Approximating the Riemannian Metric from Point Clouds via Manifold Moving Least Squares

Barak Sober
Ingrid Daubechies
Robert Ravier

July 21, 2020

Abstract

The approximation of both geodesic distances and shortest paths on point cloud sampled from an embedded submanifold $M$ of Euclidean space has been a long-standing challenge in computational geometry. Given a sampling resolution parameter $h$, state-of-the-art discrete methods yield $O(h^2)$ approximations. In this paper, we investigate the convergence of such approximations made by Manifold Moving Least-Squares (Manifold-MLS), a method that constructs an approximating manifold $M^h$ using information from a given point cloud that was developed by Sober & Levin in 2019. They showed that the Manifold-MLS procedure approximates the original manifold with approximation order of $O(h^k)$ provided that the original manifold $M \in C^k$ is closed, i.e. $M$ is a compact manifold without boundary. In this paper, we show that under the same conditions, the Riemannian metric of $M^h$ approximates the Riemannian metric of $M$; i.e., $M$ is nearly isometric to $M^h$. Explicitly, given points $p_1, p_2 \in M$ with geodesic distance $\rho_M(p_1, p_2)$, we show that their corresponding points $p^h_1, p^h_2 \in M^h$ have a geodesic distance of $\rho_M^h(p^h_1, p^h_2) = \rho_M(p_1, p_2) + O(h^{k-1})$. We then use this result, as well as the fact that $M^h$ can be sampled with any desired resolution, to devise a naive algorithm that yields approximate geodesic distances with rate of convergence $O(h^{k-1})$. We show the potential and the robustness to noise of the proposed method on some numerical simulations.

1 Introduction

The concept of a manifold is fundamental to models in numerous research fields (e.g., mathematical physics, linear algebra, computer graphics, computer aided design). In recent years, manifold-based modeling has also permeated into general high dimensional data analysis as a method of circumventing the curse of dimensionality (e.g., signal processing, computer vision, biomedical imaging). These “data-driven” models, however, do not follow from a
set of equations or some other formal representation, but rather emerge from a general intuition/observation that data has low intrinsic dimensions, at least locally (see, for example, [2, 4, 17, 18, 22]). The lack of formalism means that both explicit coordinate charts and Riemannian metric structure are usually not available; in practice, the only information available is the collection of samples, i.e., point clouds.

An essential tool for data analysis in modern data-driven scenarios is the ability to measure intrinsic geodesic distances. The approximation of geodesic distances and geodesic paths on point clouds and their manifolds has been studied extensively in computer graphics literature, primarily in the context of 3D point clouds that represent 2D surfaces. The focus of these studies is on performing efficient approximations of high approximation order (for example, [7,13,19]). In most of these methods there is a strong assumption that a meshing of the manifold exists, which is a nontrivial requirement for general point clouds. To the best of our knowledge, the highest order of convergence proven in the literature is $O(h)$ (for a resolution parameter $h$) shown in the Fast Marching Method of Kimmel and Sethian [7]. Numerically, the work in [9] reports $O(h^2)$ convergence rates in numerical simulations. Nonetheless, much less attention have been given to the general manifold case, where a triangulation cannot be achieved easily. Memoli and Sapiro [11] proposed an approach for general submanifold of $\mathbb{R}^D$ with convergence rates of $O(\sqrt{h})$.

Many works deal with the lack of triangulation by instead using standard shortest path algorithms, such as the well-known Dijkstra algorithm [5], after imposing a graph structure on the point cloud based on distance information. Though nonexhaustive, see, for example, [10,14–16,22]. However, approximating geodesic curves and distances from discrete samples in such a way may result in poor approximations if the sampling resolution is insufficient. Figure 1. gives an example of such a failure. As discussed in Section 4, in such cases, it is possible (though by no means guaranteed) that the Euclidean distance between points on the manifold serve as a better approximation to the geodesic distance.

Figure 1: Shortest path from discrete samples of a 2-sphere. (A) The original manifold $\mathcal{M}$; the shortest path computed between $p_1, p_2$ (marked by red triangles) on discrete samples of the manifold (the path marked by the red line). (B) The sample set of $\mathcal{M}$ and the shortest path connecting points $p_1, p_2$ marked by the red triangles.
The main contribution of the current work is theoretical. In [21], Sober and Levin proposed a procedure called Manifold Moving Least Squares (Manifold-MLS) that defined an implicit manifold from a given point cloud. Explicitly, given a Point Cloud sampled from a manifold \( M \in \mathbb{C}^k \) \((k \geq 2)\) with sampling resolution \( h \) (such as the fill distance; see (3) below), they introduce a moving least squares (MLS) procedure to create an approximating manifold \( M^h \). They show that \( P : M \rightarrow M^h \), the Manifold-MLS mapping from \( M \) to \( M^h \), is a diffeomorphism. Similar to standard MLS approximations, the approximating manifold exists as an abstract object and can be realized at a point based upon a local neighborhood, requiring no global processing. Sober and Levin also show that the Hausdorff distance between \( M^h \) and \( M \) is bounded by \( O(h^k) \).

In this paper, we show that the Riemannian metric of \( M^h \) approximates the Riemannian metric of \( M \) as well. Specifically, let \( p_1, p_2 \in M \) with geodesic distance \( \rho_M(p_1, p_2) \) then

\[
\rho_{M^h}(p^h_1, p^h_2) = \rho_M(p_1, p_2) + O(h^{k-1}),
\]

where \( p^h_1, p^h_2 \in M^h \) are \( O(h^k) \) away from \( p_1, p_2 \), respectively, and \( \rho_{M^h}(p^h_1, p^h_2) \) is the Riemannian distance between \( p^h_1, p^h_2 \) on \( M^h \) (see Theorem 1). We prove this by showing that \( P \) is actually an approximate isometry given that \( M \) is sampled densely enough. In other words, the differential of \( P \) is close to the identity map. In Section 4, we show how this result can be utilized by developing a simple brute-force algorithm that significantly improves the estimation of geodesics and their distances. The proposed algorithm can be shown to converge with a rate of \( O(h^{k-1}) \), a significant improvement on other results. We provide experimental evidence showing the potential of using the Manifold-MLS as well as its robustness to noise. We believe that the current work will motivate the future development of novel, efficient ways of approximating geodesics on point clouds.

2 Definitions, Assumptions and the Approximation

The analysis carried in Section 3 builds upon the definition of the Manifold-MLS as introduced in [21], which culminated in a bound on the difference between the approximated manifold \( M \) and its approximant with respect to the Hausdorff distance. However, since we wish to analyze the convergence of geodesic paths, we ultimately need to bound the distortion of first derivatives as well. Therefore, we will start by explaining some recent results in derivative estimation of standard MLS and carry on to formally define the Manifold-MLS approximation along with the sampling assumptions that will be used in the proofs below.

2.1 Moving Least-Squares Approximation of Derivatives

A key concept we aim to use in our analysis is the approximation property of the natural derivatives of the MLS, as reported in [12].

Let \( f : \Omega \rightarrow \mathbb{R} \) be sampled at locations \( X = \{x_i\}_{i=1}^n \). Then, we define the local polynomial \( \pi^*(x \mid \xi) \) by

\[
\pi^*(x \mid \xi) = \arg\min_{\pi \in \Pi^d_{k-1}} \sum_{i=1}^n \left| f(x_i) - \pi(x_i) \right|^2 \theta(\|\xi - x_i\|),
\]

where \( \Pi^d_{k-1} \) is the set of degree-\( d \) polynomials in \( \mathbb{R}^d \). This is a common form of MLS approximation and is known to be a good approximation to the true derivative of \( f \) at \( \xi \).
where $\Pi_{k-1}^d$ is the space of polynomial of degree $\leq (k-1)$ over a $d$-dimensional domain, $\theta$ is a locally supported (or fast decaying) weight function, and $\|\cdot\|$ denotes the Euclidean norm. We define the MLS function by

$$s_{f,X}^{\text{MLS}}(x) \triangleq \pi^*(0 \mid x).$$

(2)

The theory developed in [12] shows that the $\alpha$-order derivatives of $s_{f,X}^{\text{MLS}}(x)$ approximate the derivatives of $f(x)$ with convergence rates of $O(h^{k-|\alpha|})$ for $f, \theta \in C^k$ and $|\alpha| \leq k - 1$. However, the analysis of [12] is described in a more general setting, where the approximation order is derived for arbitrary fractional Sobolev norm and with a domain with Lipschitz boundary. Since we are interested in the ordinary derivatives of the MLS approximation (more precisely, just first order derivatives) in a boundaryless domain, in this subsection we describe and justify the main result that we wish to use in this paper, which is presented in (13): the MLS approximant’s first order natural derivatives have optimal convergence rates to $f$’s derivatives.

Let $X = \{x_i\}_{i=1}^n$ be samples of a domain $\Omega \subset \mathbb{R}^d$. Furthermore, we assume that the samples are a quasi-uniform sample set.

**Definition 1** (quasi-uniform sample set). A set of data sites $X = \{x_1, \ldots, x_n\}$ is said to be quasi-uniform with respect to a domain $\Omega$ and a constant $c_{\text{qu}} > 0$ if

$$\delta_X \leq h_{X,\Omega} \leq c_{\text{qu}} \delta_X,$$

where $h_{X,\Omega}$ is the fill distance

$$h_{X,\Omega} = \sup_{x \in \Omega} \min_{x_i \in X} \|x - x_i\|,$$

(3)

and $\delta_X$ is the separation radius defined by

$$\delta_X := \frac{1}{2} \min_{i \neq j} \|x_i - x_j\|.$$

Throughout the paper we will denote for short $h := h_{X,\Omega}$.

Then, as described in (1), we define the MLS local polynomial $\pi^*(x \mid \xi)$ as

$$\pi^*(x \mid \xi) = \arg\min_{\pi \in \Pi_{k-1}^d} \sum_{i=1}^n |f(x_i) - \pi(x_i)|^2 \theta_h(\|\xi - x_i\|),$$

(4)

where now $\theta_h$ is a locally supported weight function consistent across scales (i.e., $\theta_h(th) = \Phi(t)$). Furthermore, we assume that $\theta_h \in C^k$ and has a compact support of size $sh$, and we denote by $I(x) = \{i \mid \|x - x_i\| \leq sh, x_i \in X\}$. Levin showed in [8] that, given enough samples in $I(x)$ (i.e., if the least-squares matrix is invertible), the MLS approximation defined in (2) is given by

$$s_{f,X}^{\text{MLS}}(x) = \sum_{i \in I(x)} a_i^*(x)f(x_i),$$

(5)
where the coefficients $a_i^*(x)$ are determined by minimizing the quadratic form

$$\sum_{i \in I(x)} a_i(x)^2 \frac{1}{\theta_h(\|x - x_i\|)}$$

under the constraints

$$\sum_{i \in I(x)} a_i(x) \pi(x_i) = \pi(x), \ \forall \pi \in \Pi_{k-1}(\mathbb{R}^d).$$

