m^4 \gamma))) |\lambda_k| \leq 1 + O(\kappa m^4 \gamma)
\]
\[
\Rightarrow |\lambda_k| \leq 1 + O((d - m)\kappa m^4 \gamma)).
\]

Since \( |\lambda_k| = \max_i |\lambda_i| \), it establishes our claim.

Since \( \sum_{i=1}^{d-m} \lambda_i \cdot \nabla \theta_{z,i}(\tilde{y}) \) is a unit vector, we get

\[
\left\| \sum_{i=1}^{d-m} \lambda_i \cdot \nabla \theta_{z,i}(\tilde{y}) \right\|^2 = \sum_{i=1}^{d-m} \lambda_i^2 \cdot \left\| \nabla \theta_{z,i}(\tilde{y}) \right\|^2 + \sum_{i \neq j} \lambda_i \lambda_j \cdot \nabla \theta_{z,i}(\tilde{y})^t \cdot \nabla \theta_{z,j}(\tilde{y}) = 1,
\]

which implies that

\[
1 - O(\kappa m^4 \gamma) - O(\kappa m^4 \gamma) \sum_{i \neq j} |\lambda_i \lambda_j| \leq \sum_{i=1}^{d-m} \lambda_i^2 \leq 1 + O(\kappa m^4 \gamma) + O(\kappa m^4 \gamma) \sum_{i \neq j} |\lambda_i \lambda_j|.
\]

Using the above relations concerning \( \lambda_i \)'s, we get an upper bound of the right hand side of \( (36) \) as follows.

\[
\sum_{i=1}^{d-m} \lambda_i^2 + O(\kappa m^4 \gamma) \left( \sum_{i=1}^{d-m} |\lambda_i| \right)^2 \leq 1 + O(\kappa m^4 \gamma) + O(\kappa m^4 \gamma) \sum_{i \neq j} |\lambda_i \lambda_j|
\]
\[
\leq 1 + O(\kappa m^4 \gamma) + O(\kappa m^4 \gamma) \cdot (d^2 + O(d^2(d - m)\kappa m^4 \gamma))
\]
\[
\leq 1 + O(d^2 \kappa m^4 \gamma).
\]

29
Symmetrically, we get a lower bound of the left hand side of (36):

\[
\sum_{i=1}^{d-m} \lambda_i^2 - O(km^4 \gamma) \left( \sum_{i=1}^{d-m} |\lambda_i| \right)^2 \geq 1 - O(d^2 km^4 \gamma).
\]

Thus, we simplify (36) to

\[
(1 - O(d^2 km^4 \gamma)) \cdot \|y - \bar{y}\|^2 \leq \|g_2(y)\|^2 \leq (1 + O(d^2 km^4 \gamma)) \cdot \|y - \bar{y}\|^2. \tag{38}
\]

In other words, \(\|g_2(y)\|\) is a good approximation of the distance from \(y\) to the zero-set of \(g_2\).

Next, we give a lower bound on \(\cos \angle y' \bar{y} \bar{y}'\). Consider the dot product \((y' - y)^t \cdot (\bar{y} - y)\). By expanding \(B_{\varphi,y}^t \cdot (a_y - y)\), we get

\[
y' - y = B_{\varphi,y} \cdot B_{\varphi,y}^t \cdot (a_y - y) = B_{\varphi,y} \cdot (-g_2(y)).
\]

Since \(B_{\varphi,y}\) consists of orthogonal unit column vectors, we get

\[
\|y' - y\| = \|B_{\varphi,y} \cdot (-g_2(y))\| = \|g_2(y)\|. \tag{39}
\]

Therefore,

\[
(y' - y)^t \cdot (\bar{y} - y) = \|\varphi_2\| \cdot \|y - \bar{y}\| \cdot \cos \angle y' \bar{y} \bar{y}' \leq \sqrt{1 + O(d^2 km^4 \gamma)} \cdot \|y - \bar{y}\|^2 \cdot \cos \angle y' \bar{y} \bar{y}'. \tag{40}
\]

