

## DISTANCE FUNCTIONS AND GEODESICS ON SUBMANIFOLDS OF $\mathbb{R}^d$ AND POINT CLOUDS\*

FACUNDO MÉMOLI† AND GUILLERMO SAPIRO†

**Abstract.** A theoretical and computational framework for computing intrinsic distance functions and geodesics on submanifolds of  $\mathbb{R}^d$  given by point clouds is introduced and developed in this paper. The basic idea is that, as shown here, intrinsic distance functions and geodesics on general co-dimension submanifolds of  $\mathbb{R}^d$  can be accurately approximated by extrinsic Euclidean ones computed inside a thin offset band surrounding the manifold. This permits the use of computationally optimal algorithms for computing distance functions in Cartesian grids. We use these algorithms, modified to deal with spaces with boundaries, and obtain a computationally optimal approach also for the case of intrinsic distance functions on submanifolds of  $\mathbb{R}^d$ . For point clouds, the offset band is constructed without the need to explicitly find the underlying manifold, thereby computing intrinsic distance functions and geodesics on point clouds while skipping the manifold reconstruction step. The case of point clouds representing noisy samples of a submanifold of Euclidean space is studied as well. All the underlying theoretical results are presented along with experimental examples for diverse applications and comparisons to graph-based distance algorithms.

**Key words.** geodesic distance, point clouds, manifolds, high dimensions, eikonal equations, random coverings, fast marching

**AMS subject classifications.** 65D18, 57A07, 68U05, 60D05

**DOI.** 10.1137/S003613990342877X

**1. Introduction.** One of the most popular sources of point clouds are three-dimensional (3D) shape acquisition devices, such as laser range scanners, with applications in geoscience, art (e.g., archival study), medicine (e.g., prosthetics), manufacturing (from cars to clothes), and security (e.g., recognition), among other disciplines. These scanners generally provide raw data in the form of (noisy) unorganized point clouds representing surface samples, and often produce very large numbers of points (tens of millions, for example, for the *David* model used in this paper). With the increasing popularity and very broad applications of this source of data, it is natural and important to work directly with such representations, without having to go through the intermediate step of fitting a surface to each (a step that can add computational complexity and introduce errors). See, for example, [11, 18, 20, 29, 33, 45, 46, 56, 58] for a few recent works with this type of data. Note that point clouds can also be used as primitives for visualization (e.g., [12, 33, 59]), as well as for editing [72].

Another important field where point clouds are found is in the representation of high-dimensional manifolds by samples (see, for example, [36, 44, 67]). This type of high-dimensional and general codimensional data appears in almost all disciplines, from computational biology to image analysis and financial data. Due to the extremely high number of dimensions in this case, it is impossible to perform manifold reconstruction, and the work needs to be done directly on the raw data, meaning the

---

\*Received by the editors May 28, 2003; accepted for publication (in revised form) August 25, 2004, published electronically April 14, 2005. This work was supported by Office of Naval Research grant ONR-N00014-97-1-0509, the Presidential Early Career Award for Scientists and Engineers (PECASE), and a National Science Foundation CAREER Award.

<http://www.siam.org/journals/siap/65-4/42877.html>

†Electrical and Computer Engineering Department, University of Minnesota, Minneapolis, MN 55455, and Instituto de Ingeniería Eléctrica, Universidad de la República, Montevideo, Uruguay (memoli@ece.umn.edu, guille@ece.umn.edu). The research of the first author is also supported by CSIC-Uruguay.

point cloud. Also in this area, large amounts of data are becoming available, from neuroscience experiments with neural recording of millions of points to large image and protein databases.

Note that in general a point cloud representation is codimension free, in contrast with other popular representations such as triangular meshes. Some operations, such as the union of point clouds acquired from multiple views, are much easier when performed directly on the representations than when performed on the triangular meshes obtained from them. This paper addresses one of the most fundamental operations in the study and processing of submanifolds of Euclidean space, the computation of intrinsic distance functions and geodesics. We show that these computations can be made by working directly with the point cloud, without the need for reconstructing the underlying manifold. Even if possible (for example, at low dimensions), the meshing operation is avoided, saving computations and improving accuracy. The distance computation itself is performed in computationally optimal time. We present the corresponding theoretical results, experimental examples, and basic comparisons to mesh-based distance algorithms.<sup>1</sup> The results are valid for general dimensions and codimensions, and for (underlying) manifolds with or without boundary. These results include the analysis of noisy point clouds obtained from sampling the manifold. We provide bounds on the accuracy of the computations that depend on the sampling rate and pattern as well as on the noise, thereby addressing real manifold sampling scenarios.

A number of key building blocks are part of the framework introduced here. The first one is based on the fact that distance functions intrinsic to a given submanifold of  $\mathbb{R}^d$  can be accurately approximated by Euclidean distance functions computed in a thin offset band that surrounds this manifold. This concept was first introduced in [49], where convergence results were given for hypersurfaces (codimension one submanifolds of  $\mathbb{R}^d$ ) without boundary. This result is reviewed in section 2. In this paper, we first extend these results to general codimensions and deal with manifolds with or without boundary in section 3. Interestingly, we also show that the approximation is true not only for the intrinsic distance function but also for the intrinsic minimizing geodesic.

The approximation of intrinsic distance functions (and geodesics) by extrinsic Euclidean ones permits us to compute them using computationally optimal algorithms in Cartesian grids (as long as the discretization operation is permitted, memorywise;<sup>2</sup> see sections 7.1 and 8). These algorithms are based on the fact that the distance function satisfies a Hamilton–Jacobi partial differential equation (see section 2), for which consistent and fast algorithms have been developed in Cartesian grids [35, 62, 63, 69].<sup>3</sup> (See [40] for extensions to triangular meshes, and [68] for other Hamilton–Jacobi equations.) That is, due to these results, we can use computationally optimal algorithms in Cartesian grids (with boundaries) also to compute distance functions, and from them geodesics,<sup>4</sup> intrinsic to a given manifold, and in a computationally

---

<sup>1</sup>Theoretical results on the accuracy of the technique for 3D mesh-based computationally optimal distance computation proposed in [40] have not been reported to the best of our knowledge.

<sup>2</sup>This is of course just a limitation of a straightforward implementation that doesn't avoid allocating memory to empty grids and works in the embedding dimension, and not a limitation of the theoretical and computational frameworks here developed.

<sup>3</sup>Tsitsiklis first described an optimal-control type of approach to solving the Hamilton–Jacobi equation, while independently Sethian and Helmsen both developed techniques based on upwind numerical schemes.

<sup>4</sup>Geodesics are the integral curves corresponding to the gradient directions of the intrinsic distance function, and are obtained by back-propagating in this gradient direction from the target point to the source point.

optimal fashion. Note that, in contrast with the popular Dijkstra algorithm, these numerical techniques are consistent; they converge to the true distance when the grid is refined. Dijkstra’s algorithm suffers from digitization bias due to metrication error when implemented on a grid (if no new graph edges are added to account for the new diagonals in each successive level of refinement of the grid); see [52, 53].

Once these basic results are available, we can then move on and deal with point clouds. The basic idea here is to construct the offset band directly from the point cloud, without the intermediate step of manifold reconstruction.<sup>5</sup> This is addressed in section 4 and section 5 for noise-free points and manifold samples, and in section 6 for points considered to be noisy samples of the manifold. In these cases, we explicitly compute the probability that the constructed offset band contains the underlying manifold. As we expect, this probability is a function of the number of point samples, the noise level, the size of the offset, and the basic geometric characteristics of the underlying manifold. This then covers the most realistic scenario, where the manifold is randomly sampled and the samples contain noise, thereby providing bounds that relate the error to the quality of the data. In the experimental section, section 7, we present a number of important applications. These applications are given to show the importance of this novel computational framework, and are by no means exhaustive. The data used in these examples were obtained from real acquisition devices, following laser scanning and photometric stereo. Concluding remarks are presented in section 8, where we also report the directions our research is taking.

To conclude this introduction, we should note that, to the best of our knowledge, the only additional work explicitly addressing the computation of distance functions and geodesics for point clouds is the one reported in [9, 67].<sup>6</sup> The comparison of performance in the presence of noise for our framework and the one proposed in [9, 67] is deferred to Appendix A.<sup>7</sup>

**2. Preliminary results and notation.** In this section we briefly review the main results in [49], where the idea of approximating intrinsic distances and geodesics by extrinsic ones was first introduced.

**2.1. Notation.** First, we introduce some basic notation that will be used throughout the article. For a compact and connected set  $\Omega \in \mathbb{R}^d$ ,  $d_\Omega(\cdot, \cdot)$  denotes the intrinsic distance between any two points of  $\Omega$ , measured by paths constrained to be in  $\Omega$ . We will also assume the convention that if  $A \subset \mathbb{R}^d$  is compact, and  $x, y$  are not both in  $A$ , then  $d_A(x, y) = D$  for some constant  $D \gg \max_{x, y \in A} d_A(x, y)$ . Given a  $k$ -dimensional submanifold  $\mathcal{M}$  of  $\mathbb{R}^d$ ,  $\Omega_{\mathcal{M}}^h$  denotes the set  $\{x \in \mathbb{R}^d : d(\mathcal{M}, x) \leq h\}$  (here the distance  $d(\cdot, \cdot)$  is the Euclidean one). This is basically an  $h$ -offset of  $\mathcal{M}$ . To state that the sequence of functions  $\{f_n(\cdot)\}_{n \in \mathbb{N}}$  uniformly converges to  $f(\cdot)$  as  $n \uparrow \infty$ , we frequently write  $f_n \xrightarrow{n} f$ . For a given event  $\mathcal{E}$ ,  $\mathbb{P}(\mathcal{E})$  stands for its probability of occurring. For a random variable (R.V. from now on)  $X$ , its mean value is denoted by  $\mathbb{E}(X)$ . By

<sup>5</sup>Recent results such as those reported in [57] provide efficient techniques for constructing such bands for point cloud data.

<sup>6</sup>In addition to studying the computation of distance functions on point clouds, [9, 67] address the important combination of this with multidimensional scaling for manifold analysis. Prior work on using geodesics and multidimensional scaling can be found in [61].

<sup>7</sup>While concluding this paper, we learned of a recent extension to Isomap reported in [31]. This paper is also mesh-based, and follows the geodesics approach in Isomap with a novel neighborhood/connectivity approach and a number of interesting theoretical results and novel dimensionality estimation contributions. Further analysis of Isomap, as a dimensionality reduction technique, can be found in [19].

$X \sim \mathbf{U}[A]$  we mean that the R.V.  $X$  is *uniformly distributed* in the set  $A$ . For a function  $f : \Omega \rightarrow \mathbb{R}$  and a subset  $A$  of  $\Omega$ ,  $f|_A : A \rightarrow \mathbb{R}$  denotes the restriction of  $f$  to  $A$ . For a smooth function  $f : \Omega \rightarrow \mathbb{R}$ ,  $Df$ ,  $D^2f$ , and  $D^3f$  stand for the first, second (Hessian matrix), and third differential, respectively, of  $f$ . Given a point  $x$  on the complete manifold  $\mathcal{S}$ ,  $B_{\mathcal{S}}(x, r)$  will denote the (intrinsic) open ball of radius  $r > 0$  centered at  $x$ , and  $B(y, r)$  will denote the *Euclidean* ball centered at  $y$  of radius  $r$ . Finally,  $\log x$  will denote the natural logarithm of  $x \in \mathbb{R}^+$ .