In addition, it has been shown that $s_{f,X}^{\text{MSE}}(x) \in C^k$ [8, 24], and that it has high approximation order in both the $L_\infty$ and Sobolev sense (see [8, 24] for the $L_\infty$ analysis, [1] for a first order Sobolev, and [12] for arbitrary fractional Sobolev norm). The key idea behind these results stems from the fact that $s_{f,X}$ is exact for polynomials. Namely,

$$\sum_{i \in I(x)} a_i^*(x) \pi(x_i) = \pi(x)$$

for $\pi \in \Pi_{k-1}(\mathbb{R}^d)$. Thus,

$$E_{f,X}(x) = s_{f,X} - f(x)$$

$$= s_{f,X} - \pi(x) + \pi(x) - f(x)$$

$$= \sum_{i \in I(x)} a_i^*(x)(f(x_i) - \pi(x_i)) + \pi(x) - f(x).$$

Then, for a fixed but arbitrary $\hat{x} \in \Omega$ we can choose $\pi$ to be the Taylor approximation of $f(x)$ around $\hat{x}$ and then we get for all $x \in B_{sh}(\hat{x})$

$$|E_{f,X,B_{sh}(\hat{x})}(\hat{x})| \leq \sum_{i \in I(\hat{x})} |a_i^*(x)| \left| (f(x_i) - \pi(x_i)) \right| + \left| \pi(x) - f(x) \right|$$

$$\leq \max_{x \in B(\hat{x},sh)} \left| \pi(x) - f(x) \right| \left( \sum_{i \in I(\hat{x})} |a_i^*(x)| + 1 \right),$$

where the number of indices in $I(x)$ are bounded by $N = \rho s^d$ for some constant $\rho$ (namely, the bound is independent of $h$; see Lemma A.1 in the Appendix). Since the $a_i^*$ depend smoothly on the location of their neighbors in a $sh$-sized neighborhood, the summation can be bounded by a constant independent of $h, x,$ and the specific sample set $X$. See the Appendix for a more rigorous discussion of this fact. Since we chose $\pi$ to be the $(k-1)$ degree Taylor expansion of $f(x)$ we get

$$\max_{x \in B(\hat{x},sh)} \left| f(x) - \pi(x) \right| \leq c \left| f \right|_{C^k(B_{sh}(\hat{x}))} h^k,$$

where $\left| f \right|_{C^k(B_{sh}(\hat{x}))}$ is the semi-norm defined by

$$\left| f \right|_{C^k(B_{sh}(\hat{x}))} \overset{\Delta}{=} \max_{\left| \alpha \right| = k} \max_{x \in B_{sh}(\hat{x})} |\partial^\alpha f(x)|,$$
where \( \alpha = (\alpha_1, \ldots, \alpha_d) \) is a multi-index of order \( |\alpha| = k \) and

\[
\partial^\alpha f(x) = \partial^\alpha_1 \partial^\alpha_2 \cdots \partial^\alpha_d f(x).
\]

As a result, we achieve the bound

\[
|s_{f,X}(x) - f(x)| = |E_{f,X,B_{sh}(\hat{x})}(x)| \leq C |f|_{C^k(B_{sh}(\hat{x}))} h^k.
\]  

(9)

If, on the other hand, one wishes to measure the error in the derivatives we get

\[
\partial^\alpha E_{f,X,B}(x) = \sum_{i \in I(x)} (\partial^\alpha a^*_i(x))(f(x_i) - \pi(x_i)) + \partial^\alpha (\pi(x) - f(x)).
\]

Note that in the current case, \( \partial^\alpha \) is scalar-valued as \( f : \mathbb{R}^d \rightarrow \mathbb{R} \). Then, for a fixed but arbitrary \( \hat{x} \in \Omega \) we can choose \( \pi \) to be the Taylor approximation of \( f(x) \) around \( \hat{x} \) and get

\[
|\partial^\alpha E_{f,X,B}(\hat{x})| \leq \max_{x \in B(\hat{x},sh)} |f(x) - \pi(x)| \sum_{i \in I(\hat{x})} |\partial^\alpha a^*_i(\hat{x})| + |\partial^\alpha \pi(x) - \partial^\alpha f(x)|.
\]  

(10)

Since \( \pi \) is the Taylor expansion of \( f(x) \) at \( \hat{x} \) and \( \|x - \hat{x}\| \leq sh \) we get that

\[
|\partial^\alpha \pi(x) - \partial^\alpha f(x)| \leq c_1 |f|_{C^k(B_{sh}(\hat{x}))} h^{k - |\alpha|}.
\]

Then, using arguments similar to Theorem 3.11 in [12] we can achieve for all \( x \in B_{sh}(\hat{x}) \)

\[
|\partial^\alpha a^*_i(x)| \leq c_2 h^{-|\alpha|}.
\]  

(11)

We can combine (7) and the above bounds with the fact that the number of indices in \( I(x) \) is bounded by \( N = \rho s^d \) to get the following lemma.

**Lemma 1.** Let \( f \in C^k(\mathbb{R}^d) \) be a scalar-valued function. Let \( X \subset \Omega \) be a quasi-uniform unbounded set. Suppose \( \theta_h \in C^k \) is supported on \( [0, sh] \). Then, if for a fixed but arbitrary \( \hat{x} \in \Omega \) the set \( X \) is \( \Pi_{k-1}(\mathbb{R}^d) \)-unisolvent (i.e., the least-squares matrix resulting from (4) is invertible), we get for all \( x \in B_{sh}(\hat{x}) \)

\[
|\partial^\alpha s_{f,X}(\hat{x}) - \partial^\alpha f(\hat{x})| = |\partial^\alpha E_{f,X}(\hat{x})| \leq C |f|_{C^k(\hat{x})} h^{k - |\alpha|},
\]  

(12)

where \( C \) is some constant independent of \( f \) and \( h \).

**Remark 1.** Lemma 1 is adapted from Theorem 3.12 of [12], which handles a much more general case.

**Remark 2.** Wendland invest a considerable effort in showing what is the tightest support for \( \theta_h \) such that the MLS problem has a unique solution (i.e., that the least-squares matrix is invertible). In order to achieve that he introduces the Interior Cone Condition [24], which is then inherited into the work of Mirzaei [12]. The main challenge in these investigations is the fact that the domain has boundary. Nevertheless, in the work of Levin [8] and the following Manifold-MLS works [20, 21] as well as in the current paper the domain is without boundary, and thus, the Interior Cone Condition is not necessary (see the Appendix for a more detailed discussion).
Although the above discussion was conducted on scalar-valued function \( f \), a similar line of arguments will suit the case where \( f : \mathbb{R}^d \rightarrow \mathbb{R}^D \). Note that the coefficients \( a_i^*(x) \) depend only on the sites of the sample set \( X \) and not on the values. Thus, \( a_i^*(X) \) are the same for all \( D \) coordinates in the target domain, and we can write

\[
\partial^\alpha E_{f,X}(x) = \sum_{i \in I(x)} (\partial^\alpha a_i^*(x))(f(x_i) - \pi(x_i)) + \partial^\alpha (\pi(x) - f(x)),
\]

where \( a_i^*(x) : \mathbb{R}^d \rightarrow \mathbb{R} \) and \( \pi : \mathbb{R}^d \rightarrow \mathbb{R}^D \). Then, the error given in (10) for a fixed but arbitrary \( \hat{x} \in \Omega \) and \( \pi \) being the coordinate-wise Taylor expansion of \( f \) around \( \hat{x} \) can be written for the \( \ell \)th coordinate (\( 1 \leq \ell \leq D \)) as

\[
\partial^\alpha E_{f,X}^\ell(x) = \sum_{i \in I(x)} (\partial^\alpha a_i^*(x))(f^\ell(x_i) - \pi^\ell(x_i)) + \partial^\alpha (\pi^\ell(x) - f^\ell(x)).
\]

Then using the same considerations on a coordinate-wise level we arrive at

**Lemma 2.** Let \( f : \mathbb{R}^d \rightarrow \mathbb{R}^D \) and \( f \in C^k \). Let \( X \subset \Omega \) be a quasi uniform unbounded set. Suppose \( \theta_h \in C^k \) is supported on \([0, sh]\) with vanishing derivatives at the end points. Then, if for a fixed but arbitrary \( \hat{x} \in \Omega \) the set \( X \) is \( \Pi_{k-1}(\mathbb{R}^d) \)-unisolvent (i.e., the least-squares matrix resulting from (4) is invertible), we get for all \( 1 \leq \ell \leq D \) and all \( x \in B_{sh}(\hat{x}) \)

\[
|\partial^\alpha s_{f,X}^\ell(x) - \partial^\alpha f^\ell(x)| = |\partial^\alpha E_{f,X}^\ell(x)| \leq C \|f\|_{C^k(B_{sh}(\hat{x}))} h^{k-|\alpha|},
\]

where

\[
\|f\|_{C^k(B_{sh}(\hat{x}))} \Delta \max_{1 \leq \ell \leq D} |f^\ell|_{C^k(B_{sh}(\hat{x}))},
\]

and \( C \) is some constant independent of \( f \) and \( h \).

### 2.2 Manifold Moving Least-Squares (Manifold-MLS)

We now turn to presenting the Manifold-MLS [21], which generalizes the MLS to approximate submanifolds of \( \mathbb{R}^D \) based upon scattered data. Let us begin with the sampling assumptions.

**Clean Sampling Assumptions**

1. \( \mathcal{M} \subset C^k \) \((k \geq 2)\) is a closed (i.e., compact and boundaryless) submanifold of \( \mathbb{R}^D \).

2. \( R = \{r_i\}_{i=1}^n \subset \mathcal{M} \) is a quasi-uniform sample set with respect to the domain \( \mathcal{M} \) (see Definition 1); we denote the fill-distance of our sample \( h_{R,M} \) by \( h \) for short.

Given a point \( r \) near \( \mathcal{M} \) the Manifold Moving Least-Squares (Manifold-MLS) projection of \( r \) is defined through two sequential steps:

1. Find a local \( d \)-dimensional affine space \((q(r), H(r))\) that approximates the sampled points. Explicitly, \( H(r) \) is a linear space and the origin is set to \( q(r) \). So, \( \{q+h| h \in H\} \) is the affine subspace that approximates the sampled points.
2. Define the projection of $r$ using a local polynomial approximant $\pi : H \simeq \mathbb{R}^d \to \mathbb{R}^D$ of $\mathcal{M}$ over the new coordinate system. Explicitly, we denote by $x_i$ the projections of $r_i - q$ onto $H$ and then define the samples of a function $\varphi$ by $\varphi(x_i) = r_i$.

**Step 1 - the local coordinate system**

Let

$$J_1(q, H \mid r) = \sum_{i=1}^{n} d(r_i - q, H)^2 \theta_1(\|r_i - q\|),$$

(15)

where $\theta_1(t)$ is a fast decaying radial weight function. We wish to find a $d$-dimensional affine space $H(r)$, and a point $q(r)$ on $H(r)$, such that

$$(q(r), H(r)) = \arg\min_{q \in \mathbb{R}^D, H \in \text{Gr}(d, D)} J_1(q, H \mid r)$$

(16)

under the constraints

1. $r - q \perp H$
2. $q \in B_\mu(r)$
3. $\#(R \cap B_h(q)) \neq 0$,

where $d(r_i - q, H)$ is the Euclidean distance between $r_i - q$ and the linear subspace $H$, $B_\mu(r)$ is an open ball of radius $\mu$ around $r$ limiting the Region Of Interest (ROI), and $h$ is the fill distance $h_{\mathcal{M}, R}$ of the sampling assumptions. Note, that $\mu$ must be limited by the injectivity radius of $\mathcal{M}$ (i.e., manifold’s reach), but we omit the relevant discussions from our paper and instead refer interested readers to [20,21].

**Step 2 - the weighted least-squares**

Upon obtaining a local coordinate system we now pronounce the manifold in $B_\mu(r)$, the restricted ROI, as a function $\varphi : H(r) \to \mathbb{R}^D$. The approximation in the local coordinates is performed by means of finding a polynomial $\pi^*(x) \in \Pi_{k-1}^d$ which minimizes

$$J_2(\pi(x) \mid q, H) = \sum_{i=1}^{n} \|\pi(x_i) - r_i\|^2 \theta_2(x_i),$$

(17)

where $x_i = P_H(r_i - q)$ is the orthogonal projection of $r_i - q$ onto $H \in \text{Gr}(d, D)$, and $\theta_2(t)$ is a fast decaying radial weight function consistent across scales (i.e., $\theta_2(x_i - q) = \theta_h(\|x_i - q\|)$ and $\theta_h(th) = \Phi(t)$). That is,

$$\pi^*(x \mid r) = \arg\min_{\pi \in \Pi_{k-1}^d} J_2(\pi(x) \mid q(r), H(r)).$$

(18)

We define the Manifold-MLS projection as

$$P^h_k(r) = \pi^*(0 \mid r).$$

(19)

Using this procedure we can now define the approximating manifold as

$$\mathcal{M}^h = \{P^h_k(p) \mid \forall p \in \mathcal{M}\},$$

(20)
Figure 2: An illustration of a curve $\gamma^h(t)$ on $\mathcal{M}^h$ and the corresponding curve $\gamma_h(t)$ on $\mathcal{M}$.

where $h = h_{M,R}$ is the fill-distance of the sample-set (see Fig. 2).