Recall that \(\sum_{i=1}^{d-m} \lambda_i \cdot \nabla g_{z,i}(\bar{y})\) is the unit vector \((y - \bar{y})/\|y - \bar{y}\|\). By expanding \((y' - y)^t \cdot (\bar{y} - y)\), we get

\[
(y' - y)^t \cdot (\bar{y} - y) = (B_{\varphi,y} \cdot \varphi_2(y))^t \cdot \|y - \bar{y}\| \cdot \sum_{i=1}^{d-m} \lambda_i \cdot \nabla g_{z,i}(\bar{y})
\]

\[
= \left( \sum_{i=1}^{d-m} \varphi_{z,i}(y) \cdot f_{\nu_i}(y) \right)^t \cdot \|y - \bar{y}\| \cdot \sum_{i=1}^{d-m} \lambda_i \cdot \nabla g_{z,i}(\bar{y})
\]

\[
= \sum_{i=1}^{d-m} \sum_{j=1}^{d-m} \varphi_{z,i}(y) \cdot \lambda_j \cdot f_{\nu_i}(y)^t \cdot \nabla g_{z,j}(\bar{y}) \cdot \|y - \bar{y}\|
\]

\[
= \sum_{i=1}^{d-m} \varphi_{z,i}(y) \cdot \|y - \bar{y}\| \cdot \beta_i,
\]

where \(\beta_i = \lambda_i + \sum_{i=1}^{d-m} (\pm \lambda_i) \cdot O(km^4 \gamma)\) for \(i \in [1, d - m]\). Note the similarity between the \(\beta_i\)'s and the vector in (35). Therefore, \(\|y - \bar{y}\| \cdot \beta_i = \varphi_{z,i}(y) + \|y - \bar{y}\| \cdot \sum_{i=1}^{d-m} (\pm \lambda_i) \cdot O(km^4 \gamma) \geq \varphi_{z,i}(y) - O((d - m)km^4 \gamma) \cdot \|y - \bar{y}\|\) as \(|\lambda_i| \leq 1 + O((d - m)km^4 \gamma)\). Hence,

\[
(y' - y)^t \cdot (\bar{y} - y) \geq \sum_{i=1}^{d-m} \varphi_{z,i}(y)^2 - O((d - m)km^4 \gamma) \cdot \|y - \bar{y}\| \cdot \sum_{i=1}^{d-m} |\varphi_{z,i}(y)|
\]

\[
\geq \|\varphi_2(y)^2 - O((d - m)km^4 \gamma) \cdot \|y - \bar{y}\| \cdot \sqrt{d - m} \cdot \|\varphi_2(y)\|
\]

\[
\geq \|\varphi_2(y)^2 - O((d - m)^{3/2} km^4 \gamma) \cdot \|y - \bar{y}\| \cdot \|\varphi_2(y)\|.
\]

Substituting (38) into the above, we get

\[
(y' - y)^t \cdot (\bar{y} - y) \geq (1 - O(d^2 km^4 \gamma)) \cdot \|y - \bar{y}\|^2.
\]
Combining \((40)\) with the above inequality gives
\[\cos \angle y'y\bar{y} \geq 1 - O(d^2 km^4 \gamma).\]

Finally, consider triangle \(y'y\bar{y}\). By the cosine law, we have
\[\|y' - \bar{y}\|^2 = (\|y' - y\|^2 + \|y - \bar{y}\|^2 - 2\|y' - y\| \|y - \bar{y}\| \cos \angle y'y\bar{y})^{1/2}.\]

By \((38)\) and \((39)\), \(\|y' - \bar{y}\|^2 \leq (1 + O(d^2 km^4 \gamma)) \|y - \bar{y}\|^2\). Therefore,
\[\|y' - \bar{y}\|^2 \leq \|y - \bar{y}\| \left(2 + O(d^2 km^4 \gamma) - 2(1 - O(d^2 km^4 \gamma)) (1 - O(d^2 km^4 \gamma))\right)^{1/2}\]
\[\leq \gamma^{1/4} \|y - \bar{y}\|\]
whenever \(\varepsilon_0\) is small enough so that \(\gamma^{1/4} = O(\varepsilon_0^{1/4})\) cancels the \(O(dm^2 \sqrt{\kappa})\) factor. This requires \(\varepsilon_0\) to decrease as \(d\) and \(\kappa\) increase.  

By combining Lemmas \(7.1\) and \(7.2\), we prove that the projection operator will bring an initial point to a point in \(Z_\varphi \cap \hat{M}\) in the limit.