**2.2. Prelude.** In [49], we presented a new approach for the computation of weighted intrinsic distance functions on hyper-surfaces. We proved convergence theorems and addressed the fast, computationally optimal, computation of such approximations; see comments after Theorem 1 below. The key starting idea is that distance functions satisfy the (intrinsic) Eikonal equation, a particular case of the general class of Hamilton–Jacobi partial differential equations. Given  $p \in \mathcal{S}$  (a hypersurface in  $\mathbb{R}^d$ ), we want to compute  $d_{\mathcal{S}}(p, \cdot) : \mathcal{S} \rightarrow \mathbb{R}^+ \cup \{0\}$ , the intrinsic distance function from every point on  $\mathcal{S}$  to  $p$ . It is well known that the distance function  $d_{\mathcal{S}}(p, \cdot)$  satisfies, in the viscosity sense (see [47]), the equation

$$\begin{cases} \|\nabla_{\mathcal{S}} d_{\mathcal{S}}(p, x)\| = 1 & \forall x \in \mathcal{S}, \\ d_{\mathcal{S}}(p, p) = 0, \end{cases}$$

where  $\nabla_{\mathcal{S}}$  is the intrinsic differentiation (gradient). Instead of solving this intrinsic Eikonal equation on  $\mathcal{S}$ , we solve the corresponding extrinsic one in the offset band  $\Omega_{\mathcal{S}}^h$ :

$$\begin{cases} \|\nabla_x d_{\Omega_{\mathcal{S}}^h}(p, x)\| = 1 & \forall x \in \Omega_{\mathcal{S}}^h; \\ d_{\Omega_{\mathcal{S}}^h}(p, p) = 0, \end{cases}$$

where  $d_{\Omega_{\mathcal{S}}^h}(p, \cdot)$  is the Euclidean distance and therefore now the differentiation is the usual one.

**THEOREM 1** (see [49]). *Let  $p$  and  $q$  be any two points on the smooth (orientable, without boundary) hypersurface  $\mathcal{S}$ ; then  $|d_{\mathcal{S}}(p, q) - d_{\Omega_{\mathcal{S}}^h}(p, q)| \leq C_{\mathcal{S}}\sqrt{h}$  for small enough  $h$ ,<sup>8</sup> where  $C_{\mathcal{S}}$  is a constant depending on the geometry of  $\mathcal{S}$ .*

This simplification of the intrinsic problem into an extrinsic one permits the use of the computationally optimal algorithms mentioned in the introduction. This makes computing intrinsic distances, and from them geodesics, as simple and computationally efficient as computing them in Euclidean spaces. Moreover, as detailed in [49], the approximation of the intrinsic distance  $d_{\mathcal{S}}$  by the extrinsic Euclidean one  $d_{\Omega_{\mathcal{S}}^h}$  is never less accurate than the numerical error of these algorithms.

In [49], the result above was limited to hypersurfaces of  $\mathbb{R}^d$  (codimension one submanifolds of  $\mathbb{R}^d$ ) without boundary, and the theory was applied to implicit surfaces, where computing the offset band is straightforward. It is the purpose of the present work to extend Theorem 1 to deal with (1) submanifolds of  $\mathbb{R}^d$  of any codimension and possibly with boundary,<sup>9</sup> (2) convergence of geodesic curves in addition to distance functions, (3) submanifolds of  $\mathbb{R}^d$  represented as point clouds and (4) random sampling of submanifolds of  $\mathbb{R}^d$  in the presence of noise. We should note that Theorem 1 holds even when the metric is not the one inherited from  $\mathbb{R}^d$ , obtaining weighted distance

<sup>8</sup>“Small enough  $h$ ” means that  $h < 1/\max_i \kappa_i(\mathcal{S})$ , where  $\kappa_i(\mathcal{S})$  is the  $i$ th principal curvature of  $\mathcal{S}$ . This guarantees having smoothness in  $\partial\Omega_{\mathcal{S}}^h$ ; see [49].

<sup>9</sup>We will later impose some convexity conditions on the boundary in order to get rate of convergence estimates. However, the uniform convergence in itself doesn’t require other hypotheses beyond smoothness.

functions; see [49]. Although we will not present these new results in such generality, this is a simple extension that will be reported elsewhere.

**3. Submanifolds of  $\mathbb{R}^d$  with boundary.** We first extend Theorem 1 to more general manifolds, and we deal not only with distance functions but also with geodesics. The first extension is important for the learning of high-dimensional manifolds from samples and for scanned open volumes. The extension to geodesics is important for path planning on surfaces and for finding special curves such as crests and valleys; see [8, 49].

First we need to recall some results that will be key ingredients in our proofs below. All our results rest upon a certain degree of smoothness of geodesics in manifolds with boundary. We use “shortest path” and “minimizing geodesic” interchangeably.

**THEOREM 2** (see [1]). *Let  $\mathcal{M}$  be a  $C^3$  Riemannian manifold with  $C^1$  boundary  $\partial\mathcal{M}$ . Then any shortest path of  $\partial\mathcal{M}$  is  $C^1$ .*

We will eventually need more regularity on the geodesics than simply  $C^1$ . This is achieved by requiring more regularity of the boundary.

**THEOREM 3** ([48]). *Let  $\mathcal{U} : \mathbb{R}^d \rightarrow \mathbb{R}$  be a  $C^3$  function such that for some  $h \in \mathbb{R}$*

(i) *the interior of  $\{x \in \mathbb{R}^d \mid \mathcal{U}(x) = h\}$  is nonempty and there we have  $D\mathcal{U}(x) \neq 0$ .*

(ii) *the “obstacle”  $\{x \in \mathbb{R}^d \mid \mathcal{U}(x) \geq h\}$  is compact.*

*Let  $p$  and  $q$  be any two points in the same connected component of  $\{x \in \mathbb{R}^d \mid \mathcal{U}(x) \leq h\}$ ; then the shortest (constrained) path joining both points is  $C^1$  and has Lipschitz first derivative.*

We now present the usual definition of length, as follows.

**DEFINITION 1.** *Let  $\alpha : [a, b] \rightarrow \mathbb{R}^d$  be a curve, then we define its length  $\mathbf{L}(\alpha)$  as*

$$\mathbf{L}(\alpha) \triangleq \sup_{a=t_0 < \dots < t_N=b} \sum_{k=0}^{N-1} \|\alpha(t_{k+1}) - \alpha(t_k)\|.$$

*Remark 1.* Note that if  $\alpha$  is Lipschitz with constant  $\mathcal{L}_\alpha$ , then  $\mathbf{L}(\alpha) = \int_a^b \|\dot{\alpha}(t)\| dt$  and  $\mathbf{L}(\alpha) \leq \mathcal{L}_\alpha(b - a)$ .

**PROPOSITION 1.** *Let  $\mathcal{S}$  be a smooth compact submanifold of  $\mathbb{R}^d$  with boundary  $\partial\mathcal{S}$ . Let  $x, y$  be any two points in  $\mathcal{S}$ . Then  $d_{\Omega_S^h}(x, y)$  converges pointwise as  $h \downarrow 0$ .*

*Proof.* Since  $\Omega_S^h \subseteq \Omega_S^{h'}$  if  $h' \geq h$ , we have that  $d_{\Omega_S^h}(x, y) \geq d_{\Omega_S^{h'}}(x, y)$ . Also, for any  $h > 0$ ,  $d_{\Omega_S^h}(x, y) \leq d_S(x, y) \leq \text{diam}(\mathcal{S}) < +\infty$ . Hence, the sequence  $\{d_{\Omega_S^h}(x, y)\}_{h>0}$  (for fixed  $x$  and  $y$  over  $\mathcal{S}$ ) is bounded and nondecreasing, and therefore it converges to the supremum of its range.  $\square$

**THEOREM 4.** *Let  $\mathcal{S}$  be a compact  $C^2$  submanifold of  $\mathbb{R}^d$  with (possibly empty) smooth boundary  $\partial\mathcal{S}$ . Let  $x, y$  be any two points in  $\mathcal{S}$ . Then we have*

1. *uniform convergence of the distances:*

$$d_{\Omega_S^h} |_{\mathcal{S} \times \mathcal{S}}(\cdot, \cdot) \xrightarrow{h \downarrow 0} d_S(\cdot, \cdot);$$

2. *convergence of the geodesics: Let  $x$  and  $y$  be joined by a unique minimizing geodesic  $\gamma_S : [0, 1] \rightarrow \mathcal{S}$  over  $\mathcal{S}$ , and let  $\gamma_h : [0, 1] \rightarrow \Omega_S^h$  be a  $\Omega_S^h$ -minimizing geodesic; then*

$$\gamma_h \xrightarrow{h \downarrow 0} \gamma_S.$$

*Proof.* Given our hypothesis on  $\mathcal{S}$ , and according to [26], there exists  $H > 0$  such that  $\partial\Omega_{\mathcal{S}}^h$  is  $C^{1,1}$  for all  $0 < h \leq H$ . Then Theorem 2 guarantees that for  $0 < h \leq H$ ,  $\gamma_h : [0, 1] \rightarrow \Omega_{\mathcal{S}}^h$ , the  $\Omega_{\mathcal{S}}^h$  length-minimizing geodesic joining  $x$  and  $y$  is of class  $C^1$ . Since  $d_{\Omega_{\mathcal{S}}^h}(x, y) \leq d_{\mathcal{S}}(x, y) \leq \text{diam}(\mathcal{S}) < +\infty$  for any  $h \in (0, H]$ , we see that we can admit our  $\Omega_{\mathcal{S}}^h$ -geodesics to have Lipschitz constant  $\mathcal{L} \leq \text{diam}(\mathcal{S})$ . Obviously, the set  $\Omega_{\mathcal{S}}^H$  is bounded, and then the family  $\{\gamma_h\}_{0 < h \leq H}$  is bounded and equicontinuous. Hence, by the Ascoli–Arzelá theorem, there exist a subsequence  $\{\gamma_{h_k}\}_{k \in \mathbb{N}}$  and a curve  $\gamma_0 \in C^0([0, 1], \mathcal{S})$  such that  $\max_{t \in [0, 1]} \|\gamma_{h_k}(t) - \gamma_0(t)\| \xrightarrow{h_k \downarrow 0} 0$ .

Moreover, by writing  $|\gamma_0(t) - \gamma_0(t')| \leq |\gamma_{h_k}(t) - \gamma_0(t)| + |\gamma_{h_k}(t') - \gamma_0(t')| + \mathcal{L}|t - t'|$  and using the (pointwise) convergence of  $\gamma_{h_k}$  towards  $\gamma_0$ , we find that  $\mathcal{L}$  is also a Lipschitz constant for  $\gamma_0$ . Then we have  $\gamma_0 \in C^{0,1}([0, 1], \mathcal{S})$ .