As shown in [21], given that $\theta_1, \theta_2$ are smooth and $h$ small enough (see the Injectivity Conditions of Section 3.1), we achieve that $\mathcal{M}^h$ is a smooth sub-manifold of $\mathbb{R}^D$ and that

$$\|\mathcal{M}^h - \mathcal{M}\|_H \leq M_0 \cdot h^k,$$

(21)

where $\|\cdot\|_H$ denotes the Hausdorff norm and $M$ is some constant.

**Remark 3.** Note [21] uses the notion of a $h$-$\rho$-$\delta$ sample set rather than a quasi-uniform sample set assumption of Definition 1. However, we show in Appendix A that these two definitions are the same. Therefore, throughout the paper we refer only to the quasi-uniform assumption and use the results of [21] as-is.

**Remark 4.** We emphasize that, throughout the exposition of the Manifold-MLS, we have denoted by $h$ the fill distance $h_{R,M}$. From here on, $h$ will be used in this meaning, and whenever we wish to define a fill-distance with respect to a different sample or a different domain we will denote it explicitly.

### 2.3 Main Result

Computing geodesic distances on clouds of points are done regularly through finding shortest paths on distance graphs (e.g., by means of Dijkstra algorithm). Given a limited sample of $\mathcal{M}$ the computation of shortest paths may be problematic as shown in Fig. 1. Nonetheless, the Manifold-MLS approximant $\mathcal{M}^h$ can be sampled in any desired resolution, by taking points $r$ in a neighborhood of $\mathcal{M}$ and computing $P^h_k(r)$. Thus, the geodesic distances of $\mathcal{M}^h$ can be attained up to machine precision. The main contribution of the current paper is in showing that the geodesic distances (as well as paths) computed on $\mathcal{M}^h$ approximate geodesic distances of $\mathcal{M}$ up to $O(h^{k-1})$. 

9
Theorem 1. Let the Sampling Assumptions of Section 2.2 hold and let the Injectivity Conditions of Section 3.1 hold. Let \( \theta_2(x) = \theta_h(\|x - q\|) \) of (17) be such that \( \lim_{t \to 0} \theta_h(t) = \infty \) (i.e., the Manifold-MLS is interpolatory). Let \( \rho_M, \rho_{M^h} \) denote the natural distances on \( M \) and \( M^h \) correspondingly, and let \( p_1^h = \mathcal{P}_k^h(p_1) \) and \( p_2^h = \mathcal{P}_k^h(p_2) \). Then, there exists \( h_0 \) such that for all \( h \leq h_0 \) we have

\[
\rho_{M^h}(p_1^h, p_2^h) = \rho_M(p_1, p_2) + O(h^{k-1})
\]

The complete proof of this theorem can be found in Section 3 below.

3 Proof of Theorem 1

3.1 The Approximate Isometry

As explained in Section 2.2 the first step of the Manifold-MLS is aimed at finding a local coordinate system \((q, H)\) by which we can describe the manifold locally as the image of a function \( \varphi : H \simeq \mathbb{R}^d \to \mathbb{R}^D \). In other words, \( \varphi(x) \) can be viewed as a local parametrization of the manifold. Then, in the second step of the Manifold-MLS we perform locally weighted least-squares polynomial estimation of this function \( \varphi \). Using this local polynomial estimation of the local parametrization, we show that the metric tensor of \( M^h \) resembles the one of \( M \).

The one-to-one relationship: Let \( p^h \) be some point in \( M^h \). Then, by its construction, we know that there exists \( p_h \in M \) such that \( \mathcal{P}_k^h(p_h) = p^h \) along with the corresponding coordinate domain \((q(p_h), H(p_h))\). In Lemma 3 below, we show the existence of \( \varphi_{p_h} : H(p_h) \to M \) such that \( \varphi_{p_h}(0) = p_h \). Furthermore, we also show that if the Injectivity Conditions presented in [21] hold, then this is a one-to-one correspondence. Translating these conditions to our sampling case we get

Injectivity Conditions

1. The functions \( \theta_1(t) \) and \( \theta_h(t) \) of (16) and (18) are monotonically decaying and supported on \([0, c_1 h]\), where \( c_1 > 3 \).

2. Suppose that \( \theta_1(c_2 h) > c_3 > 0 \) and \( \theta_h(c_2 h) > c_3 > 0 \), for some constant \( c_2 < c_1 \).

3. Set \( \mu = rch(M)/2 \) in constraint 2 of (16), where \( rch(M) \) denotes the manifold’s reach (i.e., the injectivity radius; see [6]).

Lemma 3. Let the Sampling Assumptions of Section 2.2 and the Injectivity Conditions of Section 3.1 hold. Then, for all \( h \leq h_0 \) and all \( p \in M \) we have:

1. There exists a unique \( p^h \in M^h \) and a unique coordinate domain \((q(p), H(p))\) \( \in \mathbb{R}^D \times Gr(d, D) \).
2. There exist neighborhoods $W_p \subset \mathcal{M}, U \subset H(p)$ of $p, q(p)$ correspondingly such that
\[ \varphi_p : U \to \mathcal{M} \subset \mathbb{R}^D, \quad \varphi^h_p : U \to \mathcal{M}^h \subset \mathbb{R}^D \]
and
\[ \varphi_p[U] = W_p \subset \mathcal{M}, \quad \varphi^h_p[U] = W^h \subset \mathcal{M}^h \]

3. Furthermore,
\[ \varphi_p(0) = p, \quad \varphi^h_p(0) = \pi^*(0 \mid p) = p^h \]

**Proof.** Under the Sampling Assumptions and the Injectivity Conditions, the conditions for Lemmas 4.13 and 4.18 of [21] hold. Therefore, we know that both $(q(p), H(p)) : \mathcal{M} \to \mathbb{R}^D \times Gr(d, D)$ and $\mathcal{P}_k^h : \mathcal{M} \to \mathcal{M}^h$ are injective. This proves item 1 of the lemma. From Lemma 4.4. of [21] we know that $H(p)$ converges to $T_p\mathcal{M}$ as $h \to 0$. Thus, there exists a small enough $h$ such that locally $\mathcal{M}$ is the image of the following function:

\[ \varphi_p : H(p) \to W_p \subset \mathcal{M} \]
\[ x \mapsto q(p) + \underbrace{x}_{\text{origin}} + \underbrace{\phi_p(x)}_{\text{"tangential"}} + \underbrace{\tilde{\phi}_p(x)}_{\text{"normal"}}, \quad (22) \]

where $\phi_p(x) \in H(p)\perp$. The function $\varphi_p$ comprises three parts: a shift to the local origin $q(p)$, a movement along $H(p)$ (this is what we mean by “tangential”) and a movement along $H(p)\perp$ (normal direction to $H(p)$); see Fig. 3 for an illustration.

In addition, since $\mathcal{M}^h$ is constructed by a local polynomial above $H$, which is smoothly varying, we can create $\varphi^h_p$, a parametrization from $H$ to $\mathcal{M}^h$ with similar components.

From Constraint 1 of (16) we know that

\[ p - q(p) \perp H(p). \]

Thus, the tangential component is null and we have

\[ \varphi_p(0) = p. \]

Furthermore from (19) we know that

\[ \pi^*(0 \mid p) = p^h. \]

Thus, we achieve Results 2 and 3 as required. \qed

In other words, under the Injectivity Conditions, $\pi^*(x \mid p)$ is a local weighted polynomial least-squares approximation to $\varphi_p$. Thus, from Lemma 2 we get that for $x \in B_{sh}(0) \subset H(p)$

\[ \|\varphi^h_p(x) - \varphi_p(x)\|_\infty \leq ch^k \|\varphi_p\|_{C^k(B_{sh}(0))}, \quad (23) \]
Figure 3: Illustration of $\varphi_p$: The $xy$-plane is $H(p)$; The local origin $q(p)$, which is mapped by $\phi_p$ to $p$ is marked by the red dot; the vector $x$ represents a “tangential” movement; and $\phi_p(x)$ is a normal movement.

where

$$
\|\varphi_p\|_{C^k(B_{sh}(0))} = \max_{|\alpha|=k} \left| \partial^\alpha \varphi_p(x) \right|.
$$

Furthermore, as $M$ is compact and the dependency of $\varphi_p$ on $p$ is smooth (resulting from the fact that $H(p)$ vary smoothly – see Theorem 4.11 in [21]) we can bound $\|\varphi_p\|_{C^k(B_{sh}(0))}$ independently of $p \in M$ by a constant and get

$$
\|\varphi_p^h(x) - \varphi_p^h(0)\|_\infty \leq C_0 h^k,
$$

for all $x \in B_{sh}(0)$. Note, that this implies that

$$
\|p^h - p\|_\infty = \|\varphi_p^h(0) - \varphi_p(0)\|_\infty \leq C h^k.
$$

In Lemma 4 below, we extend this result to the derivatives as well. Namely, we show that for $x \in B_{sh}(0)$

$$
\left\| \partial_v \varphi_p^h(x) - \partial_v \varphi_p(x) \right\|_\infty \leq C_1 h^{k-1}.
$$

Lemma 4. Let the Sampling Assumptions of Section 2.2 and the Injectivity Conditions of Section 3.1 hold, and let $\theta_2(x) = \theta_h(\|x - q\|)$ of (17) be such that $\lim_{t \to 0} \theta_h(t) = \infty$ (i.e., $M^h$ interpolates $M$ at the samples $r_i$). Then, there exists a constant $h_0$ such that for all $h \leq h_0$, all $p \in M$, any direction $v \in \mathbb{R}^d$, and all $x \in B_{sh}(0) \subset H(p)$

$$
\left\| \partial_v \varphi_p(x) - \partial_v \varphi_p^h(x) \right\|_\infty \leq C_1 h^{k-1},
$$

$$
\left\| \partial_v \pi_u(x \mid p) - \partial_v \varphi_p(x) \right\|_\infty \leq C_2 h^{k-1},
$$

$$
\left\| \partial_v \pi_u(x \mid p) - \partial_v \varphi_p^h(x) \right\|_\infty \leq C_3 h^{k-1},
$$

(25)
where \(C_1, C_2, C_3\) are constants independent of \(v\) or \(p\), and \(\partial_v \pi^*(x \mid p) = \frac{\partial}{\partial v} \bigg|_{x=x^*} \pi^*(x \mid p)\) denotes the directional derivative of the polynomial in the direction \(v\) with respect to the first input evaluated at \(x \in B_{sh}(0)\); i.e., taking \(p\) to be constant.

**Proof.** Let us fix an arbitrary point \(p \in \mathcal{M}\). According to Lemma 3 we have the local coordinate system \((q,H) = (q(p),H(p))\) and the local functions \(\varphi_p : H \rightarrow \mathcal{M}\) and \(\varphi_p^h : H \rightarrow \mathcal{M}^h\), defined on some open ball around the origin \(B_0(s) \subset H \simeq \mathbb{R}^d\). Furthermore, we define the sample sites on \(H\) by means of

\[
X = \{ x_i = P_H(r_i - q) \mid \| r_i - p \| \leq \mu \}_{i=1}^n \cap B_{sh}(0),
\]

and by Lemma 6 we know that \(h_{X,B_0(sh)} \leq c \cdot h\) for some constant \(c\).