**Theorem 7.1** Let \(\varphi\) be the function for a uniform \((\varepsilon, \kappa)\)-sample of an \(m\)-dimensional compact smooth manifold \(M\) in \(\mathbb{R}^d\) as specified in Theorem \(6.1\). Define the projection operator \(x_{i+1} = x_i + B_{\varphi,x_i} \cdot B_{\varphi,x_i}^T \cdot (a_{x_i} - x_i)\), where \(a_{x_i} = \sum_{p \in P} \omega(x_i, p) \cdot p\). There exists \(\varepsilon_0 \in (0,1)\) that decreases as \(d\) and \(\kappa\) increase such that if \(\varepsilon \leq \varepsilon_0\), then for any initial point \(x_0\) at distance \(w \gamma\) or less from some sample point, where \(\gamma\) is the input neighborhood radius, the following properties hold.

- \(\lim_{i \to \infty} x_i \in Z_\varphi \cap \hat{M}\), where \(\hat{M}\) is the set of points within a distance of \(\varepsilon\) from \(M\).
- For all \(i > 0\), \(\|x_i - \nu(x_0)\| = O(m^{7/2} \gamma^2) = O(m^{7/2} \varepsilon^2).\)

**Proof.** For any point \(x\), let \(\tilde{x}\) denote the nearest point in \(Z_\varphi \cap \hat{M}\) to \(x\). By Lemma \(7.1\)
\[\|x_1 - \nu(x_0)\| = O(m^{7/2} \gamma^2).\]
Let \(b\) be the nearest point in \(Z_\varphi \cap \hat{M}\) to \(\nu(x_0)\). Since \(\|b - \nu(x_0)\| = O(m^{5/2} \gamma^2)\) by Theorem \(6.1\), triangle inequality implies that for a small enough \(\varepsilon_0\),
\[\|x_1 - \tilde{x}_1\| \leq \|x_1 - b\| \leq \|b - \nu(x_0)\| + \|x_1 - \nu(x_0)\| \leq O(m^{5/2} \gamma^2) + O(m^{7/2} \gamma^2) = O(m^{7/2} \gamma^2).\]

Since \(\|x_1 - \nu(x_0)\| = O(m^{7/2} \gamma^2)\), Lemma \(7.2\) is applicable to \(x_1\). It ensures that \(\|x_2 - \tilde{x}_2\| \leq \|x_2 - \tilde{x}_1\| \leq \gamma^{1/4} \cdot \|x_1 - \tilde{x}_1\| = O(m^{7/2} \gamma^{9/4}),\) which is smaller than \(O(m^{7/2} \gamma^2)\) and so Lemma \(7.2\) is applicable to \(x_2\). Repeating this argument gives
\[\|x_i - \tilde{x}_i\| \leq \|x_i - \tilde{x}_{i-1}\| = O(m^{7/2} \gamma^{(7+i)/4}).\]
This proves that \(\lim_{i \to \infty} x_i \in Z_\varphi \cap \hat{M}\). By triangle inequality,
\[\|x_i - x_{i-1}\| \leq \|x_i - \tilde{x}_{i-1}\| + \|x_{i-1} - \tilde{x}_{i-1}\| = O(m^{7/2} \gamma^{(7+i)/4}) + O(m^{7/2} \gamma^{(6+i)/4}) = O(m^{7/2} \gamma^{(7+i)/4}).\]
Therefore, for a small enough $\varepsilon_0$,
\[
\|x_i - \nu(x_0)\| \leq \sum_{j=2}^{i} \|x_j - x_{j-1}\| + \|x_1 - \nu(x_0)\|
\]
\[
< \sum_{j=2}^{i} O(m^{7/2} \gamma^{(7+j)/4}) + O(m^{7/2} \gamma^2)
\]
\[
= O(m^{7/2} \gamma^2).
\]

8 Conclusion

We define a function $\varphi$ from a uniform $(\varepsilon, \kappa)$-sample of a compact smooth manifold $M$ in $\mathbb{R}^d$ such that the zero-set of $\varphi$ near $M$ is a faithful reconstruction of $M$. Moreover, we give a projection operator that will yield a point on the zero-set near $M$ in the limit by iterative applications. More work is needed to improve the angular error of $O(m^2 \sqrt{\kappa \varepsilon})$, which is weaker than the $O(\varepsilon)$ angular error offered by provably good simplicial reconstructions. It would also be desirable for $\varepsilon$ to depend on $m$ only instead of $d$. Another natural question is how to deal with non-smooth manifolds and non-manifolds.