Now, since  $\gamma_0$  lies on  $\mathcal{S}$  but may not be a shortest path, we have that its (finite) length is greater than or equal to  $d_{\mathcal{S}}(x, y)$ . We also have the trivial inequality  $d_{\mathcal{S}}(x, y) \geq d_{\Omega_{\mathcal{S}}^h}(x, y)$ . Putting this all together, we obtain

$$\mathbf{L}(\gamma_h) = d_{\Omega_{\mathcal{S}}^h}(x, y) \leq d_{\mathcal{S}}(x, y) \leq \mathbf{L}(\gamma_0).$$

Therefore

$$\limsup_{h \downarrow 0} \mathbf{L}(\gamma_h) = \limsup_{h \downarrow 0} d_{\Omega_{\mathcal{S}}^h}(x, y) \leq d_{\mathcal{S}}(x, y) \leq \mathbf{L}(\gamma_0).$$

Note that  $\mathbf{L}(\gamma_0) = \mathbf{L}(\lim_{h_k \downarrow 0} \gamma_{h_k}) \leq \liminf_{h_k \downarrow 0} \mathbf{L}(\gamma_{h_k})$ . This is the semicontinuity of length, an immediate consequence of its definition; see [41].

Since  $\liminf_{h_k \downarrow 0}(\cdot) \leq \limsup_{h_k \downarrow 0}(\cdot) \leq \limsup_{h \downarrow 0}(\cdot)$ , we see that  $\limsup_{h \downarrow 0} d_{\Omega_{\mathcal{S}}^h}(x, y) = \limsup_{h \downarrow 0} \mathbf{L}(\gamma_h)$  equals  $d_{\mathcal{S}}(x, y)$  for all  $x$  and  $y$  in  $\mathcal{S}$ . From Proposition 1, we find that in fact  $\lim_{h \downarrow 0} d_{\Omega_{\mathcal{S}}^h}(x, y)$  exists and equals  $d_{\mathcal{S}}(x, y)$ . Then, we have that the function  $d_{\Omega_{\mathcal{S}}^h}|_{\mathcal{S} \times \mathcal{S}}(\cdot, \cdot)$  satisfies the following:

- (i)  $d_{\Omega_{\mathcal{S}}^h}|_{\mathcal{S} \times \mathcal{S}} : \mathcal{S} \times \mathcal{S} \rightarrow \mathbb{R} \cup \{0\}$  is continuous for each  $H > h > 0$ ;
- (ii) for each  $(x, y) \in \mathcal{S} \times \mathcal{S}$ ,  $\{d_{\Omega_{\mathcal{S}}^h}|_{\mathcal{S} \times \mathcal{S}}(x, y)\}_h$  is nondecreasing;
- (iii)  $d_{\Omega_{\mathcal{S}}^h}|_{\mathcal{S} \times \mathcal{S}}(\cdot, \cdot)$  converges pointwise towards  $d_{\mathcal{S}}(\cdot, \cdot)$ , which is continuous.

Then by Dini’s uniform convergence theorem (see [6]) we can conclude that the convergence is uniform.

We can also see that  $\gamma_0$  must be a minimizing geodesic of  $\mathcal{S}$  since from the above chain of equalities  $\mathbf{L}(\gamma_0) = d_{\mathcal{S}}(x, y)$ . Then, if there was only one such curve joining  $x$  with  $y$ , we would have uniform convergence (along any subsequence!) of  $\gamma_h$  towards  $\gamma_0$ .<sup>10</sup>  $\square$

*Remark 2.* In Theorem 4, the convergence (of distances) is uniform, but we will have forfeited rate of convergence estimates unless we impose additional conditions on  $\partial\mathcal{S}$ , as we do in Corollary 3. Note that the new setting is wider than the one considered in Theorem 1 since the codimension of the underlying manifold is not necessarily 1. This is very important for applications such as dimensionality reduction, where the dimension of the underlying manifold is unknown beforehand.

**COROLLARY 1.** *Let  $\mathcal{S}$  and  $\partial\mathcal{S}$  satisfy the hypotheses of Theorem 4. Let  $\{\Sigma_i\}_{i \in \mathbb{N}}$  be a family of compact of sets in  $\mathbb{R}^d$  such that  $\mathcal{S} \subseteq \Sigma_i$  for all  $i \in \mathbb{N}$  and  $d_{\mathcal{H}}(\Sigma_i, \mathcal{S}) \xrightarrow{i \uparrow +\infty} 0$ .*

<sup>10</sup>This follows from the fact that uniform convergence of  $\gamma_h$  to  $\gamma_0$  is equivalent to the statement that for any subsequence  $\{\gamma_{h_i}\}$  there exists a further subsubsequence  $\{\gamma_{h_{i_k}}\}$  uniformly converging to  $\gamma_0$ .

Then,

$$d_{\Sigma_i}(\cdot, \cdot)|_{\mathcal{S} \times \mathcal{S}} \xrightarrow{i \uparrow +\infty} d_{\mathcal{S}}(\cdot, \cdot),$$

where  $d_{\mathcal{H}}$  stands for the Hausdorff distance between sets.

We now present a uniform rate of convergence result for the distance in the band in the case  $\partial\mathcal{S} = \emptyset$ , and from this we deduce Corollary 3 below, which deals with the case  $\partial\mathcal{S} \neq \emptyset$ . This result generalizes the one presented in [49] because it allows for any codimension.

**THEOREM 5.** *Under the hypotheses of Theorem 4, with  $\partial\mathcal{S} = \emptyset$ , we have that for small enough  $h > 0$ ,*

$$(1) \quad \max_{(x,y) \in \mathcal{S} \times \mathcal{S}} \left| d_{\Omega_h^{\mathcal{S}}}|_{\mathcal{S} \times \mathcal{S}}(x, y) - d_{\mathcal{S}}(x, y) \right| \leq C_{\mathcal{S}} \sqrt{h},$$

where the constant  $C_{\mathcal{S}}$  does not depend on  $h$ . Also, we have the “relative” rate of convergence bound

$$(2) \quad 1 \leq \sup_{\substack{x,y \in \mathcal{S} \\ x \neq y}} \frac{d_{\mathcal{S}}(x, y)}{d_{\Omega_h^{\mathcal{S}}}(x, y)} \leq 1 + C_{\mathcal{S}} \sqrt{h}.$$

*Proof.* This is a remake of our proof of the main theorem in [49]; therefore we skip some technical details which can be found there. Throughout the proof we will sometimes write  $d_h$  instead of  $d_{\Omega_h^{\mathcal{S}}}$  for the sake of notational simplicity. We will denote by  $k (\leq n - 1)$  the dimension of  $\mathcal{S}$ .

Let  $\gamma_0$  be the arc length parametrized  $\mathcal{S}$ -shortest path joining the points  $x, y \in \mathcal{S}$ ; clearly, we have  $\text{trace}(\gamma_0) \subset \mathcal{S}$ . Let  $\gamma_h$  be the  $\Omega_h^{\mathcal{S}}$  arc length parametrized shortest path joining  $x$  and  $y$ , which, as we know from Theorem 4, uniformly converges toward  $\gamma_0$ . For a number  $H$  as in the proof of Theorem 4, we have  $\gamma_h \in C^{1,1}([0, d_h], \mathcal{S})$ , and also  $\eta : \Omega_{\mathcal{S}}^H \rightarrow \mathbb{R}$  defined by  $\eta(x) \triangleq \frac{1}{2}d^2(x, \mathcal{S})$  is smooth; see Appendix B. We define the projection operator  $\Pi_{\mathcal{S}} : \Omega_{\mathcal{S}}^H \rightarrow \mathcal{S}$  by  $\Pi_{\mathcal{S}}(x) = x - D\eta(x)$ . We refer the reader to Appendix B for properties of  $\Pi_{\mathcal{S}}$  and  $\eta$  which we use below.

Now,  $d_{\Omega_h^{\mathcal{S}}}(x, y) = \mathbf{L}(\gamma_h) \leq d_{\mathcal{S}}(x, y) \leq \mathbf{L}(\Pi_{\mathcal{S}}(\gamma_h))$ ; then

$$\begin{aligned} d_{\mathcal{S}}(x, y) - d_{\Omega_h^{\mathcal{S}}}(x, y) &\leq |\mathbf{L}(\Pi_{\mathcal{S}}(\gamma_h)) - \mathbf{L}(\gamma_h)| \\ &\leq \int_0^{d_h} \left\| \overline{\Pi_{\mathcal{S}}(\gamma_h(t)) - \gamma_h(t)} \right\| dt \\ &= \int_0^{d_h} \left\| \overline{D\eta(\gamma_h(t))} \right\| dt \\ &\leq \sqrt{d_h \int_0^{d_h} \dot{V}(t) \cdot \dot{V}(t) dt} \quad (\text{by Cauchy-Schwarz inequality}) \\ &\leq \sqrt{d_h \int_0^{d_h} V(t) \cdot \ddot{V}(t) dt} \quad (\text{integrating by parts; see below}), \end{aligned}$$

where  $V(t) \triangleq D\eta(\gamma_h(t))$  and  $V(0) = V(d_h) = 0$ ; see Appendix B.

Also  $V(t) = D^2\eta(\gamma_h(t))\dot{\gamma}(t)$ , and since  $\gamma_h$  is Lipschitz and  $\eta$  is smooth,  $\ddot{V}(t)$  exists almost everywhere and  $\ddot{V}(t) = D^3\eta(\gamma_h(t))[\dot{\gamma}_h(t), \dot{\gamma}_h(t)] + D^2\eta(\gamma_h(t))\ddot{\gamma}(t)$  at points of

existence. Then since  $D^3\eta D\eta = D^2\eta(I - D^2\eta)$  and  $D^2\eta D\eta = D\eta$  (see Appendix B),

$$\begin{aligned} V \cdot \ddot{V} &= D^3\eta(\gamma_h)[D\eta(\gamma_h), \dot{\gamma}_h, \dot{\gamma}_h] + D^2\eta[\ddot{\gamma}_h, D\eta(\gamma_h)] \\ &= (D^2\eta(\gamma_h) (I - D^2\eta(\gamma_h))) [\dot{\gamma}_h, \dot{\gamma}_h] + \ddot{\gamma}_h \cdot D\eta(\gamma_h). \end{aligned}$$

The matrix  $\Lambda(t) \triangleq D^2\eta(\gamma_h(t))(I - D^2\eta(\gamma_h(t)))$  filters out normal components and has eigenvalues associated with the tangential bundle given by

$$\lambda_i(t) = \frac{d(t)\lambda_i(0)}{(1 + d(t)\lambda_i(0))^2} \quad \text{for } 1 \leq i \leq k,$$

where we let  $d(t) = d(\gamma_h(t), \mathcal{S})$ . Note that  $\max_{1 \leq i \leq k} |\lambda_i(t)|$  can be bounded by  $d(t)$  times a certain finite constant  $K'$  independent of  $h$ .