Under these settings we can now define \(s_{\varphi,X}(x), s_{\varphi^h,X}(x)\), the MLS approximants of \(\varphi_p\) and \(\varphi_p^h\) respectively. As we know that \(\lim_{t \to 0} \vartheta(t) = \infty\), we get that the Manifold-MLS is interpolatory (see Section 3.2 in [20]) and so

\[
\varphi_p(x_i) = r_i = \varphi_p^h.
\]

Therefore, we get that

\[
s_{\varphi,X}(x) = s_{\varphi^h,X}(x) = \sum_{x_i \in X} a_i^*(x) r_i.
\]

Using Lemma 2, we get that for \(x \in B_{sh}(0)\)

\[
\| \partial_v s_{\varphi,X}(x) - \partial_v \varphi_p(x) \|_\infty \leq \tilde{c}_1 \| \varphi_p \|_{C^k(B_{sh}(0))} h^{-k-1},
\]

and

\[
\| \partial_v s_{\varphi^h,X}(x) - \partial_v \varphi_p^h(x) \|_\infty = \| \partial_v s_{\varphi,X}(x) - \partial_v \varphi_p^h(x) \|_\infty \leq \tilde{c}_2 \| \varphi_p^h \|_{C^k(B_{sh}(0))} h^{-k-1},
\]

where

\[
\| f \|_{C^k(B_{sh}(0))} \triangleq \max_{1 \leq \ell \leq D} \max_{x \in B_{sh}(0)} | \partial^\ell f(x) |.
\]

We know from Theorem 4.11 of [21] that \(H(p)\) depend smoothly on \(p\) we get that both \(\varphi_p(x)\) and \(\varphi_p^h(x)\) vary smoothly with \(p\). Combining this with the fact that \(\mathcal{M}\) is compact we can give a global constant bound for both \(\| \varphi_p \|_{C^k(B_{sh}(0))}\) and \(\| \varphi_p^h \|_{C^k(B_{sh}(0))}\). Thus we obtain

\[
\| \partial_v s_{\varphi,X}(x) - \partial_v \varphi_p(x) \|_\infty \leq c_1 h^{-k-1}
\]

\[
\| \partial_v s_{\varphi,X}(x) - \partial_v \varphi_p^h(x) \|_\infty \leq c_2 h^{-k-1}
\]

Then, by the triangle inequality we achieve

\[
\| \partial_v \varphi_p^h(x) - \partial_v \varphi_p(x) \|_\infty \leq \| \partial_v \varphi_p^h(x) - \partial_v s_{\varphi,X}(x) \| + \| \partial_v s_{\varphi,X}(x) - \partial_v \varphi_p(x) \| \\
\leq c_1 h^{-k-1} + c_2 h^{-k-1} = C_1 h^{-k-1}.
\]

Finally, we note that the polynomial \(\pi^*(x \mid p)\) which minimizes (18) coincides with the polynomial \(\pi^*(x \mid 0)\) minimizing (4) with respect to the domain \(H(p)\). As described in
the alternative minimization problem of (6), it reproduces polynomials in \( \Pi_{k-1} \), and more specifically, the Taylor polynomial of degree \( k-1 \). Therefore, we get for \( x \in B_{sh}(0) \)
\[
\| \partial_v \pi^*(x \mid p) - \partial_v \varphi_p(x) \|_\infty \leq C_2 h^{k-1} \\
\| \partial_v \pi^*(x \mid p) - \partial_v \varphi^h_p(x) \|_\infty \leq C_3 h^{k-1}
\]
as required. \( \square \)

An immediate consequence of Lemma 4 is that the differential of \( \varphi^h_p \) approximates the differential of \( \varphi_p \)

**Lemma 5.** Let the conditions of Lemma 4 hold, and let \( \mathcal{D}_0 \varphi_p, \mathcal{D}_0 \varphi^h_p \) denote the differentials of \( \varphi_p, \varphi^h_p \) at 0 respectively. Then,
\[
\mathcal{D}_0 \varphi_p = \mathcal{D}_0 \varphi^h_p + X,
\]
where \( X \) is a linear operator with operator norm
\[
\| X \|_{op} \leq C h^{k-1},
\]
for some constant \( C \).

**Proof.** Let us define
\[
X = \mathcal{D}_0 \varphi_p - \mathcal{D}_0 \varphi^h_p,
\]
which is linear as a subtraction of two linear operators. Recall that vectors \( v \in T_0 H(p) = H(p) \) naturally correspond to directional derivatives; explicitly,
\[
\mathcal{D}_0 \varphi_p \cdot v = \partial_v \varphi_p(0) \\
\mathcal{D}_0 \varphi^h_p \cdot v = \partial_v \varphi^h_p(0).
\]
From Lemma 4 we know that for any such \( v \) we get
\[
\| X \cdot v \| = \| \mathcal{D}_0 \varphi_p \cdot v - \mathcal{D}_0 \varphi^h_p \cdot v \| = \| \partial_v \varphi_p(0) - \partial_v \varphi^h_p(0) \| \leq C_1 h^{k-1},
\]
since the Euclidean norm respects the inequality \( \| x \| \leq \| x \|_\infty \) (essentially, this claim will be true in any given norm due to the equivalence of norms in finite dimensions). Thus,
\[
\| X \|_{op} = \max_{\| v \| = 1} \| X \cdot v \| \leq C h^{k-1}
\]
\( \square \)

**Corollary 1.** Define the local map \( \mu_p : W_p \to W^h \) by \( \varphi^h_p \circ \varphi^{-1}_p \). Then, its differential
\[
\mathcal{D}_p \mu_p = \text{Id} + Y,
\]
where
\[
\| Y \|_{op} \leq C h^{k-1},
\]
for some constant \( C \).
Proof. Let us write $\mu_p$ as the following composition

$$\mu_p = \varphi_p^h \circ \varphi_p^{-1} : W_p \to W^h,$$

where $\varphi_p : U \to W_p \subset M$ and $\varphi_p^h : U \to W^h \subset M^h$ as defined in Lemma 3. Then, by the chain rule we get

$$D_p \mu_p = D_0 \varphi_p^h \circ D_p \varphi_p^{-1}.$$

Furthermore, from the fact that $Id = D_p(Id)$ we know

$$Id = D_p(\varphi_p \circ \varphi_p^{-1}) = D_0 \varphi_p \circ D_p \varphi_p^{-1},$$

and so we get

$$D_p \varphi_p^{-1} = (D_0 \varphi_p)^{-1}.$$

In addition, we know from Lemma 5 that

$$D_0 \varphi_p = D_0 \varphi_p^h + X,$$

where $\|X\|_{op} \leq c_1 h^{k-1}$. Applying $(D_0 \varphi_p)^{-1}$ to both sides of the equation results with

$$Id = D_0 \varphi_p^h \circ (D_0 \varphi_p)^{-1} + X \circ (D_0 \varphi_p)^{-1}$$

$$Id - D_p P^h_k = X \circ (D_0 \varphi_p)^{-1}.$$

Denoting $X \circ (D_0 \varphi_p)^{-1}$ by $Y$ we get the desired expression, and

$$\|Y\|_{op} = \|X \circ (D_0 \varphi_p)^{-1}\|_{op} \leq \|X\|_{op} \|(D_0 \varphi_p)^{-1}\|_{op} = \|X\|_{op} \frac{1}{\|D_0 \varphi_p\|_{op}}.$$

Since the manifold is differentiable and compact, and $H$ approximates the tangent, we can bound $\|D_0 \varphi_p\|_{op}$ from below by $0 < 1/c_2$ a constant independent of $p$. Therefore,

$$\|Y\|_{op} \leq c_1 c_2 h^{k-1} = Ch^{k-1},$$

as requested.

Lemma 6. Let the Sampling Assumptions of Section 2.2 and the Injectivity Conditions of Section 3.1 hold. Let $\theta_1, \theta_2$ of (15) and (17) be with supports of size $c_1 h$ and $sh$ respectively and let $s \leq c_1$. Fix an arbitrary point $p \in M$ and denote the corresponding coordinate system resulting from (16) by $(q, H) = (q(p), H(p))$, and let

$$X = \{x_i = P_H(r_i - q) \mid \|r_i - p\| \leq \mu\}_{i=1}^n \cap B_0(sh),$$

where $s$ is some constant and $\mu$ is smaller than the injectivity radius (i.e., the manifold’s reach). Then, for $h_{R,M}$ sufficiently small we get

$$h_{X,B_0(sh)} \leq c \cdot h_{R,M},$$

where $c$ is some constant independent of $p$. 

---

15
Proof. The proof outline is as follows

1. We first look at the sample set projected onto the tangent $T_p\mathcal{M}$,

   \[
   \tilde{X} = \{ \tilde{x}_i = P_{T_p\mathcal{M}}(r_i - p) \mid \|r_i - p\| \leq \mu \}_{i=1}^n \cap B_0(c_1h) \]

   and its fill-distance
   \[
   h_{\tilde{X}, B_0(c_1h)} \leq h_{R,M}.
   \]

2. Then by Lemma A.1 from Appendix A we achieve that there exists a constant $\rho$ such that the number of samples in $\tilde{X}$ is bounded by $c_1 \cdot \rho^d$.

3. As a result, we achieve that $J_1(p, T_p\mathcal{M}) = O(h^4)$.

4. Since $(q(p), H(p))$ are the minimizers of (16) and since $(p, T_p\mathcal{M})$ are in the search space of the minimization problem of (16), we get that

   \[
   J_1(q(p), H(p)) \leq J_1(p, T_p\mathcal{M}) = O(h^4)
   \]

5. Since the number of samples in $X$ is bounded by $c_1 \cdot \rho^d$ as well we achieve that

   \[
   d(r_i - q, H) = O(h^2)
   \]

6. Thus, for small enough $h_{R,M}$ we get that

   \[
   h_{X, B_0(c_1h)} \leq c \cdot h_{R,M}.
   \]

Looking at the projected sample

\[
\tilde{X} = \{ \tilde{x}_i = P_{T_p\mathcal{M}}(r_i - p) \mid \|r_i - p\| \leq \mu \}_{i=1}^n \cap B_0(c_1h),
\]

and measuring the fill-distance of that set we get

\[
h_{\tilde{X}, B_0(c_1h)} = \sup_{\tilde{x} \in B_0(c_1h)} \min_{\tilde{x} \in \tilde{X}} \|\tilde{x} - \tilde{x}_i\| = \sup_{r \in \phi[B_0(c_1h)]} \min_{r_i \in \phi[B_0(c_1h)]} \|P_{T_p\mathcal{M}}(r - p) - P_{T_p\mathcal{M}}(r_i - p)\|,
\]

where $r = \phi(\tilde{x})$ is the unique point which satisfies $P_{T_p\mathcal{M}}(r - p) = \tilde{x}$. Thus,

\[
h_{\tilde{X}, B_0(c_1h)} \leq \sup_{r \in \phi[B_0(c_1h)]} \min_{r_i \in \phi[B_0(c_1h)]} \|r - r_i\| \leq h_{R,M}.
\]

Furthermore, since the tangent is a linear approximation at $p$ we know that for $r_i \in \phi[B_0(c_1h)]$

\[
d(r_i - p, T_p\mathcal{M}) = O(h^2).
\]
Thus, for any $\varepsilon$ there is small enough $h$ such that

$$\phi[B_0(c_1h)] \subset B_p((c_1 + \varepsilon)h)).$$

From the fact that $R$ is a quasi-uniform set and Lemma A.1 we know that

$$\#(R \cap B_p((c_1 + \varepsilon)h)) \leq (c_1 + \varepsilon) \cdot \rho^d$$

for some constant $\rho$. Thus, we achieve that

$$\# \tilde{X} \leq (c_1 + \varepsilon) \cdot \rho^d$$

as well.