Acknowledgment The authors would like to thank the anonymous reviewers for helpful comments, pointing out mistakes in an earlier version that we subsequently corrected, and suggesting the removal of some slack in the bounds on Hausdorff distance and angular error.

References

[1] A. Adamson and M. Alexa. Point-sampled cell complexes. ACM Transactions on Graphics, 25 (2006), 671–680.

[2] M. Alexa, J. Behr, D. Cohen-OR, S. Fleishman, D. Levin, and C.T. Silva. Computing and rendering point set surfaces. IEEE Transactions on Visualization and Computer Graphics, 9 (2003), 3–15.

[3] E. Aamari, J. Kim, F. Chazal, B. Michel, A. Rinaldo, and L. Wasserman. Estimating the reach of a manifold. arXiv:1705.04565 [math.ST], 2017.

[4] M. Belkin, J. Sun, and Y. Wang. Constructing Laplace operator from point clouds in $\mathbb{R}^d$. Proceedings of the ACM-SIAM Annual Symposium on Discrete Algorithms, 2009, 1031–1040.

[5] M. Bertalmío, L.T. Cheng, S. Osher, and G. Sapiro. Variational problems and partial differential equations on implicit surfaces. Journal of Computational Physics, 174 (2001), 759–780.

[6] A. Bjorck and G. Golub. Numerical methods for computing angles between linear subspaces. Mathematics of Computation, 27 (1973), 579–594.

[7] J.-D. Boissonnat and F. Cazals. Smooth surface reconstruction via natural neighbor interpolation of distance functions. Computational Geometry: Theory and Applications, 22 (2002), 185–203.
[8] J.-D. Boissonnat and A. Ghosh. Manifold reconstruction using tangential Delaunay complexes. Proceedings of the 26th Annual Symposium on Computational Geometry, 2010, 324–333.

[9] J.-D. Boissonnat, L.J. Guibas and S.Y. Oudot. Manifold reconstruction in arbitrary dimensions using witness complexes. Discrete and Computational Geometry, 42 (2009), 37–70.

[10] J.C. Carr, R.K. Beatson, J.B. Cherrie, T.J. Mitchell, W.R. Fright, B.C. McCallum, and T.R. Evans. Reconstruction and representation of 3D objects with radial basis functions. Proceedings of SIGGRAPH, 2001, 67–76.

[11] S.-W. Cheng and M.-K. Chiu. Tangent estimation from point samples. Discrete and Computational Geometry, 56 (2016), 505–557.

[12] S.-W. Cheng and M.-K. Chiu. Dimension detection via slivers. Proceedings of the 20th Annual ACM-SIAM Symposium on Discrete Algorithms, 2009, 1001–1010.

[13] S.-W. Cheng, T.K. Dey and E.A. Ramos. Manifold reconstruction from point samples. Proceedings of the 16th Annual ACM-SIAM Symposium on Discrete Algorithms, 2005, 1018–1027.

[14] S.-W. Cheng, Y. Wang, and Z. Wu. Provable dimension detection using principal component analysis. International Journal of Computational Geometry and Applications, 18 (2008), 414–440.

[15] T.K. Dey, Z. Dong, and Y. Wang. Parameter-free topology inference and sparsification for data on manifolds. Proceedings of the 28th Annual ACM-SIAM Symposium on Discrete Algorithms, 2017, 2733–2747.

[16] T. K. Dey, J. Giesen, S. Goswami and W. Zhao. Shape dimension and approximation from samples. Discrete and Computational Geometry, 29 (2003), 419–434.

[17] T.K. Dey and J. Sun. An adaptive MLS surface for reconstruction with guarantees. Proceedings of the Eurographics Symposium on Geometry Processing, 2005, 43–52.

[18] J. Dorsey and P. Hanrahan. Digital materials and virtual weathering. Scientific American, 282 (2000), 282–289.