On the other hand, we can bound  $|\ddot{\gamma}_h(t)|$  almost anywhere by a finite constant, say  $K$ , which takes into account the maximal curvature of all the boundaries  $\partial\Omega_{\mathcal{S}}^h$ ,  $0 < h < H$ , but does not depend on  $h$ .

Putting all this together, we find (recall that  $\|D\eta(x)\| = \sqrt{2\eta(x)} = d(x, \mathcal{S})$ ; see Appendix B)

$$\begin{aligned} \left(d_{\mathcal{S}}(x, y) - d_{\Omega_{\mathcal{S}}^h}(x, y)\right)^2 &\leq d_h \int_0^{d_h} \Lambda(t)[\dot{\gamma}_h, \dot{\gamma}_h] dt \\ &\quad + d_h \int_0^{d_h} \|\ddot{\gamma}_h\| \|D\eta(\gamma_h)\| dt \\ &\leq K' \max_{t \in [0, d_h]} d(t) d_h^2 + K \max_{t \in [0, d_h]} d(t) d_h^2. \end{aligned}$$

Now, remembering that  $d_h$  stands for  $d_{\Omega_{\mathcal{S}}^h}(x, y)$ , that  $\text{trace}(\gamma_h) \subset \Omega_{\mathcal{S}}^h$ , and defining  $C = K + K'$ , we arrive with only a little simple additional work, at the relations (1) or (2).  $\square$

*Remark 3.* Note that, as the simple case of a circle in the plane shows, the rate of convergence is at most  $C \cdot h$ .

We immediately obtain the following corollary, which will be useful ahead.

**COROLLARY 2.** *Let  $p \in \mathcal{S}$  and  $r \leq H$ ; then  $B(p, r) \cap \mathcal{S} \subseteq B_{\mathcal{S}}(p, r(1 + C_{\mathcal{S}}\sqrt{r}))$ .*

*Proof.* Let  $q \in B(p, r) \cap \mathcal{S}$ ; then by (2),  $d_{\mathcal{S}}(p, q) \leq d_{\Omega_{\mathcal{S}}^r}(p, q)(1 + C_{\mathcal{S}}\sqrt{r})$ . However,  $q \in B(p, r) \subset \Omega_{\mathcal{S}}^r$ , and thus  $d_{\Omega_{\mathcal{S}}^r}(p, q) = \|p - q\| \leq r$ , which completes the proof.  $\square$

**DEFINITION 2.** (see [21]) *We say that the compact manifold  $\mathcal{S}$  with boundary  $\partial\mathcal{S}$  is strongly convex if for every pair of points  $x$  and  $y$  in  $\mathcal{S}$  there exists a unique minimizing geodesic joining them whose interior is contained in the interior of  $\mathcal{S}$ .*

Using basically the same procedure as in Theorem 5 with the convexity hypotheses above, we can prove the following corollary, whose (sketched) proof is presented in Appendix C.

**COROLLARY 3.** *Under the hypotheses of Theorem 2, and assuming  $\mathcal{S}$  to be strongly convex, we have for small enough  $h > 0$  the same conclusions of Theorem 5 (rate of convergence).*

*Remark 4.* Note that in case  $\partial\mathcal{S} \neq \emptyset$  is not strongly convex, then obviously the same statement of Corollary 3 remains valid for any strongly convex subset of  $\mathcal{S}$ .

To conclude, in this section we extended the results in [49] to geodesics and distance functions in general codimension manifolds with or without (smooth) boundary, thereby covering all possible manifolds in common shape, graphics, visualization, and

learning applications.<sup>11</sup> We are now ready to extend this to manifolds represented as point clouds.

**4. Distance functions on point clouds.** We are now interested in making computations on manifolds represented as point clouds, i.e., *sampled manifolds*. In the case of this paper we will restrict ourselves to the computation of intrinsic distances.<sup>12</sup> Let  $\mathcal{P}_n \triangleq \{p_1, \dots, p_n\}$  be a set of  $n$  different points sampled from the compact submanifold  $\mathcal{S}$  and define<sup>13</sup>

$$\Omega_{\mathcal{P}_n}^h \triangleq \bigcup_{i=1}^n B(p_i, h).$$

Let  $h$  and  $\mathcal{P}_n$  be such that  $\mathcal{S} \subseteq \Omega_{\mathcal{P}_n}^h$ . We then have  $(\mathcal{S} \subseteq) \Omega_{\mathcal{P}_n}^h \subseteq \Omega_{\mathcal{S}}^h$ . We now want to consider  $d_{\Omega_{\mathcal{P}_n}^h}(p, q)$  for any pair of points  $p, q \in \mathcal{S}$  and prove some kind of proximity to the real distance  $d_{\mathcal{S}}(p, q)$ . The argument carries over easily since

$$d_{\Omega_{\mathcal{S}}^h}(p, q) \leq d_{\Omega_{\mathcal{P}_n}^h}(p, q) \leq d_{\mathcal{S}}(p, q),$$

and hence

$$(3) \quad 0 \leq d_{\mathcal{S}}(p, q) - d_{\Omega_{\mathcal{P}_n}^h}(p, q) \leq d_{\mathcal{S}}(p, q) - d_{\Omega_{\mathcal{S}}^h}(p, q),$$

and the rightmost quantity can be bounded by  $C_{\mathcal{S}} h^{1/2}$  (see section 3) in the case that  $\partial\mathcal{S}$  is either convex or void. In general, without hypotheses on  $\partial\mathcal{S}$  other than some degree of smoothness, we can also work out uniform convergence since by virtue of Theorem 4 the upper bound in (3) uniformly converges to 0. The key condition is  $\mathcal{S} \subset \Omega_{\mathcal{P}_n}^h$ , something that can obviously be coped with using the compactness of  $\mathcal{S}$ .<sup>14</sup> We can then state the following claim.

**THEOREM 6** (uniform convergence for point clouds). *Let  $\mathcal{S}$  be a compact smooth submanifold of  $\mathbb{R}^d$  possibly with boundary  $\partial\mathcal{S}$ . Then the following hold:*

1. General case: *Given  $\varepsilon > 0$ , there exists  $h_{\varepsilon} > 0$  such that for all  $0 < h \leq h_{\varepsilon}$  one can find finite  $n(h)$  and a set of points  $\mathcal{P}_{n(h)}(h) = \{p_1(h), \dots, p_{n(h)}(h)\}$  sampled from  $\mathcal{S}$  such that*

$$\max_{p, q \in \mathcal{S}} \left( d_{\mathcal{S}}(p, q) - d_{\Omega_{\mathcal{P}_{n(h)}(h)}^h}(p, q) \right) \leq \varepsilon.$$

2.  $\partial\mathcal{S}$  is either void or convex: *For every sufficiently small  $h > 0$  one can find finite  $n(h)$  and a set of points  $\mathcal{P}_{n(h)}(h) = \{p_1(h), \dots, p_{n(h)}(h)\}$  sampled from  $\mathcal{S}$  such that*

$$\max_{p, q \in \mathcal{S}} \left( d_{\mathcal{S}}(p, q) - d_{\Omega_{\mathcal{P}_{n(h)}(h)}^h}(p, q) \right) \leq C_{\mathcal{S}} \sqrt{h}.$$

<sup>11</sup>Although in this paper we consider only manifolds with constant codimension, many of the results are extendible to variable codimensions, and this will be reported elsewhere.

<sup>12</sup>Note that having the intrinsic distance allows us to compute basic intrinsic properties of the manifold; see e.g., [13].

<sup>13</sup>The balls now used are defined with respect to the metric of  $\mathbb{R}^d$ ; they are not intrinsic.

<sup>14</sup>By compactness, given  $h > 0$ , we can find finite  $N(h)$  and points  $p_1, p_2, \dots, p_{N(h)} \in \mathcal{S}$  such that  $\mathcal{S} = \cup_{i=1}^{N(h)} B_{\mathcal{S}}(p_i, h)$ . But since for  $p \in \mathcal{S}$ ,  $B_{\mathcal{S}}(p, h) \subset B(p, h) \cap \mathcal{S}$ , we also get  $\mathcal{S} \subset \cup_{i=1}^{N(h)} B(p_i, h)$ .

In practice, one must worry about both the number of points and the radii of the balls. Obviously, there is a tradeoff between both quantities. If we want to use few points, in order to cover  $\mathcal{S}$  with the balls we have to increase the value of the radius. Clearly, there exists a value  $H$  such that for values of  $h$  smaller than  $H$  we do not change the topology; see [3, 4, 5]. This implies that the number of points must be larger than a certain lower bound. This result can be generalized to ellipsoids which can be locally adapted to the geometry of the point cloud [15], or from minimal spanning trees. Note that we are interested in the smallest possible offset of the point cloud that covers  $\mathcal{S}$ . Further comments on this are presented below and are also the subject of current efforts to be reported elsewhere.

The practical significance of the previous Theorem is clear. Part 1 says that in general, given a desired precision for the computation of the distance, we have a maximum *nonzero* value for the radius of all the balls, below which we can always find a *finite* number of points sampled from the manifold for which the “ $\Omega$ -set” formed by those points achieves the desired accuracy;<sup>15</sup> that is, we can choose the radius at our convenience *within* a certain range which depends on this level of accuracy. Part 2 says more, since it actually links  $\varepsilon$  to  $h_\varepsilon$ . It basically says that the radius of the balls must be of the order of the square of the desired error.

**5. Extension to random sampling of manifolds.** In practice, we really do not have too much control over the way in which points are sampled by the acquisition device (e.g., scanner) or given by the learned sampled data. Therefore it is more realistic to make a probabilistic model of the situation and then try to conveniently estimate the probability of achieving a prescribed level of accuracy as a function of the number of points and the radii of the balls. It will be interesting to see how geometric quantities of  $\mathcal{S}$  enter in those bounds we will establish. However, since the bounds are based in local volume computations and all manifolds are locally Euclidean, those curvature dependent quantities will be asymptotically negligible.

We now present a simple model for the current setting, while results for other models can be developed from the derivations below. Here we assume that the points in  $\mathcal{P}_n$  are independently and identically sampled on the submanifold  $\mathcal{S}$  with the uniform probability law;<sup>16</sup> we will write this as  $p_i \sim \mathbf{U}[\mathcal{S}]$ . For simplicity of exposition, we will restrict ourselves to the case when  $\mathcal{S}$  has no boundary.<sup>17</sup> Also, we deal only with uniform independently and identically distributed (i.i.d.) sampling; results for other sampling models, including those adapted to the manifold geometry, can be easily obtained following the developments below and will be reported elsewhere.