Now we show that $J_1(q(p), H(p) \mid p) = O(h^4)$ by showing that

$$J_1(q(p), H(p) \mid p) \leq J_1(p, T_pM \mid p) = O(h^4).$$

The coordinate system $(q(p), H(p))$ is the one minimizing the following energy

$$J_1(q, H \mid r) = \sum_{i=1}^n d(r_i - q, H)^2 \theta_1(||r_i - q||)$$

under the constraints of (16):

1. $r - q \perp H$
2. $q \in B_\mu(r)$
3. $\# (R \cap B_h(q)) \neq 0$

We note that a possible choice of coordinates that respect these three constraints is given by taking $q = p$ and $H = T_pM$. Thus, we achieve

$$J_1(q(p), H(p) \mid p) \leq J_1(p, T_pM \mid p)$$

, and

$$J_1(p, T_pM \mid p) = \sum_{i=1}^n d(r_i - p, T_pM)^2 \theta_1(||r_i - p||).$$

Since $T_pM$ is the linear approximation to $M$ at $p$, we know that for $r_i$ in the support of $\theta_1$ (i.e., $||r_i - p|| \leq c_1h$)

$$\|P_{T_pM}(r_i - p)\| \leq c_1hR,M,$$

and thus,

$$d(r_i - p, T_pM) \leq c_2h^2_{R,M}.$$

Furthermore, since $\# \tilde{X} \leq (c_1 + \varepsilon)\rho^d$ we achieve that

$$J_1(p, T_pM \mid p) = O(h^4)$$
and, as a result
\[ J_1(q(p), H(p) \mid p) = O(h^4) \]
as well. From Lemma 4.4 of [21] we know that under the Injectivity Conditions as \( h \to 0 \) we achieve that \( q(p) \) approaches \( p \), and if we denote \( \epsilon = \|q(p) - p\| \), then
\[ \|P_H(p) - P_{T_{\nu}M}\|_{op} \leq O(h + \epsilon^2). \]
Thus, for sufficiently small \( h \) we get that
\[ \#X \leq (s + \bar{\epsilon}) \cdot \rho^d \]
and we get that
\[ d(r_i - q(p), H(p)) = O(h^2). \]
Finally,
\[
\begin{align*}
h_{X,B_0(sh)} &= \sup_{x \in B_0(sh)} \min_{x_i \in X} \|x - x_i\| \\
&= \sup_{r \in \phi_p[B_0(sh)]} \min_{r_i \in \phi_p[B_0(sh)]} \|P_H(r - q) - P_H(r_i - q)\| \\
&\leq \sup_{r \in \phi_p[B_0(sh)]} \min_{r_i \in \phi_p[B_0(sh)]} \|r - r_i\| + O(h^2).
\end{align*}
\]
Thus, for sufficiently small \( h \) we obtain
\[ h_{X,B_0(sh)} \leq c \cdot h_{R,M}. \]

### 3.2 Geodesic Paths on \( M^h \) Converge to Geodesic Paths in \( M \)

Let us fix points \( p_1, p_2 \in M \) and denote a path between them by \( \gamma(t) : [a, b] \to M \) explicitly, \( \gamma(a) = p_1, \gamma(b) = p_2 \). Then, the length of the curve is defined as
\[ L_g(\gamma(t)) = \int_a^b \|\dot{\gamma}(t)\|_g dt, \]
and the geodesic distance can be defined by the piece-wise smooth curve \( \gamma(t) : [a, b] \to M \) of minimal length such that \( \gamma(a) = p_1, \gamma(b) = p_2 \). Namely,
\[ \rho_M(p_1, p_2) = \min_{\gamma(t)} L_g(\gamma(t)) = \min_{\gamma(t)} \int_a^b \|\dot{\gamma}(t)\|_g dt. \tag{26} \]
Notice, that one can always choose the arc-length parameter defined by
\[ s(t) = \int_a^t \|\dot{\gamma}(t)\|_g dt \]
and then get
\[ L_g(\gamma(t)) = L_g(\gamma(s(t))) = \int_a^b ds. \]
Thus, when looking for a geodesic curve (i.e., one that minimizes the length) we can limit the discussion to curves with arc-length parameter; that is \( \|\dot{\gamma}(t)\|_g = 1 \), up to the zero measurable set where \( \gamma \) is not smooth.
A (Simplified) Local Case: Let $\gamma^*(t) : [a, b] \to \mathcal{M}$ denote a geodesic curve such that $\gamma^*(a) = p_1, \gamma^*(b) = p_2$ and $\|\dot{\gamma}^*(t)\|_g = 1$. That is, 

$$\rho_{\mathcal{M}}(p_1, p_2) = L_g(\gamma^*(t)).$$

If $p_1, p_2$ belong to an open neighborhood $W_{p_1} \subset \mathcal{M}$ such that 

$$\varphi_{p_1} : U \to W_{p_1}$$

and $U \subset B_{sh}(0) \subset H(p_1)$, then we can look at the map $\mu_{p_1} : W_{p_1} \to W_{p_1}^h$ defined by 

$$\mu = \varphi_{p_1}^b \circ \varphi_{p_1}^{-1}$$

and from Corollary 1 we know that its differential 

$$D_{p_1} \mu_{p_1} = Id + Y,$$

where $\|Y\|_{op} \leq ch^{k-1}$. 

In this case, we could look at the pushed forward curve 

$$\gamma^h(t) = \mu_{p_1}(\gamma^*(t)).$$

By the chain rule, we get 

$$\dot{\gamma}^h(t) = \partial_{t} \mid_{t=t} (\mu_{p_1} \circ \gamma^*(t)) = D_{p_1} \mu_{p_1} \cdot \partial_t \gamma^*(t) = D_{p_1} \mu_{p_1} \cdot \dot{\gamma}^*(t).$$

Using Corollary 1 we achieve that 

$$\dot{\gamma}^h(t) = \dot{\gamma}^*(t) + Y \cdot \dot{\gamma}^*(t),$$

and thus, 

$$\|\dot{\gamma}^h(t)\|_{g^h} = \|\dot{\gamma}^*(t) + Y \cdot \dot{\gamma}^*(t)\|_g$$

$$\leq \|\dot{\gamma}^*(t)\|_g + \|Y \cdot \dot{\gamma}^*(t)\|_g$$

$$\leq \|\dot{\gamma}^*(t)\|_g + \|Y\|_{op} \|\dot{\gamma}^*(t)\|_g$$

$$= 1 + \|Y\|_{op} \|\dot{\gamma}^*(t)\|_g$$

$$\leq 1 + ch^{k-1} \tag{27}$$

Since the manifold $\mathcal{M}$ is closed (i.e., $\rho_{\mathcal{M}}(p_1, p_2)$ can be bounded globally), the length of $\gamma^h(t)$ can now be bounded by 

$$L_{g^h}(\gamma^h(t)) \leq \int_a^b 1 + ch^{k-1} dt = L_g(\gamma^*) + ch^{k-1} L_g(\gamma^*). \tag{28}$$

Assuming the lengths of curves are bounded we get 

$$L_{g^h}(\gamma^h(t)) \leq \rho_{\mathcal{M}}(p_1, p_2) + c_1 h^{k-1}.$$
From the fact that the geodesic curve has minimal length we get
\[ \rho_{\mathcal{M}^h}(\tilde{p}_1^h, \tilde{p}_2^h) \leq \rho_{\mathcal{M}}(p_1, p_2) + c_1 h^{k-1}, \]
where \( \tilde{p}_1^h = \mu_{p_1}(p_1) \) and \( \tilde{p}_2^h = \mu_{p_2}(p_2) \). Unfortunately, this analysis does not solve the issue. Explicitly, even though
\[ \mu_{p_1}(p_1) = \tilde{p}_1^h = p_1^h = \mathcal{P}_k^h(p_1), \]
it is apparent that
\[ \mu_{p_1}(p_2) = \tilde{p}_2^h \neq p_2^h = \mathcal{P}_k^h(p_2). \]
Therefore, we need to bridge this gap between \( p_1^h \) and \( p_2^h \) and make sure that it is \( O(h^{k-1}) \) as well. By doing this, we will achieve the desired bound from the triangle inequality.

**The (Real) Global Case:** Let us now remove the assumption that \( p_1, p_2 \) belong to such a neighborhood \( W \) that can be mapped by \( \varphi_{p_1} \). In this case we sample the curve \( \gamma^*(t) \) at the points \( \{z_j = \gamma^*(t_j)\}_{j=1}^{J+1} \) densely enough such that each pair \( z_j, z_{j+1} \) belong to an open neighborhood \( W_j \subset B_{\text{sh}}(z_j) \) that can be mapped by
\[ \varphi_{z_j} : U_j \rightarrow W_j. \]
Then, we look at the set of points
\[
\begin{align*}
    z_j^h &= \mathcal{P}_k^h(z_j) \in \mathcal{M}^h, \quad (j = 1, \ldots, J+1) \\
    z_{j+1}^h &= \mu_{z_j}(z_{j+1}) \in \mathcal{M}^h, \quad (j = 1, \ldots, J).
\end{align*}
\]
Note, that by definition
\[ \mu_{z_j}(z_j) = z_j^h. \]
Thus, essentially the set of points we are looking at are
\[
\begin{align*}
    z_j^h &= \mu_{z_j}(z_j) \in \mathcal{M}^h, \quad (j = 1, \ldots, J+1) \\
    z_{j+1}^h &= \mu_{z_j}(z_{j+1}) \in \mathcal{M}^h, \quad (j = 1, \ldots, J).
\end{align*}
\]
Using these sets of points we construct a piece-wise smooth curve \( \gamma^{h,J}(t) \). We begin by defining the first segment \( \gamma_1^h(t) \) connecting \( z_1^h \) with \( z_2^h \) by pushing the curve \( \gamma_1^* \) (the restriction of \( \gamma^*(t) \) between \( z_1 \) and \( z_2 \)) with \( \mu_{z_1} \). Explicitly,
\[ \gamma_1^h(t) = \mu_{z_1} \circ \gamma_1^*(t) = \mu_{z_1} \circ \gamma^*(t)|_{[z_1,z_2]} \mathcal{M}. \]
Note, that \( \gamma_1^h(t) \) ends at \( \mu_{z_1}(z_2) = z_2^h \). Then, the following segment denoted by \( \tilde{\gamma}_1^h(t) \) is defined through the geodesic path connecting \( z_2^h \) and \( z_2^h = \mathcal{P}_k^h(z_2) = \mu_{z_2}(z_2) \). Subsequently,
\[ \gamma_2^h(t) = \mu_{z_1} \circ \gamma_2^*(t) = \mu_{z_2} \circ \gamma^*(t)|_{[z_2,z_3]} \mathcal{M}, \]
and \( \tilde{\gamma}_2^h(t) \) is defined through the geodesic path connecting \( z_3^h \) and \( z_3^h = \mathcal{P}_k^h(z_3) = \mu_{z_3}(z_3) \). We define the rest of the piece-wise smooth curve in the same manner (see Figure 4). Let us denote
\[ \ell_j^h = L_{\gamma_j^h}(\gamma_j^h(t)), \quad (j = 1, \ldots, J) \]
\[ \ell_j^* = L_g(\gamma_j^*(t)), \quad (j = 1, \ldots J) \]

Then, from the considerations portrayed in the Localized Case (i.e., (27)) we get that

\[ \ell_j^h \leq \ell_j^*. \]

Furthermore, from Lemma 7 we achieve that

\[ L_g^h(\gamma_j^h(t)) \leq c_1 h^k. \]

Therefore,

\[
L_g^h(\gamma^{h,J}(t)) = \sum_{j=1}^{J} L_g^h(\gamma_j^h(t)) + L_g^h(\tilde{\gamma}_j^h(t)) \\
\leq \sum_{j=1}^{J} \ell_j^h + c_1 h^k \\
\leq \sum_{j=1}^{J} \ell_j^* + c_2 h^{k-1} \ell_j^* + c_1 h^k \\
= L_g(\gamma^*(t)) + L_g(\gamma^*(t)) \cdot c_2 h^{k-1} + Jc_1 h^k,
\]

where the second inequality comes from (28). It is clear that for sufficiently small \( h \) we have \( J = O(L_g(\gamma^*(t))/h) \), and so we obtain

\[ L_g^h(\gamma^{h,J}(t)) \leq \rho_M(z_1, z_{J+1}) + \rho_M(z_1, z_{J+1}) \cdot c_1 h^{k-1} + c_3 h^{k-1}. \]

Since \( \mathcal{M} \) is closed (i.e., compact and bounded) and of bounded reach we know that the length of curves on \( \mathcal{M} \) are bounded, and we obtain

\[ \rho_M^h(p_1^h, p_2^h) \leq \rho_M(p_1, p_2) + C_1 h^{k-1}, \]

where \( p_1 = z_1, p_2 = z_{J+1} \) and \( p_1^h = z_1^h, p_2^h = z_{J+1}^h \).