[19] S.C. Eisenstat and I.C.F. Ipsen. Relative perturbation bounds for eigenspaces and singular vector subspaces. Proceedings of the 5th SIAM Conference on Applied Linear Algebra, 1994, 62–66.

[20] D.G. Feingold, R.S. Varga, Block diagonally dominant matrices and generalizations of the Gershgorin circle theorem. Pacific Journal of Mathematics, 12 (1962), 1241–1250.

[21] A. Galántai and Cs.J. Hegedűs. Jordan’s principal angles in complex vector spaces. Numerical Linear Algebra with Applications, 13 (2006), 589–598.

[22] J.B. Greer. An improvement of a recent Eulerian method for solving PDEs on general geometries. Journal of Scientific Computing, 29 (2006), 321–352.

[23] J. Giesen and U. Wagner. Shape dimension and intrinsic metric from samples of manifolds with high codimension. Discrete and Computational Geometry, 32 (2004), 245–267.

[24] G.H. Golub and C.F. van Loan. Matrix Computations, Johns Hopkins University Press, 1996.
[25] M. Hein and J.-Y. Audibert. Intrinsic dimensionality estimation of submanifolds in Euclidean space. *Proceedings of the 22nd International Conference on Machine Learning*, 2005, 289–296.

[26] H. Hoppe, T. DeRose, T. Duchamp, J. McDonald, and W. Stuetzle. Surface reconstruction from unorganized points. *Proceedings of SIGGRAPH*, 1992, 71–78.

[27] R. Horn and C.R. Johnson. *Topics in Matrix Analysis*, Cambridge University Press, 1991.

[28] R. Kolluri. Provably good moving least squares. *ACM Transactions on Algorithms*, 2 (2008), article no. 18.

[29] D. Levin. Mesh-independent surface interpolation. In *Geometric Modeling for Scientific Visualization*, G. Brunett, B. Hamann, K. Mueller, and L. Linsen, eds., Springer-Verlag, 2003.

[30] E. Levina and P.J. Bickel. Maximum likelihood estimation of intrinsic dimension. *Advances in Neural Information Processing Systems*, 17 (2005), 777—784.

[31] J. Liang and H. Zhao. Solving partial differential equations on point clouds. *SIAM Journal on Scientific Computing*, 35 (2013), 1461–1486.

[32] A.V. Little, M. Maggioni, and L. Rosasco. Multiscale geometric methods for data sets I: multiscale SVD, noise and curvature. *Computer Science and Artificial Intelligence Laboratory Technical Report*, MIT-CSAIL-TR-2012-029, CBCL-310, September 8, 2012.

[33] F. Mémoli, G. Sapiro and P. Thompson. Implicit brain imaging. *NeuroImage*, 23 (2004), 179–188.

[34] B. Mendelson. *Introduction to Topology*, Dover Pubblications Inc., New York, third edition, 1990.

[35] J. Miao and A. Ben-Israel. On principal angles between subspaces in \( \mathbb{R}^n \). *Linear Algebra and its Applications*, 171 (1992), 81–98.

[36] T.G. Myers and J.P.F. Charpin. A mathematical model for atmospheric ice accretion and water flow on a cold surface. *International Journal of Heat and Mass Transfer*, 47 (2004), 5483–5500.

[37] T.G. Myers, J.P.F. Charpin, and S.J. Chapman. The flow and solidification of a thin fluid film on an arbitrary three-dimensional surface. *Physics of Fluids*, 14 (2002), 2788–2803.

[38] S.J. Ruuth and B. Merriman. A simple embedding method for solving partial differential equations on surfaces. *Journal of Computational Physics*, 227 (2008), 1843–1964.

[39] B. Schölkopf, A. Smola, and K.-R. Müller. Nonlinear component analysis as a kernel eigenvalue problem. *Neural Computation*, 10 (1998), 1299–1319.

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

[41] G. Turk. Generating textures on arbitrary surfaces using reaction-diffusion. *Proceedings of SIGGRAPH*, 1991, 289–298.

[42] H. Wendland. Piecewise polynomial, positive definite and compactly supported radial functions of minimal degree. *Advances in Computational Mathematics*, 4 (1995), 389–396.
[43] A. Witkin and M. Kass. Reaction-diffusion textures, \textit{Proceedings of SIGGRAPH}, 1991, 299–308. 25 (1991), 299–308.