We have to define the way in which we are going to measure accuracy. A possibility for such a measure is (for each  $\varepsilon > 0$ )

$$(4) \quad \mathbb{P} \left( \max_{p, q \in \mathcal{S}} \left( d_{\mathcal{S}}(p, q) - d_{\Omega_{\mathcal{P}_n}^h}(p, q) \right) > \varepsilon \right).$$

There is a potential problem with this way of testing accuracy, since we are assuming that when we use the approximate distance, we will be evaluating it on  $\mathcal{S}$ . This might seem a bit awkward since we don't exactly know all the surface but just

<sup>15</sup>We are considering the case when all the balls have the same radii.

<sup>16</sup>This means that for any subset  $A \subseteq \mathcal{S}$  and any  $p_i \in \mathcal{P}_n$ ,  $\mathbb{P}(p_i \in A) = \frac{\mu(A)}{\mu(\mathcal{S})}$ , where  $\mu(\cdot)$  stands for the measure (area/volume) of the set.

<sup>17</sup>In order to extend the results in this section to the case  $\partial\mathcal{S} \neq \emptyset$ , the same considerations discussed in [9] remain valid in our case.

some points on it. Moreover, a more natural and real-problem-motivated approach would be to measure the discrepancy over  $\mathcal{P}_n$  itself (see section 7 ahead), over part of this set, or over another *trial* set of points  $\mathcal{Q}_m$ .

However, since for any set of points  $\mathcal{Q}_m \subset \mathcal{S}$  we have that the following inclusion of events,

$$\left\{ \max_{p,q \in \mathcal{Q}_m} \left( d_{\mathcal{S}}(p,q) - d_{\Omega_{\mathcal{P}_n}^h}(p,q) \right) > \varepsilon \right\} \subseteq \left\{ \max_{p,q \in \mathcal{S}} \left( d_{\mathcal{S}}(p,q) - d_{\Omega_{\mathcal{P}_n}^h}(p,q) \right) > \varepsilon \right\},$$

holds, bounding (4) suffices for dealing with any of the possibilities mentioned above. Note that we are somehow considering  $d_{\Omega_{\mathcal{P}_n}^h}$  defined for all pairs of points in  $\mathcal{S} \times \mathcal{S}$ , even if it might happen that  $\mathcal{S} \cap \Omega_{\mathcal{P}_n}^h \neq \mathcal{S}$ . In any case we extend  $d_{\Omega_{\mathcal{P}_n}^h}$  to all  $\Omega_{\mathcal{S}}^h \times \Omega_{\mathcal{S}}^h$  by a large constant, say  $k \operatorname{diam}(\mathcal{S})$ ,  $k \gg 1$ .

Let us spell out a few definitions so as to avoid an overload of notation:

$$(5) \quad \mathcal{E}_\varepsilon \triangleq \left\{ \max_{p,q \in \mathcal{S}} \left( d_{\mathcal{S}}(p,q) - d_{\Omega_{\mathcal{P}_n}^h}(p,q) \right) > \varepsilon \right\},$$

$$(6) \quad \mathcal{J}_{h,n} \triangleq \{ \mathcal{S} \subseteq \Omega_{\mathcal{P}_n}^h \}.$$

Now, since  $\mathcal{E}_\varepsilon = (\mathcal{E}_\varepsilon \cap \mathcal{J}_{h,n}) \cup (\mathcal{E}_\varepsilon \cap \mathcal{J}_{h,n}^c)$ , using the union bound and then Bayes rule, we have

$$\begin{aligned} \mathbb{P}(\mathcal{E}_\varepsilon) &\leq \mathbb{P}(\mathcal{E}_\varepsilon \cap \mathcal{J}_{h,n}) + \mathbb{P}(\mathcal{E}_\varepsilon \cap \mathcal{J}_{h,n}^c) \\ &= \mathbb{P}(\mathcal{E}_\varepsilon | \mathcal{J}_{h,n}) \mathbb{P}(\mathcal{J}_{h,n}) + \mathbb{P}(\mathcal{E}_\varepsilon | \mathcal{J}_{h,n}^c) \mathbb{P}(\mathcal{J}_{h,n}^c) \end{aligned}$$

↓

$$(7) \quad \mathbb{P}(\mathcal{E}_\varepsilon) \leq \mathbb{P}(\mathcal{E}_\varepsilon | \mathcal{J}_{h,n}) + \mathbb{P}(\mathcal{J}_{h,n}^c).$$

It is clear now that we must find a convenient lower bound for the second term in the previous expression, the probability of covering all  $\mathcal{S}$  with the union of balls. (The first term will be dealt with using the convergence theorems presented in previous sections.) For this we need a few lemmas.

LEMMA 1. *Let  $K$  be an upper bound for the sectional curvatures of  $\mathcal{S}$  ( $\operatorname{diam}(\mathcal{S}) = k$ ) and  $x \in \mathcal{S}$  be a fixed point. Then, under the hypotheses on  $\mathcal{P}_n$  described above, there exist a constant  $\omega_k > 0$  and a function  $\theta_{\mathcal{S}}(\cdot)$  with  $\lim_{h \downarrow 0} \frac{\theta_{\mathcal{S}}(h)}{h^{k+1}} = 0$  such that for small enough  $h > 0$*

$$(8) \quad \mathbb{P}(\{x \notin \Omega_{\mathcal{P}_n}^h \cap \mathcal{S}\}) \leq \left( 1 - \frac{\omega_k h^k + \theta_{\mathcal{S}}(h)}{\mu(\mathcal{S})} \right)^n.$$

Moreover, one can further expand the right-hand side of (8) as

$$\left( 1 - \frac{\omega_k h^k (1 - K c_k h^2) + \phi_{\mathcal{S}}(h)}{\mu(\mathcal{S})} \right)^n$$

for some  $c_k$  depending only on the dimension  $k$  of  $\mathcal{S}$  and a function  $\phi_{\mathcal{S}}$  such that  $\frac{\phi_{\mathcal{S}}(h)}{h^{k+2}} \rightarrow 0$  as  $h \downarrow 0$ .

*Proof.*

$$(9) \quad \mathbb{P}(\{x \notin \Omega_{\mathcal{P}_n}^h \cap \mathcal{S}\}) = \mathbb{P}\left(\left\{\bigcap_{i=1}^n \{x \notin B(p_i, h) \cap \mathcal{S}\}\right\}\right)$$

$$(10) \quad = \mathbb{P}\left(\left\{\bigcap_{i=1}^n \{p_i \notin B(x, h) \cap \mathcal{S}\}\right\}\right)$$

$$(11) \quad = \prod_{i=1}^n \mathbb{P}(\{p_i \notin B(x, h) \cap \mathcal{S}\})$$

$$(12) \quad = \prod_{i=1}^n (1 - \mathbb{P}(\{p_i \in B(x, h) \cap \mathcal{S}\})).$$

Since  $B_{\mathcal{S}}(x, h) \subseteq B(x, h) \cap \mathcal{S}$ ,<sup>18</sup> then  $\mu(\mathcal{S} \cap B(x, h)) \geq \mu(B_{\mathcal{S}}(x, h))$ . On the other hand, note that

$$\begin{aligned} \mathbb{P}(\{p_i \in B(x, h) \cap \mathcal{S}\}) &= \frac{\mu(\mathcal{S} \cap B(x, h))}{\mu(\mathcal{S})} \\ &\geq \frac{\mu(B_{\mathcal{S}}(x, h))}{\mu(\mathcal{S})}. \end{aligned}$$

Finally, as shown in Appendix D, one can lower bound  $\mu(B_{\mathcal{S}}(x, h))$  using information on the curvatures of  $\mathcal{S}$ , by means of the Bishop–Günther volume comparison theorem. More precisely, we can write

$$\mu(B_{\mathcal{S}}(x, h)) \geq \min_{\zeta \in \mathcal{S}} \mu(B_{\mathcal{S}}(\zeta, h)) \geq \omega_k h^k + \theta_{\mathcal{S}}(h),$$

where  $\frac{\theta_{\mathcal{S}}(h)}{h^q} \rightarrow 0$  when  $h \rightarrow 0$  for  $q \leq k + 1$ . Therefore, from (9) we obtain

$$\mathbb{P}(\{x \notin \Omega_{\mathcal{P}_n}^h \cap \mathcal{S}\}) \leq \left(1 - \frac{\omega_k h^k + \theta_{\mathcal{S}}(h)}{\mu(\mathcal{S})}\right)^n.$$

The last assertion follows from Proposition 3.  $\square$

*Remark 5.* Note that we cannot, however, from (8), conclude that  $\mathbb{P}(\mathcal{S} \not\subseteq \Omega_{\mathcal{P}_n}^h) \leq \left(1 - \frac{\omega_k h^k + \theta_{\mathcal{S}}(h)}{\mu(\mathcal{S})}\right)^n$ . In order to upper bound  $\mathbb{P}(\mathcal{S} \not\subseteq \Omega_{\mathcal{P}_n}^h)$  we will first estimate  $\mathbb{P}(B_{\mathcal{S}}(x, \delta) \not\subseteq \Omega_{\mathcal{P}_n}^h)$  for any  $x \in \mathcal{S}$  and small  $\delta > 0$ . Then we will use the compactness of  $\mathcal{S}$  by covering it with a finite  $\delta$ -net consisting of  $\mathcal{N}(\mathcal{S}, \delta)$  points, and conclude by using the union bound. Yet another intermediate step will therefore be to estimate the covering number  $\mathcal{N}(\mathcal{S}, \delta)$ .

LEMMA 2. *Under the hypotheses of the previous lemma, let  $\delta \in (0, h)$ ; then*

$$(13) \quad \mathbb{P}(B_{\mathcal{S}}(x, \delta) \not\subseteq \Omega_{\mathcal{P}_n}^h) \leq \left(1 - \frac{\omega_k (h - \delta)^k + \theta_{\mathcal{S}}(h - \delta)}{\mu(\mathcal{S})}\right)^n.$$

*Proof.* We find  $\alpha$  and  $\beta$  such that  $\{B_{\mathcal{S}}(q, \delta) \subseteq \Omega_{\mathcal{P}_n}^h\} \supseteq \{q \in \Omega_{\mathcal{P}_n}^{\alpha h + \beta \delta}\}$ . Note first that for any  $x \in B_{\mathcal{S}}(q, \delta)$ ,  $|x - q| \leq d_{\mathcal{S}}(x, q) \leq \delta$ . Assume that the event  $\{q \in \Omega_{\mathcal{P}_n}^{\alpha h + \beta \delta}\}$

<sup>18</sup>Consider  $z \in B_{\mathcal{S}}(x, h)$ ; then  $d_{\mathcal{S}}(x, z) \leq h$ , but always  $d(x, z) \leq d_{\mathcal{S}}(x, z)$ , and thus  $d(x, z) \leq h$ , which implies  $z \in B(x, h) \cap \mathcal{S}$ .

holds. Then for some  $p_r \in \mathcal{P}_n$ ,  $q \in B(p_r, \alpha h + \beta \delta)$ ; that is,  $|q - p_r| \leq \alpha h + \beta \delta$ . Now, note that

$$|x - p_r| \leq |x - q| + |q - p_r| \leq \alpha h + (\beta + 1)\delta.$$

If we force the rightmost number to be  $h$ , we find that we must have  $(1 + \beta)\delta = (1 - \alpha)h$ , and then  $\alpha h + \beta \delta = h - \delta$ . Then we have found  $B_S(q, \delta) \subseteq B(p_r, h - \delta) \subset \Omega_{\mathcal{P}_n}^h$ . Hence (using (8)),  $\mathbb{P}(B_S(q, \delta) \subseteq \Omega_{\mathcal{P}_n}^h) \geq \mathbb{P}(q \in \Omega_{\mathcal{P}_n}^{h-\delta} \geq 1 - (1 - \frac{\omega_k(h-\delta)^k + \theta_S(h-\delta)}{\mu(\mathcal{S})})^n)$ .  $\square$

We also need the next lemma, whose proof is deferred to Appendix C.