By symmetry, we can show

\[ \rho_M(p_1, p_2) \leq \rho_M(p_1^h, p_2^h) + C_2 h^{k-1}, \]

and the proof of Theorem 1 is concluded.

\[ \square \]

**Lemma 7.** Let \( \rho_M(z_j, z_{j+1}) \leq sh \) then for sufficiently small \( h \)

\[ \rho_M^h(z_{j+1}, z_{j+1}^h) \leq Ch^k. \]

**Proof.** For simplicity we assume that \( q(z_j) = 0 \). We first remind that

\[ z_{j+1}^h = P_k^h(z_{j+1}), \]
and thus by (24)

$$\|z_{j+1}^h - z_{j+1}\| \leq c_1 h^k.$$ 

Accordingly, we have

$$\|P_{H(z_j)}(z_{j+1}) - P_{H(z_j)}(z_{j+1}^h)\| \leq c_2 h^k.$$ 

On the other hand we have

$$P_{H(z_j)}(z_{j+1}) = P_{H(z_j)}(z_{j+1}^h),$$

and so

$$\|P_{H(z_j)}(z_{j+1}^h) - P_{H(z_j)}(z_{j+1}^h)\| \leq c_2 h^k$$

as well. Now, since $\varphi_{z_j}^h$ so is $\phi_{z_j}^h : H(z_j) \to H^1_{\mathcal{M}}$ (as defined in (22); see Figure 3) is a smooth function, and through taking the zero\textsuperscript{th} order Taylor approximation around $x_0 = P_{H(z_j)}(z_{j+1})$ we know that

$$\phi_{z_j}(x_0 + \varepsilon) = \phi_{z_j}^h(x_0) + c_3 \varepsilon,$$

where $c_3$ depends on the first derivatives of $\varphi_{z_j}^h$. Since $\varphi_{z_j}^h$ vary smoothly with the change in $z_j$ and $\mathcal{M}$ is compact, we can bound $c_3$ by a uniform constant independent of $z_j$. Therefore, by taking $\varepsilon = P_{H(z_j)}(z_{j+1}) - P_{H(z_j)}(z_{j+1}^h)$ we get

$$\|\varepsilon\| \leq c_2 h^k,$$

and so, since

$$\|\tilde{z}_{j+1}^h - z_{j+1}^h\| = \left\|\varphi_{z_j}^h(P_{H(z_j)}(\tilde{z}_{j+1}^h)) - \varphi_{z_j}^h(P_{H(z_j)}(z_{j+1}^h))\right\|,$$

we achieve

$$\|\tilde{z}_{j+1}^h - z_{j+1}^h\| = \sqrt{\varepsilon^2 + \left(\phi_{z_j}^h(P_{H(z_j)}(\tilde{z}_{j+1}^h)) - \phi_{z_j}^h(P_{H(z_j)}(z_{j+1}^h))\right)^2} \leq c_2 h^k + c_3 h^k = c_4 h^k.$$ 

Finally, for sufficiently small $h$

$$\|\tilde{z}_{j+1}^h - z_{j+1}^h\| \approx \rho_{\mathcal{M}}(\tilde{z}_{j+1}^h, z_{j+1}^h),$$

Figure 4: Illustration of $\gamma_{h,J}(t)$ (the piece-wise smooth curve in blue) drawn near the limit curve $\gamma_h(t)$ (in red)
and we get
\[ \rho_{M^h}(z^h_{j+1}, z^h_{j+1}) \leq Ch^k, \]
as required.

## 4 Numerical Simulations

We now propose a simple algorithm to approximate geodesic distances based upon the aforementioned theoretical analysis. The basic idea is to sample \( M^h \) with many more samples than in the original point cloud to arrive at accurate approximations of geodesic distances on \( M^h \).

### 4.1 Naive Geodesic Approximation

Let \( R = \{r_i\}_{i=1}^n \) be a sample-set of \( M \) a smooth \( d \)-dimensional submanifold of \( \mathbb{R}^D \). Then, we create \( X = \{x_i\}_{i=1}^N (N = K^d \cdot n) \), a new set of samples of \( M^h \) by means of Algorithm 1. Explicitly, for each point in the original set \( r_i \) find \((q(r_i), H(r_i))\) of (16) (for an implementation of this see [20, 21]). We then create a uniform grid \( \{\tilde{x}_{ij}\}_{j=1}^{K^d} \) on \( H \) around the origin. In other words, in each basis direction on \( H \) take \( K \) uniform samples from \(-\sigma, \sigma\). Then, find \( x_{ij} \) for \( j = 1, \ldots, K^d \) by computing the \( k^{th} \) degree Manifold-MLS projection of \( \tilde{x}_{ij} \) onto \( M^h \); i.e.,

\[ x_{ij} = P_k(\tilde{x}_{ij}). \]

**Algorithm 1** Re-sampling of \( M^h \)

1: **Input:**
   \( R = \{r_i\}_{i=1}^n \subset M \subset \mathbb{R}^D \)
   \( d \)
   \( K \)
   \( k \)
   \( \sigma \)

2: **Output:**
   \( X = \{x_j\}_{j=1}^{K^d \cdot n} \subset M^h \subset \mathbb{R}^D \) - a new sample set of \( M^h \)

3: for \( r_i \in R \) do

4: Find \((q(r_i), H(r_i))\) of (16) \( \triangleright \) see [20, 21] for implementation details

5: Create \( \{\tilde{x}_{ij}\}_{j=1}^{K^d} \) a uniform grid on \( H \) \( \triangleright \) sample between \([-\sigma, \sigma]\) in each direction.

6: For \( j = i \cdot K + j \) project \( x_j = P_k(\tilde{x}_{ij}) \) \( \triangleright \) see [20, 21] for implementation details

7: end for

Finally, after obtaining the new sample set \( X \), given \( p_1, p_2 \in M \) we compute

\[ p^h_1 = P_k(p_1), \quad p^h_2 = P_k(p_2), \]

and compute the shortest path between \( p^h_1 \) and \( p^h_2 \) with respect to the new cloud of points \( X \).

In the experiments below we have pre-computed \( \sigma \) by the heuristic approach presented in Algorithm 2.
4.2 Qualitative Examples

In the first qualitative experiment, we have drawn uniformly $20^2$ samples of a 2-sphere embedded in $\mathbb{R}^3$ (Fig. 6B). Using Algorithms 1 and 2, with the parameters $k = 3, K = 5, d = 2$ we have created a sample set $X \subset \mathcal{M}^h$ of cardinality $20^2 \times 5^2$. Fig. 6 shows a comparison between the shortest path computed on $\mathbb{R}$ compared with the shortest path computed on $X$ (both paths were computed by Dijkstra algorithm). It is apparent that the initial sample set $\mathbb{R}$ has insufficient resolution, and thus provides very poor estimation of the geodesic path, whereas the generated sample $X$ yields a very accurate representation of the geodesic using the same data.

On the second qualitative experiment, we took a point cloud obtained from the digitized talus bone of an *Aotus trivirgatus*, a nocturnal monkey native to South America. The data was obtained from Morphosource\(^1\) and subsequently processed in Meshlab [3] (Fig. 7C-D). In this experiment we first sub-sample the manifold, then compute the shortest path (Fig. 7A). Then we up-sample the manifold using the Manifold-MLS and compute the shortest path again (Fig. 7B).

\(^1\)https://www.morphosource.org/Detail/SpecimenDetail/Show/specimen_id/549
Algorithm 2 Compute $\sigma$

\begin{itemize}
    \item $R = \{r_i\}_{i=1}^{n} \subset \mathcal{M} \subset \mathbb{R}^D$ the original sample set
    \item $d$ the intrinsic dimension of the manifold
    \item $k$ local polynomial degree
\end{itemize}

1: \textbf{Input:} $d$ \hspace{1cm} $k$

2: \textbf{Output:} $\sigma$

3: $\sigma_{\text{est}} = \text{zeros}(1,100)$

4: \textbf{for} $j = 1, \ldots, 100$ \textbf{do}

5: \hspace{1cm} Pick $r_{i'} \in R$ at random

6: \hspace{1cm} Find $\sigma_{i'}$ such that $\#(B(r_{i'}, \sigma_{i'}) \cap R) = \binom{k+d}{k}$

7: \hspace{1cm} $\sigma_{\text{est}}(j) = \sigma_{i'}$

8: \textbf{end for}

9: $\sigma \leftarrow \max(\sigma_{\text{est}})$

Figure 6: Shortest path from discrete samples of a 2-sphere. (A) The original manifold $\mathcal{M}$; the shortest path computed between $p_1, p_2$ (marked in red triangles) on discrete samples of the manifold (the path is marked by the red line); shortest path on re-sampling of the Manifold-MLS approximation to $\mathcal{M}$ (the path is marked by the green line). (B) The original samples of $\mathcal{M}$ and the shortest path connecting points $p_1, p_2$ marked by the red triangles. (C) The samples of the approximating manifold and the shortest path connecting points $p_1, p_2$ marked by the red triangles.
Figure 7: Shortest path from discrete samples of a talus bone. (A) A down-sampled Point Cloud of a talus bone and the shortest path computed between $p_1, p_2$ (marked in red triangles) on the sub-sampled manifold (the path is marked by the red line) (B) The samples of the $\mathcal{M}^h$ and the shortest path connecting points $p_1, p_2$ marked by the red triangles. (C-D) Two angles of both paths overlaying the original mesh of the talus bone. The path computed on the sparse sample set (red) and the path computed on $\mathcal{M}^h$ (green).
Table 1: Accuracy of geodesic distance estimation. The experiment was conducted on randomly chosen $d$-dimensional spheres of radius 0.5 embedded in $[0, 1]^{20} \subset \mathbb{R}^{20}$, for which $n$ uniformly distributed samples were taken. Column $R$ represents the error of applying the Dijkstra algorithm on the original set $R$. Column $X_1, X_3$ show the results of applying the Dijkstra algorithm on newly sampled set using $K = 3$ and degrees $k = 1, 3$ of Manifold-MLS respectively. Column Euclidean, represents the error of computing simple Euclidean distance between the points in $\mathbb{R}^{20}$. The root mean squared error was computed over 100 realizations.

| $d$ | $n$ (samples) | $R$ (RMSE%) | $X_1$ (RMSE%) | $X_3$ (RMSE%) | Euclidean (RMSE%) |
|---|---|---|---|---|---|
| 2 | 100 | 17.3% | 8.7% | 0.5% | 24.3% |
| 3 | 1000 | 36.3% | 1.1% | 3.2% | 18% |
| 4 | 10000 | 49.1% | 7.4% | 8.0% | 16% |

4.3 A Quantitative Experiment

In this experiment we aimed at performing a large scale comparison between the accuracy of estimating geodesic distance based upon shortest paths on the original samples of $\mathcal{M}$ with that of the approximating manifold. Although in both cases we compute the geodesic estimation by means of the Dijkstra algorithm, using the generated sample $X$ is assumed to be a close estimation of the geodesic distance of $\mathcal{M}^h$, since we can resample $\mathcal{M}^h$ using Algorithm 1 as fine as we wish. Nonetheless, as can be seen in Table 1 in order to provide good estimation to the geodesic distances the enlarging factor $K$ need not be to large (in all the experiments below we have used $K = 3$).