LEMMA 3 (bounding the covering number). *Under the hypotheses of Lemma 2 and further assuming  $\mathcal{S}$  to be compact, we have that for any small enough  $\delta > 0$  there exists a  $\delta$ -covering of  $\mathcal{S}$  with cardinality*

$$(14) \quad \mathcal{N}(\mathcal{S}, \delta) \leq \frac{\mu(\mathcal{S})}{\omega_k(\delta/2)^k + \theta_S(\delta/2)}.$$

PROPOSITION 2. *Let the set of hypotheses sustaining all of the previous lemmas hold. Let also  $([0, 1] \ni) x_h \triangleq \frac{\omega_k(h/2)^k + \theta_S(h/2)}{\mu(\mathcal{S})}$ , where  $\omega_k$  and  $\theta_S$  are given as in the proof of Lemma 1. Then*

$$(15) \quad \mathbb{P}(\mathcal{S} \not\subseteq \Omega_{\mathcal{P}_n}^h) \leq \frac{e^{-nx_h}}{x_h}.$$

*Proof.* Consider a finite  $\frac{h}{2}$ -net covering  $\mathcal{S}$  given by Lemma 3, that is,  $\mathcal{S} = \bigcup_{i=1}^{\mathcal{N}(\mathcal{S}, \frac{h}{2})} B_S(q_i, \frac{h}{2})$ ; then

$$\begin{aligned} \mathbb{P}(\mathcal{S} \not\subseteq \Omega_{\mathcal{P}_n}^h) &= \mathbb{P}\left(\bigcup_{x \in \mathcal{S}} \{x \notin \Omega_{\mathcal{P}_n}^h\}\right) \\ &= \mathbb{P}\left(\bigcup_{i=1}^{\mathcal{N}(\mathcal{S}, \frac{h}{2})} \bigcup_{x \in B_S(q_i, \frac{h}{2})} \{x \notin \Omega_{\mathcal{P}_n}^h\}\right) \\ &\leq \mathcal{N}\left(\mathcal{S}, \frac{h}{2}\right) \max_{1 \leq i \leq \mathcal{N}(\mathcal{S}, \frac{h}{2})} \mathbb{P}\left(\bigcup_{x \in B_S(q_i, \frac{h}{2})} \{x \notin \Omega_{\mathcal{P}_n}^h\}\right) \\ &= \mathcal{N}\left(\mathcal{S}, \frac{h}{2}\right) \max_{1 \leq i \leq \mathcal{N}(\mathcal{S}, \frac{h}{2})} \mathbb{P}\left(B_S\left(q_i, \frac{h}{2}\right) \not\subseteq \Omega_{\mathcal{P}_n}^h\right) \\ &= \mathcal{N}\left(\mathcal{S}, \frac{h}{2}\right) \left(1 - \min_{1 \leq i \leq \mathcal{N}(\mathcal{S}, \frac{h}{2})} \mathbb{P}\left(B_S\left(q_i, \frac{h}{2}\right) \subseteq \Omega_{\mathcal{P}_n}^h\right)\right). \end{aligned}$$

Using the lemmas above, we obtain

$$\mathbb{P}(\mathcal{S} \not\subseteq \Omega_{\mathcal{P}_n}^h) \leq \frac{(1 - x_h)^n}{x_h},$$

and we conclude by using the inequality  $1 - x \leq e^{-x}$ , valid for  $x \geq 0$ .  $\square$

It is both interesting and useful to find a relation between  $n$  (the number of points in the cloud),  $h$  (the radii of the balls), and  $k$  (the dimension of the manifold) which guarantees  $\lim_{n \uparrow +\infty, h \downarrow 0} \mathbb{P}(\mathcal{S} \not\subseteq \Omega_{\mathcal{P}_n}^h) = 0$ . For this purpose we will use Proposition 2.

Note that  $h > 0$  will be small, and also, if we are attempting to approximate  $d_S$ ,  $h$  should tend to 0.<sup>19</sup>

*Remark 6.* Note that for  $\{a_m\}_{m \in \mathbb{N}}$ ,  $a_m \downarrow 0$ ,  $\frac{e^{-ma_m}}{a_m}$  goes to zero as  $m \uparrow \infty$  if  $a_m$  is asymptotically greater than or equal to  $\frac{\log m}{m}$ . Then, in order to have the right-hand side of (15) tend to zero, we should have  $x_h \gtrsim \frac{\log n}{n}$ , and the condition relating  $h$ ,  $k$ , and  $n$  should then be<sup>20</sup>

$$(16) \quad h^k \gtrsim \left( \mu(\mathcal{S}) \frac{2^k}{\omega_k} \right) \frac{\log n}{n}.$$

Also, under this condition we can estimate the rate at which  $\frac{e^{-nx_h}}{x_h}$  approaches zero as  $n \uparrow \infty$ . For example, with  $x_h \simeq \frac{\log n}{n}$ ,  $\frac{e^{-nx_h}}{x_h} \simeq \frac{1}{\log n}$  as  $n \uparrow \infty$ . Note that, of course, we can speed up the convergence towards zero by choosing slower variations of  $x_{h_n}$  with  $n$ ; for instance, with  $x_{h_n} \simeq \frac{\log n^\gamma}{n}$ ,  $\gamma \geq 1$ , we have  $\frac{e^{-nx_h}}{x_h} \simeq \frac{1}{\gamma(\log n)n^{\gamma-1}}$  as  $n \uparrow \infty$ . Bounds for  $\mathbb{P}(\mathcal{S} \not\subseteq \Omega_{\mathcal{P}_n}^h)$  similar to ours can be found in [27]. It can be seen that our bounds are better than the ones reported in [27] for a certain range of  $k$ , the dimension of  $\mathcal{S}$ . We should point out that with our bounds we can obtain rates of convergence comparable to the optimal ones. Let us elaborate on this: In the case of the unit circle  $S^1$  it is known (see [66]) that

$$(17) \quad p_1(n, h) \triangleq 2ne^{-n\frac{h}{\pi}} \simeq \mathbb{P}(S^1 \not\subseteq \Omega_{\mathcal{P}_n}^h)$$

for  $n$  large and  $\frac{h}{\pi} \ll 1$ , whereas our bound is  $p_2(n, h) \triangleq \frac{e^{-nh/2\pi}}{h/2\pi} \gtrsim \mathbb{P}(S^1 \not\subseteq \Omega_{\mathcal{P}_n}^h)$ . Choose for  $p_1$ ,  $h_n^{(1)} = \gamma_1 \pi \frac{\log n}{n}$  and for  $p_2$ ,  $h_n^{(2)} = \gamma_2 \pi \frac{\log n}{n}$ . Plugging these expressions into the formulas for  $p_1$  and  $p_2$ , we find  $p_1 = 2n^{1-\gamma_1}$  and  $p_2 = \frac{2}{\gamma_2(\log n)} n^{1-\frac{\gamma_2}{2}}$ . Hence, by letting  $(2 >) \gamma_2 = 2\gamma_1$  (which is equivalent to  $\frac{h_n^{(2)}}{h_n^{(1)}} = 2$ ), we obtain  $p_2 \lesssim p_1$ . The *optimal* bound (17) for the case of  $S^1$  is derived using direct knowledge of the distribution of the minimal number of random arcs (of a certain fixed size) needed to cover  $S^1$  completely. This distribution is unknown for all nontrivial cases [66, 34]. In the case of the sphere  $S^2$ , also in [66], a bound of the type  $\mathbb{P}(S^2 \not\subseteq \Omega_{\mathcal{P}_n}^h \leq CN^2 e^{-DNh^2})$  is reported (for certain constants  $C$  and  $D$ ); however, the proof seems to use properties of symmetry of the sphere in a fundamental way. Other interesting bounds which could be used in this situation are those in [38].

We should finally point out that the problem of covering a certain domain (usually  $S^1$ ) with balls centered at random points sampled from this domain has been studied by many authors [66, 27, 28, 37, 65, 39, 34] and even by Shannon in [64].

We have the following interesting corollary, whose proof can be found in Appendix C.

**COROLLARY 4.** *Let  $\mathcal{S}$  be a smooth compact submanifold of  $\mathbb{R}^d$  without boundary. We have that if (16) holds, then for any  $\varepsilon > 0$*

$$\lim_{h,n} \mathbb{P}(d_{\mathcal{H}}(\mathcal{S}, \Omega_{\mathcal{P}_n}^h) > \varepsilon) = 0,$$

where  $d_{\mathcal{H}}$  is the Hausdorff distance between sets.

<sup>19</sup>For constant  $h > 0$ , by definition  $0 < x_h < 1$ , and then obviously  $\frac{e^{-nx_h}}{x_h} \rightarrow 0$  as  $n \uparrow \infty$ .

<sup>20</sup>This kind of condition is commonplace in the literature of random coverings; see, e.g., [25, 65, 22].

We are now ready to state and prove the following convergence theorem.

**THEOREM 7.** *Let  $\mathcal{S}$  be a  $k$ -dimensional smooth compact submanifold of  $\mathbb{R}^d$ . Let  $\mathcal{P}_n = \{p_1, \dots, p_n\} \subseteq \mathcal{S}$  be such that  $p_i \sim \mathbf{U}[\mathcal{S}]$  for  $1 \leq i \leq n$ . Then if  $h = h_n$  is such that  $h_n \downarrow 0$  and (16) holds as  $n \uparrow \infty$ , we have that for any  $\varepsilon > 0$ ,*

$$\mathbb{P} \left( \max_{p,q \in \mathcal{S}} \left( d_{\mathcal{S}}(p,q) - d_{\Omega_{\mathcal{P}_n}^h}(p,q) \right) > \varepsilon \right) \xrightarrow{n \uparrow \infty} 0.$$

*Proof.* We base our proof on (7). We first note that  $\mathbb{P}(\mathcal{E}_\varepsilon | \mathcal{J}_{h,n}) = 0$  for  $n$  large enough because, from considerations at the beginning of section 4,  $\max_{p,q \in \mathcal{S}} (d_{\mathcal{S}}(p,q) - d_{\Omega_{\mathcal{P}_n}^h}(p,q)) \leq C_S \sqrt{h_n}$  whenever  $\mathcal{S} \subseteq \Omega_{\mathcal{P}_n}^h$  holds. Let  $N = N(\varepsilon) \in \mathbb{N}$  be such that  $h_n < (\frac{\varepsilon}{C_S})^2$  for all  $n \geq N(\varepsilon)$ . Then, for  $n \geq N(\varepsilon)$ ,  $\mathbb{P}(\mathcal{E}_\varepsilon) = \mathbb{P}(\mathcal{J}_{h,n}^c \leq \frac{e^{-n x_{h_n}}}{x_{h_n}})$ , and since by assumption (16) holds, the right-hand side goes to 0 as  $n \uparrow \infty$ .  $\square$

*Remark 7.*

1. As can be gathered from the preceding proof, for fixed  $\varepsilon > 0$  and large  $n \in \mathbb{N}$ ,

$$\mathbb{P} \left( \max_{p,q \in \mathcal{S}} \left( d_{\mathcal{S}}(p,q) - d_{\Omega_{\mathcal{P}_n}^h}(p,q) \right) > \varepsilon \right)$$

can be upper bounded by  $\frac{e^{-n x_{h_n}}}{x_{h_n}}$ . For example, setting  $x_{h_n} = \gamma \frac{\log n}{n}$  for  $\gamma \geq 1$  yields (given  $n$  big enough)

$$\mathbb{P} \left( \max_{p,q \in \mathcal{S}} \left( d_{\mathcal{S}}(p,q) - d_{\Omega_{\mathcal{P}_n}^h}(p,q) \right) > \varepsilon \right) \leq \frac{1}{\gamma n^{\gamma-1} \log n}.$$

2. Then we see that by requiring  $\sum_{n \geq 1} \frac{e^{-n x_{h_n}}}{x_{h_n}} < \infty$  and using the Borel-Cantelli lemma, we obtain *almost sure convergence*, namely,

$$\mathbb{P} \left( \lim_{n \uparrow \infty} \max_{p,q \in \mathcal{S}} \left( d_{\mathcal{S}}(p,q) - d_{\Omega_{\mathcal{P}_n}^h}(p,q) \right) = 0 \right) = 1.$$

This can be guaranteed (for example) by setting  $x_{h_n} = \gamma \frac{\log n}{n}$  for  $\gamma > 2$ .

This concludes our study of distance functions on (noiseless) point clouds (sampled manifolds). We now turn to the even more realistic scenario where the points are considered to be *noisy* samples.

**6. Noisy sampling of manifolds.** We assume that we have some uncertainty on the actual position of the surface, and we model this as if each point in the set of sampled points is modified by a (not yet random) perturbation of magnitude smaller than  $\Delta$ . More explicitly, each  $p_i$  is given as  $p_i = p + \zeta \times \vec{v}$  for some  $\vec{v} \in S^{d-1}$ , some  $p$  in  $\mathcal{S}$ , and  $\Delta \geq \zeta \geq 0$ . Then we can guarantee that the point  $p$  from which  $p_i$  comes can be found inside  $B(p_i, \Delta) \cap \mathcal{S}$ . We are again interested in comparing  $d_{\Omega_{\mathcal{P}_n}^h} : \Omega_{\mathcal{P}_n}^h \rightarrow \mathbb{R}^+ \cup \{0\}$  with  $d_{\mathcal{S}} : \mathcal{S} \rightarrow \mathbb{R}^+ \cup \{0\}$ , but now these functions have different domains; therefore we must be careful in defining a meaningful way of relating them. If we consider

$$\mathcal{F}_{\mathcal{S}}^\Delta \triangleq \{f \mid f : \Omega_{\mathcal{S}}^\Delta \rightarrow \mathcal{S}, f(p) \in B(p, \Delta) \cap \mathcal{S}\},$$

we can compare, for some  $f \in \mathcal{F}_{\mathcal{S}}^\Delta$  and  $1 \leq i, j \leq n$ ,  $d_{\Omega_{\mathcal{P}_n}^h}(p_i, p_j)$  with  $d_{\mathcal{S}}(f(p_i), f(p_j))$ .

Note that as the perturbation's magnitude goes to zero,  $\mathcal{F}_{\mathcal{S}}^\Delta \ni f(p) \xrightarrow{\Delta \downarrow 0} p$ , for  $p \in \Omega_{\mathcal{S}}^\Delta$ . The next step is to write  $\max_{1 \leq i, j \leq n} \|d_{\Omega_{\mathcal{P}_n}^h}(p_i, p_j) - d_{\mathcal{S}}(f(p_i), f(p_j))\|$ , the biggest

error we have for our set of points. And finally, the next logical step is to look at the worst possible choice for  $f$ :

$$(18) \quad \mathcal{L}_S(\mathcal{P}_n; \Delta, h) \triangleq \sup_{f \in \mathcal{F}_S^\Delta} \max_{1 \leq i, j \leq n} \left| d_S(f(p_i), f(p_j)) - d_{\Omega_{\mathcal{P}_n}^h}(p_i, p_j) \right|.$$

We start by presenting deterministic bounds for the expression in (18), and only later will we be more (randomly) greedy and, in the spirit of Theorem 7, prove for  $\varepsilon > 0$  a result of the form  $(\mathcal{L}_S(\mathcal{P}_n; \Delta, h) > \varepsilon)$  will be a R.V.)

$$\mathbb{P}(\mathcal{L}_S(\mathcal{P}_n; \Delta, h) > \varepsilon) \xrightarrow{n \uparrow \infty} 0.$$

**6.1. Deterministic setting.** The idea is to prove that for some convenient function  $\widehat{f} \in \mathcal{F}_S^\Delta$  we can write

$$\mathcal{L}_S(\mathcal{P}_n; \Delta, h) \leq \max_{1 \leq i, j \leq n} \left| d_S(\widehat{f}(p_i), \widehat{f}(p_j)) - d_{\Omega_{\mathcal{P}_n}^h}(p_i, p_j) \right| + \lambda(h, \Delta),$$

where  $0 \leq \lambda(x, y) \xrightarrow{x, y \downarrow 0} 0$ . The natural candidate for  $\widehat{f}$  is the orthogonal projection onto  $\mathcal{S}$ ,  $\Pi_S : \Omega_S^H \rightarrow \mathcal{S}$ , whose properties are discussed in Appendix B. Then we see that we can reduce everything to bounding  $\max_{p, q \in \mathcal{S}} \|d_S(p, q) - d_{\Omega_{\mathcal{P}_n}^h}(p, q)\|$ . This is simple since if  $\mathcal{P}_n \subset \Omega_S^\Delta$ , then  $\Omega_{\mathcal{P}_n}^h \subset \Omega_S^{h+\Delta}$ , and  $d_S \geq d_{\Omega_{\mathcal{P}_n}^h|_{\mathcal{S}}} \geq d_{\Omega_S^{h+\Delta}|_{\mathcal{S}}}$ , and finally from Theorem 5,  $\|d_S - d_{\Omega_{\mathcal{P}_n}^h}\|_{L^\infty(\mathcal{S})} \leq C_S \sqrt{h + \Delta}$ .

Let  $\mathcal{S} \subset \Omega_{\mathcal{P}_n}^h$ ,  $f \in \mathcal{F}_S^\Delta$ , and  $1 \leq i, j \leq n$ . Then, after using the triangle inequality a number of times, we can write the bound

$$\begin{aligned} \left| d_S(f(p_i), f(p_j)) - d_{\Omega_{\mathcal{P}_n}^h}(p_i, p_j) \right| &\leq 2 \sup_{f \in \mathcal{F}_S^\Delta} \max_{p \in \mathcal{P}_n} d_S(f(p), \Pi_S(p)) \\ &\quad + \max_{p, q \in \mathcal{S}} \left| d_S(p, q) - d_{\Omega_{\mathcal{P}_n}^h}(p, q) \right| \\ &\quad + \max_{p, q \in \mathcal{P}_n} \left| d_{\Omega_{\mathcal{P}_n}^h}(p, q) - d_{\Omega_{\mathcal{P}_n}^h}(\Pi_S(p), \Pi_S(q)) \right|. \end{aligned}$$

The last term can be bounded by  $2\Delta$ , the one in the middle has already been discussed, and hence we are left with the first one. Using Corollary 2, we find that since  $f(p) \in B(\Pi_S(p), 2\Delta) \cap \mathcal{S}$ , then in fact  $f(p) \in B_S(\Pi_S(p), 2\Delta(1 + C_S\sqrt{\Delta}))$  and  $d_S(f(p), \Pi_S(p)) \leq 2\Delta(1 + C_S\sqrt{2}\sqrt{\Delta})$ . Summing up, under the condition  $\mathcal{S} \subset \Omega_{\mathcal{P}_n}^h$ , we obtain the desired result,

$$(19) \quad \mathcal{L}_S(\mathcal{P}_n; \Delta, h) \leq C_S \sqrt{h + \Delta} + 2\Delta(2 + \sqrt{2}C_S\sqrt{\Delta}).$$

**6.2. Random setting.** Assume that  $\{p_1, \dots, p_n\}$  is a set of i.i.d. random points such that each  $p_i \sim \mathbf{U}[\Omega_S^\Delta]$ . At this time, we want to estimate the probability of having  $\mathcal{S} \subseteq \Omega_{\mathcal{P}_n}^h$ . It is easy to see that as a first “reality compliant” condition one should have that the noise level not be too big with respect to  $h$ . We will impose  $h \geq \Delta$  for simplicity’s sake, as can be understood from the convergence theorem below. Since the techniques are similar to those used in the noise-free case, we will present its proof in Appendix C.

**THEOREM 8.** *Let  $\mathcal{S}$  be a  $k$ -dimensional smooth compact submanifold of  $\mathbb{R}^d$ . Let  $\mathcal{P}_n = \{p_1, \dots, p_n\}$  be such that  $p_i \sim \mathbf{U}[\Omega_S^\Delta]$  for  $1 \leq i \leq n$ . Then if  $h = h_n$ ,  $\Delta = \Delta_n$*

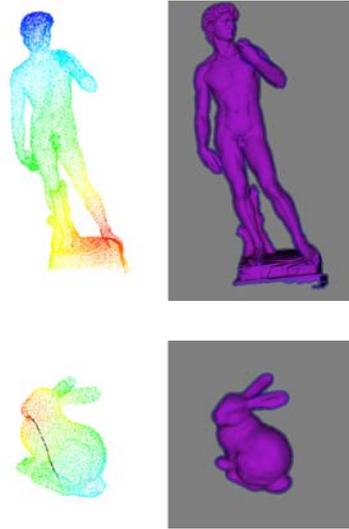


FIG. 1. *Intrinsic distance function for a point cloud. A point is selected in the head of the David, and the intrinsic distance is computed following the framework introduced here. The point cloud is colored according to the intrinsic distance to the selected point, going from bright red (far) to dark blue (close). The offset band, given by the union of balls, is shown next to the distance figure. Bottom: Same as before, with a geodesic curve between two selected points.*

are such that  $\Delta_n \leq h_n$  and  $h_n \downarrow 0$  and  $\Delta_n^k \gtrsim \frac{\log n}{n}$  as  $n \uparrow \infty$ , we have that for any  $\varepsilon > 0$ ,

$$\mathbb{P}(\mathcal{L}_S(\mathcal{P}_n; \Delta, h) > \varepsilon) \xrightarrow{n \uparrow \infty} 0.$$

We have now concluded the analysis of the most general case for noisy sampling of manifolds. Note that, although the results in this and in previous sections were presented for Euclidean balls, they can easily be extended to more general covering shapes (check Corollary 1 above), e.g., following [15, 36], or using minimal spanning trees, or from the local directions of the data [56]. In addition, the recently developed approach reported in [57] can be used for defining the offset band in an adaptive fashion. This will improve the bounds reported here. Similarly, the results can be extended to other sampling or noise models following the same techniques developed here.