The experiment was conducted on random $d$-spheres of radius 0.5 embedded in the cube $[0, 1]^{20} \subset \mathbb{R}^{20}$, with varying intrinsic dimension and $\#R = 10^d$. The error was computed by the relative root mean squared error (RMSE%)$$
\text{RMSE\%} = \frac{1}{\text{AVG}} \sqrt{\frac{1}{n} \sum_{i=1}^{n} \|x_i - \hat{x}_i\|^2},
$$where AVG is the average of the accurate values. The RMSE\% was computed over 100 realizations. Explicitly, we have generated a random sample and then measured 100 distances between pairs of samples. We compare the results of performing Dijkstra on the original sample $R$, a generated sample $X_1$ using 1-degree Manifold-MLS, and a generated sample $X_3$ using 3-degree Manifold-MLS.

4.4 Robustness to Noise

Here we took 2-dimensional random sphere of radios 0.5 embedded in the cube $[0, 1]^{20} \subset \mathbb{R}^{20}$ and drew 400 samples uniformly distributed on the sphere with additive Gaussian noise $\epsilon_i \in \mathbb{R}^{20}$ and $\epsilon_i \sim \mathcal{N}(0, \sigma)$ for $\sigma = 10^{-5}, 10^{-3}, 10^{-2}$. Explicitly, the sample set was $R = \{r_i\}_{i=1}^{400}$, and $r_i = p_i + \epsilon_i$, where $p_i \in \mathcal{M}$ and $\epsilon_i$ is as described above.

We then measured the geodesic distances between 100 pairs of randomly chosen points from the sample set. The geodesic distances were measured through applying Dijkstra on the
| Noise Level | \( R \) (RMSE\%) | \( X_1 \) (RMSE\%) | \( X_3 \) (RMSE\%) | Euclidean (RMSE\%) |
|------------|------------------|------------------|------------------|------------------|
| \( 10^{-5} \) | 17.4\% | 2.93\% | 0.36\% | 15.4\% |
| \( 10^{-3} \) | 16.0\% | 2.93\% | 0.41\% | 15.4\% |
| \( 10^{-2} \) | 30.0\% | 3.20\% | 1.83\% | 15.4\% |

Table 2: Accuracy of geodesic distance estimation. The experiment was conducted on randomly chosen 2-dimensional spheres of radius 0.5 embedded in \([0, 1]^{20} \subset \mathbb{R}^{20}\), for which 400 uniformly distributed samples were taken. The first column represents the standard deviation of the additive Gaussian noise \( \epsilon_i \). Column \( R \) represents the error of applying the Dijkstra algorithm on the original set \( R \). Column \( X_1, X_3 \) show the results of applying the Dijkstra algorithm on newly sampled set using \( K = 5 \) and degrees \( k = 1, 3 \) of Manifold-MLS respectively. Column Euclidean, represents the error of computing simple Euclidean distance between the points in \( \mathbb{R}^{20} \). The root mean squared error was computed over 100 realizations.

original sample \( R \), a generated sample \( X_1 \) using a 1-degree Manifold-MLS, and a generated sample \( X_3 \) using a 3-degree Manifold-MLS. The generated samples were created by means of Algorithm 1 using \( K = 5 \). The results of this experiment are summarized in Table 2 and show that our algorithm is robust to high levels of noise.

**Acknowledgments**

**References**

[1] M. G. Armentano. Error estimates in sobolev spaces for moving least square approximations. *SIAM Journal on Numerical Analysis*, 39(1):38–51, 2001.

[2] M. Belkin and P. Niyogi. Laplacian eigenmaps for dimensionality reduction and data representation. *Neural computation*, 15(6):1373–1396, 2003.

[3] P. Cignoni, M. Callieri, M. Corsini, M. Dellepiane, F. Ganovelli, and G. Ranzuglia. Meshlab: an open-source mesh processing tool. In *Eurographics Italian chapter conference*, volume 2008, pages 129–136. Salerno, 2008.

[4] R. R. Coifman and S. Lafon. Diffusion maps. *Applied and computational harmonic analysis*, 21(1):5–30, 2006.

[5] E. W. Dijkstra et al. A note on two problems in connexion with graphs. *Numerische mathematik*, 1(1):269–271, 1959.

[6] H. Federer. Curvature measures. *Transactions of the American Mathematical Society*, 93(3):418–491, 1959.

[7] R. Kimmel and J. A. Sethian. Computing geodesic paths on manifolds. *Proceedings of the national academy of Sciences*, 95(15):8431–8435, 1998.
[8] D. Levin. The approximation power of moving least-squares. *Mathematics of computation*, 67(224):1517–1531, 1998.

[9] M. Lichtenstein, G. Pai, and R. Kimmel. Deep eikonal solvers. In *International Conference on Scale Space and Variational Methods in Computer Vision*, pages 38–50. Springer, 2019.

[10] T. Lin and H. Zha. Riemannian manifold learning. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 30(5):796–809, 2008.

[11] F. Mémoli and G. Sapiro. Distance functions and geodesics on submanifolds of \( \mathbb{R}^d \) and point clouds. *SIAM Journal on Applied Mathematics*, 65(4):1227–1260, 2005.

[12] D. Mirzaei. Analysis of moving least squares approximation revisited. *Journal of Computational and Applied Mathematics*, 282:237–250, 2015.

[13] J. S. Mitchell, D. M. Mount, and C. H. Papadimitriou. The discrete geodesic problem. *SIAM Journal on Computing*, 16(4):647–668, 1987.

[14] A. Moscovich, A. Jaffe, and B. Nadler. Minimax-optimal semi-supervised regression on unknown manifolds. *arXiv preprint arXiv:1611.02221*, 2016.

[15] R. Ravier. *Algorithms with Applications to Anthropology*. PhD thesis, Duke University, 2018.

[16] R. J. Ravier. Eyes on the prize: Improved registration via forward propagation. *arXiv preprint arXiv:1812.10592*, 2018.

[17] L. K. Saul and S. T. Roweis. An introduction to locally linear embedding. *unpublished. Available at: http://www.cs.toronto.edu/~roweis/lle/publications.html*, 2000.

[18] L. K. Saul and S. T. Roweis. Think globally, fit locally: unsupervised learning of low dimensional manifolds. *Journal of machine learning research*, 4(Jun):119–155, 2003.

[19] J. A. Sethian. A fast marching level set method for monotonically advancing fronts. *Proceedings of the National Academy of Sciences*, 93(4):1591–1595, 1996.

[20] B. Sober, Y. Aizenbud, and D. Levin. Approximation of functions over manifolds: A moving least-squares approach. *arXiv preprint arXiv:1711.00765*, 2017.

[21] B. Sober and D. Levin. Manifold approximation by moving least-squares projection (mmls). *Constructive Approximation*, pages 1–46, 2019.

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

[23] H. Wendland. Local polynomial reproduction and moving least squares approximation. *IMA Journal of Numerical Analysis*, 21(1):285–300, 2001.

[24] H. Wendland. *Scattered data approximation*, volume 17. Cambridge university press, 2004.
A $h$-$\rho$-$\delta$ Sets, Quasi Uniform Sampling and Interior Cone Condition

The purpose of this appendix is to bridge between two different sets of assumptions that are used in the convergence analysis of MLS-based methods. We wrote this appendix in a standalone manner, so that scholars interested just in this explanation will not need to read the whole paper. Therefore, we repeat some of the definitions explained in Section 2. As mentioned above, Levin introduced in 1998 the notion of $h$-$\rho$-$\delta$ sets (see Definition A.1 below) to show optimal rates of convergence of the MLS for function approximation [8]. In his paper, Levin handles the case of unbounded and boundaryless flat domains. Then, in order to utilize Levin’s bounds, this definition reappeared in [20, 21] where similar bounds are derived for the Manifold-MLS as well as MLS approximation of functions defined over manifold domains. These works lay the foundations for the current paper.

However, in 2001, Wendland takes a seemingly different approach to achieve similar error analysis [23, 24] which reappears in [12], whose results enables a crucial step in the analysis presented in the current paper. Namely, Wendland introduces the notion of quasi-uniform sample sets and takes a more general approach to show convergence rates for any polynomial reproduction formula. Furthermore, in order to deal with boundary problems, Wendland requires that the domain maintain the Interior Cone Condition, which is the case with unbounded domains.

In the following passages we show that Levin has some redundancy in his $h$-$\rho$-$\delta$ set definition, and that, essentially, any quasi-uniform set is $h$-$\rho$-$\delta$ set. Thus, the results obtained in [21] and [20] could have been obtained using the assumption of a quasi-uniform set.

A.1 $h$-$\rho$-$\delta$ Sets and Convergence Rates

Let $X = \{x_i\}_{i=1}^{n}$ be a sample set with pairwise distinct points in $\Omega \subset \mathbb{R}^d$, and let $f(x_i)$ be the function evaluation at this point. Then, by choosing $u_j(x) : \mathbb{R}^d \rightarrow \mathbb{R}$ we can construct an approximation of the form

$$s(x) = \sum_{j=1}^{n} u_j(x) f(x_j).$$

Such an approximation is sometimes called quasi-interpolant.

The MLS approximation is defined through the following minimization. Let $\pi^*(x \mid \xi)$ be defined by

$$\pi^*(x \mid \xi) = \arg\min_{\pi \in \Pi_{k-1}} \sum_{i=1}^{n} \|f(x_i) - \pi(x_i)\|^2 \theta_h(\|\xi - x_i\|),$$

(29)

where $\theta_h$ is compactly supported on $[0, qh]$ and is consistent across scales (i.e., $\theta_h(th) = \Phi(t)$), and, $\| \cdot \|$ denotes the Euclidean norm. Then we define the MLS function as

$$s_{\text{MLS}}^{f, X}(x) \triangleq \pi^*(0 \mid x).$$

(30)

Levin showed in [8] that, given enough samples in $I(x)$, the neighborhood of $x$, the MLS
approximation defined in (30) can be pronounced as
\[ \sum_{i \in I} a_i^*(x) f(x_i), \]
where the coefficients \( a_i^*(x) \) are determined by minimizing the quadratic form
\[ \sum_{i \in I} a_i(x)^2 \frac{1}{\theta_h(x - x_i)} \]
under the constraints
\[ \sum_{i \in I} a_i(x) p(x_i) = p(x), \quad \forall p \in \Pi_{k-1}(\mathbb{R}^d). \]

In other words, the MLS approximant can be written as the following quasi-interpolant
\[ s_{f,X}^{\text{MLS}}(x) = \sum_{i \in I} a_i^*(x) f(x_i) = \pi^*(0 \mid x). \quad (31) \]

In order to show the rates of convergence for the MLS, Levin uses the following notion of an \( h\)-\( \rho\)-\( \delta \) set.

**Definition A.1** (\( h\)-\( \rho\)-\( \delta \) sets of fill distance \( h \), density \( \leq \rho \), and separation \( \geq \delta \)). Let \( \Omega \) be a \( d \)-dimensional domain in \( \mathbb{R}^n \), and consider sets of data points in \( \Omega \). We say that the set \( X = \{x_i\}_{i=1}^n \) is an \( h\)-\( \rho\)-\( \delta \) set if:

1. \( h \) is the fill distance with respect to the domain \( \Omega \)
   \[ h_{X,\Omega} = \sup_{x \in \Omega} \min_{x_i \in X} \|x - x_i\|. \]
   Throughout the appendix we will denote for short \( h := h_{X,\Omega} \).

2. \[ \# \{X \cap \overline{B}_q(y)\} \leq \rho \cdot q^d, \quad q \geq 1, \quad y \in \mathbb{R}^n. \]
   Here \( \#Y \) denotes the number of elements in a given set \( Y \), while \( \overline{B}_r(x) \) is the closed ball of radius \( r \) around \( x \).

3. \( \exists \delta > 0 \) such that
   \[ \|x_i - x_j\| \geq h \delta, \quad 1 \leq i < j \leq n. \]

**Theorem A.1** (Theorem 5 from [8]). Let \( f \in C^k(\Omega) \), and let \( \Omega \) be a compact domain. Then for fixed \( \rho \) and \( \delta \), there exists a fixed \( q > 0 \), independent of \( h \), such that \( s_{f,X}^{\text{MLS}} \) of (31) with compactly supported weight function \( \theta \) with support of size \( qh \), satisfies
\[ \|s_{f,X}^{\text{MLS}} - f(x)\|_{\Omega,\infty} \leq C_f \cdot h^k, \]
for \( h \) sufficiently small, for \( h\)-\( \rho\)-\( \delta \) sets of data points, and the constant \( C_f \) depends only on \( f \).
The proof of this theorem relies on the following observations. Let \( p \in \Pi_{k-1} \), then
\[
|s^\text{MLS}_{f,X}(x) - f(x)| \leq |s^\text{MLS}_{f,X} - p(x)| + |p(x) - f(x)|.
\]
Thus,
\[
|s^\text{MLS}_{f,X}(x) - f(x)| \leq \sum_{i \in I(x)} |a^*_i(x)(f(x_i) - p(x_i))| + \|p(x) - f(x)\|_{\Omega,\infty}
\]
\[
\leq \sum_{i \in I(x)} |a^*_i(x)| |f(x_i) - p(x_i)| + \|p(x) - f(x)\|_{\Omega,\infty}
\]
\[
\leq \left(1 + \sum_{i \in I(x)} |a^*_i(x)|\right) \|p(x) - f(x)\|_{\Omega,\infty}.
\]
Since we can choose \( p(x) \) to be any polynomial in \( \Pi_{k-1} \) we can choose it to be the Taylor expansion of \( f \) at \( x \), so
\[
|s^\text{MLS}_{f,X}(x) - f(x)| \leq \left(1 + \sum_{i \in I(x)} |a^*_i(x)|\right) \|f\|_{C^k(\Omega)},
\] (32)
where \( |f|_{C^k(\Omega)} \) is the semi-norm defined by
\[
|f|_{C^k(\Omega)} \overset{\Delta}{=} \max_{|\alpha|=k} \|\partial^\alpha f\|_{\Omega,\infty}.
\]
As a result, the key argument in providing the desired bound is through bounding the expression
\[
\sum_{i \in I(x)} |a^*_i(x)|.
\] (33)
Note, that due to the density parameter \( \rho \) of Definition A.1 the number of elements in \( I(x) \) the \( qh \)-size neighborhood of \( x \) is bounded by \( N = \rho \cdot q^d \), which is independent of \( h \). In addition, Levin shows that \( a^*_i(x) \) are continuous with respect to the installation of their neighboring samples \( x_j \). Thus, by denoting the neighbors of \( x \) by \( \{x_{i1}, \ldots, x_{iN}\} \), we can write
\[
\eta(x_{i1}, \ldots, x_{iN}) = |a^*_i(x)|,
\]
and
\[
C_{x,h} = \arg\max_{(x_{i1}, \ldots, x_{iN}) \in B_{qh}(x)} \eta(x_{i1}, \ldots, x_{iN}).
\]
Note, that the maximum exists since the search space is compact. The fact that the search space is bounded comes directly from the fact that \( (x_{i1}, \ldots, x_{iN}) \in B_{qh}(x) \). To show that it is closed, let \( \xi = (x_{i1}, \ldots, x_{iN}) \) be a limit of a sequence \( \xi_k = (x^k_{i1}, \ldots, x^k_{iN}) \) with separation parameter \( \delta \), then \( \xi \) will still have a separation parameter \( \delta \) and will still be contained in \( B_{qh}(x_i) \).
Thus, for a fixed $x$ and fixed $h$ it follows that
\[\sum_{i \in I(x)} |a_i^*(x)| \leq C_{x,h},\]
where $C_{x,h}$ is the same for all $h$-$\rho$-$\delta$ sets of points. Furthermore, since each $a_i^*(x)$ depends only on the installation if its $qh$-size neighboring points, it is clear that $C_{x,h} = C_h$ is independent of $x$. The proof is finalized by showing that if the weighting is consistent across scales (i.e., $\theta_h(th) = \Phi(t)$), then $a_i^*(x)$ are independent of $h$. Therefore, $C_h = C$ is independent of $h$ as well and
\[\|s_{f,X} - f(x)\|_{\Omega,\infty} \leq C \|f\|_{C^k(\Omega)} \cdot h^k = C_f \cdot h^k.\]

Interestingly, Wendland shows that the definition of density is redundant through a packing argument; that is, it can be derived directly from the separation parameter $\delta$. Below we portray the argument to show that given a separation parameter $\delta$ as above, there exists a density parameter $\rho$ independent of $h$.

**Lemma A.1.** Let $X$ be a sample set with fill distance $h$ and separation parameter $\delta$ (see Definition A.1 above). Then, there exists $\rho$ independent of $h$ such that
\[\# \{X \cap \overline{B}_{qh}(y)\} \leq \rho \cdot q^d, \quad q \geq 1, \quad y \in \mathbb{R}^n.\]

**Proof.** Let $I(x)$ denote the set of indices of samples $X \cap \overline{B}_{qh}(x)$, we wish to show that there exists $\rho$ such that $\#I(x) \leq \rho q^d$, where $q \geq 1$. Note, that any open ball $B_{\frac{1}{2}\delta h}(x_j)$ of radius $\frac{1}{2}\delta h$ centered at $x_j$ is disjoint to $B_{\frac{1}{2}\delta h}(x_k)$ (for $j, k \in I(x)$ and $j \neq k$). That is,
\[B_{\frac{1}{2}\delta h}(x_j) \cap B_{\frac{1}{2}\delta h}(x_k) = \emptyset.\]

On the other hand, for all $j \in I(x)$
\[B_{\frac{1}{2}\delta h}(x_j) \subset B_{qh + \frac{1}{2}\delta h}(x).\]

Thus,
\[\#I \cdot \text{Vol}\left(B_{\frac{1}{2}\delta h}(x_j)\right) \leq \text{Vol}\left(B_{qh + \frac{1}{2}\delta h}(x)\right)\]
\[\#I \cdot \text{Vol}(B_1(0)) \left(\frac{1}{2}\delta h\right)^d \leq \text{Vol}(B_1(0)) \left(qh + \frac{1}{2}\delta h\right)^d\]
\[\#I \leq \left(2q^{\frac{1}{\delta}} + 1\right)^d.\]

Now, if $\delta < 2$ we have
\[\#I \leq q^d \left(\frac{3}{\delta}\right)^d,\]
and if $\delta \geq 2$ we have
\[\leq q^d \left(\frac{3}{\rho}\right)^d.\]
A.2 Polynomial Reproduction and Interior Cone Condition

In a similar fashion, Wendland aims at proving similar convergence rates for every polynomial reproduction process.

**Definition A.2.** A process that defines for every set $X = \{x_i\}_{i=1}^n \subset \mathbb{R}^d$ a family of functions $u_i = u_i^X : \Omega \to \mathbb{R}$, $1 \leq i \leq n$, provides a local polynomial reproduction of degree $k$ on $\Omega$ if there exists constants $h_0, C_1, C_2 > 0$ such that

1. $\sum_{i=1}^n u_i(x)p(x_i) = p(x)$ for all $p \in \Pi^d_k \mid \Omega$,
2. $\sum_{i=1}^n |u_i(x)| \leq C_1$ for all $x \in \Omega$,
3. $u_i(x) = 0$ if $\|x - x_i\| > C_2 h$ and $x \in \Omega$

is satisfied for all $X$ with fill distance $h \leq h_0$.

Using this definition, and applying similar considerations to Levin, he shows the following theorem.

**Theorem A.2** (Theorem 3.2 from [24]). Suppose that $\Omega \subset \mathbb{R}^d$ is bounded. Define $\Omega^* \subset \mathbb{R}^d$ as the closure of $\cup_{x \in \Omega} B(x, C_2 h_0)$. Define $s_{f,X} = \sum_{i=1}^n u_i f(x_i)$, where $\{u_i\}$ is a local polynomial reproduction of order $(k-1)$ on $\Omega$. If $f \in C^k(\Omega^*)$ then there exists a constant $C > 0$ depending only on the constants from Definition A.2 such that

$$|f(x) - s_{f,X}(x)| \leq Ch^k |f|_{C^k(\Omega^*)}$$

for all $X$ with $h \leq h_0$. The semi-norm on the right hand side is defined by

$$|f|_{C^k(\Omega^*)} \triangleq \max_{|\alpha| = k} \|D^\alpha f\|_{L^\infty(\Omega^*)}.$$

Wendland then proceeds to show the argument regarding the MLS function approximation being a polynomial reproduction formula. Using the representation of the MLS approximation from (31), and by introducing the definition of a quasi-uniform sample set, Wendland establishes the fact that $s_{f,X}^{\text{MLS}}$ is indeed a polynomial reproduction of degree $(k-1)$ on a compact domain as detailed below.

**Definition A.3** (quasi-uniform sample set). A set of data sites $X = \{x_1, \ldots, x_n\}$ is said to be quasi-uniform with respect to a constant $c_{qu} > 0$ if

$$\delta_X \leq h_{X,\Omega} \leq c_{qu} \delta_X,$$

where $h_{X,\Omega}$ is the fill distance of Definition 3 and $\delta_X$ is the separation radius defined by

$$\delta_X := \frac{1}{2} \min_{i \neq j} \|x_i - x_j\|.$$

Note that this definition coincides with demanding just $h$ and $\delta$ conditions of the $h$-$\rho$-$\delta$ set of Definition A.1 above. Thus, by Lemma A.1 we get
Corollary A.1. $h$-$\rho$-$\delta$ sets are the same as quasi-uniform sample sets.

Then, using this notion of quasi-uniform sample set, Wendland shows the following result.

**Theorem A.3.** Suppose $\Omega \subset \mathbb{R}^d$ is compact and satisfies the interior cone condition with angle $\alpha \in (0, \pi/2)$ and radius $r > 0$. Fix $k \in \mathbb{N}$. Then there exist constants $h_0, C_1$, and $C_2$ depending only on $k, \alpha, r$ such that for a quasi-uniform sample set $X = \{x_i\}_{i=1}^n \subset \Omega$ with $h_{X,\Omega} \leq h_0$ and $\theta$ with support $2C_2h_{X,\Omega}$, the functions $a_i^*(x)$ provide local polynomial reproduction.

It is clear that, due to Theorem A.2 above, this implies that the MLS approximation has $O(h^k)$ convergence rates. The only detail that still needs to be clarified in this theorem is the interior cone condition.

**Definition A.4** (interior cone condition). A set $\Omega \subset \mathbb{R}^d$ is said to satisfy the interior cone condition if there exists an angle $\alpha \in (0, \pi/2)$ and a radius $r > 0$ such that for every $x \in \Omega$ a unit vector $\xi(x)$ exists such that the cone

$$C(x, \xi, \alpha, r) := \{x + \lambda y \mid y \in \mathbb{R}^d, \|y\| = 1, y^T \xi \geq \cos \alpha, \lambda \in [0, r]\}$$

is contained in $\Omega$.

The main difference between Levin’s proof of Theorem A.1 and Wendland’s proof of Theorem A.3 lies in the way they bound the expression

$$\sum_{i \in I(x)} |a_i^*(x)|.$$

Wendland takes a more generalized approach of showing that given a compact domain $\Omega$ that satisfies the interior cone condition, every sample set $X \subset \Omega$ has a polynomial reproduction formula (see Theorem 3.14 in [24]). Then he uses this result to show the desired bound (see the proof of Theorem 4.7 in [24]).

**A.3 Conclusion**

The main conclusion that we can draw from this discussion is that in order to obtain the approximation results reported in [8, 20, 21] it is sufficient to demand that the sample set is quasi-uniform. Furthermore, as can be seen from Levin’s result, the approximation is guaranteed when the weight function is compactly supported with a support of size $qh$, given that there exists at least $\binom{k+d}{k}$ samples in the support. Furthermore, this is also true of the analysis performed in [12], if we neglect the boundary (e.g., by taking a sub-domain $\Omega$ such that its closure is in the interior of $\Omega^*$ and restricting the error analysis to $\Omega$). As mentioned previous, since are dealing with boundaryless domain, this subtle issue is irrelevant to the discussions in the main text.