**7. Implementation details and examples.** We now present examples of distance matrices and geodesics for point clouds (Figure 1), use these computations to find intrinsic Voronoi diagrams (Figure 2; see also [42, 43, 71]); and compare the results with those obtained with mesh-based techniques (Figure 3).<sup>21</sup> We also present examples in high dimensions and use, following and extending [24], our results to compare manifolds given by point clouds. All these exercises are to exemplify the importance of computing distance functions and geodesics on point clouds, and are by no means exhaustive. The 3D data sets used come from real point cloud data and have been obtained either from range scanners (*David* model) or via photometric stereo techniques (man and woman).

<sup>21</sup>All the figures in this paper are in color. VRML files corresponding to these examples can be found at <http://mountains.ece.umn.edu/~guille/pc.htm>.



FIG. 2. Voronoi diagram for point clouds. Four points (left) and two points (right) are selected on the cloud, and the point cloud is divided (colored) according to the geodesic distance to these four points. Note that this is a surface Voronoi, based on geodesics computed with our proposed framework, not a Euclidean one.

The theoretical results presented in the previous sections show that the intrinsic distance and geodesics can be approximated by the Euclidean ones computed in the band defined (for example) by the union of balls centered at the points of the cloud. The problem is then simplified to first computing this band (no need for mesh computation, of course), and then using well-known computationally optimal techniques to compute the distances and geodesics inside this band, exactly as done in [49] for implicit surfaces (where the interested reader can also find explicit computational timings and accuracy comparisons with mesh-based approaches). The band itself can be computed in several ways, and for the examples below we have used constant radii. Locally adaptive radii can be used, based, for example, on diameters obtained from minimal spanning trees or on the recent work reported in [57]. Automatic and local estimation of  $h$  defining  $\Omega_{\mathcal{P}_n}^h$ , which will improve the bounds reported here, was not pursued in this paper and is the subject of current implementation efforts.

The software implementation of the algorithm is based on using the fast Euclidean distance computation algorithms, usually referred to as *fast marching* algorithms [35, 62, 63, 69], twice. We omit the description of this algorithm since it is well known. The starting point is defining a grid over which all the computations are performed. This amounts to choosing  $\Delta_{x_i}$ , the grid spacing in each direction  $i = 1, \dots, d$ , which will determine the accuracy of the numerical implementation (the offset band includes fewer than 10 grid points).<sup>22</sup> In the first round we compute the band  $\Omega_{\mathcal{P}_n}^h = \{x \in \mathbb{R}^d : d(\mathcal{P}_n, x) \leq h\}$  by specifying a value of zero for the function  $\Psi(x) = d(\mathcal{P}_n, x)$  on the points  $x \in \mathcal{P}_n$ . Since in general these points will not be on the grid, we use a simple multilinear interpolation procedure to specify the values on neighboring grid points. The second use of the fast distance algorithm is also simply reduced to using  $\Psi$  to define  $\Omega_{\mathcal{P}_n}^h$  by using the simple modification reported in [49]. The computation of geodesics was done using a simple Runge–Kutta gradient descent procedure, much in the way described in [49], with some obvious modifications.

<sup>22</sup>Adaptive grids inside the fixed or variable width offset band could be used as well; see, for example, [30].

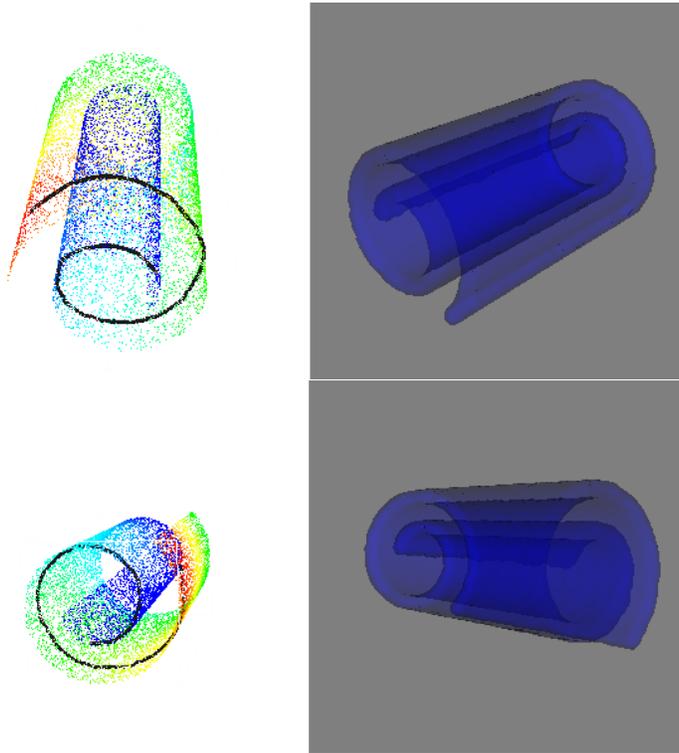


FIG. 3. *Examples of geodesic computations. This data is used to study the algorithm robustness to noise, see Appendix A.*

All the code and 3D visualization was developed in C++ using both Flujos (which is written using Blitz++; see [7]) and VTK (see [70]). For matrix manipulation and visualization of other results we used MATLAB. We are currently working on a more advanced implementation of the proposed framework that permits us to work with high-dimensional data without having the memory allocation problems that result from blind and straightforward allocation of resources to empty and nonused grids.

**7.1. High-dimensional data.** In this section we present a simple example for high-dimensional data. We embed a circle of radius 15 in  $\mathbb{R}^5$ , and use a grid of size  $34 \times 4 \times 4 \times 4 \times 34$  (with uniform spacing  $\Delta x = 1$ ) such that each of the sample points is of the form  $p_i = 15(\cos(\frac{2\pi i}{N}), 0, 0, 0, \sin(\frac{2\pi i}{N})) + (17, 2, 2, 2, 17)$ , for  $1 \leq i \leq N$ . We then use our approach to compute the (approximate) distance function  $d_h$  in a band in  $\mathbb{R}^5$ , and then the error  $e_{ij} = |d_s(p_i, p_j) - d_h(p_i, p_j)|$  for  $i, j \in \{1, \dots, N\}$ . In our experiments we used  $h = 2.5 > \Delta x\sqrt{5}$ .<sup>23</sup> We randomly sampled 500 points from the  $N = 1000$  points used to construct the union of balls to build the  $500 \times 500$  error matrix  $((e_{ij}))$ . We found  $\max_{i,j} \{e_{ij}\} = 2.0275$ , that is, a 4.3%  $L_\infty$ -error. In Figure 4 we show the histogram of all the  $(500^2)$  entries of  $((e_{ij}))$ . We should also note that when following the dimensionality reduction approach in [67], with the geodesic distance computation proposed here, the correct dimensionality of the circle was obtained.

<sup>23</sup>For a discussion on how to make a preliminary estimation of the value of  $h$ , see [49].

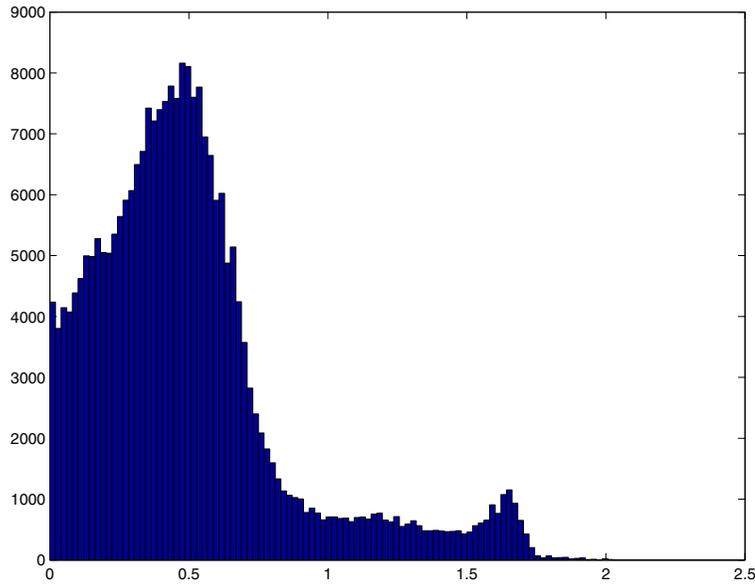


FIG. 4. Histogram for the error in the case of a circle embedded in  $\mathbb{R}^5$ .

In high dimensions, when the grid is too large, our current numerical implementation becomes unusable. The problem stems from the fact that we require too much memory space, most of which is not really used, since the computations are conducted only in a band around  $\mathcal{P} \subset \mathbb{R}^d$ . To be more precise, the memory requirements of our current direct implementation, which uses a  $d$ -dimensional array to make the computations, are  $\simeq (\max_i l_i)^d$ , whereas we really need a storage capacity of order  $\mu_k(\mathcal{S})h^{d-k}$ , where  $l_i$  is the size of  $\mathcal{P}$ 's bounding box along the  $i$ th direction,  $1 \leq i \leq d$ , and  $\mu_k(\mathcal{S})$  is the measure of the  $k$ -dimensional manifold  $\mathcal{S}$  (embedded in  $\mathbb{R}^d$ ). This memory problem is to be addressed by a computation that is not based on discretizing the whole band. (Note, of course, that the theoretical foundations presented in this paper are independent of the particular implementation.) We are currently working on addressing this specific issue.

**7.2. Object recognition.** The goal of this application is to use our framework to compare manifolds given by point clouds. The comparison is done in an intrinsic way, that is, isometrically (bending) invariant. This application is motivated by [24], where they use geodesic distances (computed using a graph-based approach) to compare 3D triangulated surfaces. In contrast with [24], we compare point clouds using our framework (which is not only based in the original raw data, but also, as shown in Appendix A, more robust to noise than mesh approaches such as those of [24] and is valid in any dimensions), and use a different procedure/similarity metric between the manifolds. The authors in [24] basically project into low-dimensional manifolds and use eigenvalues and eigenvectors of a centralized matrix related to the *distance matrices* (matrices which in each entry  $(i, j)$  have the value of the intrinsic distance between (projected) points  $p_i$  and  $p_j$  of the cloud), which are clearly not sufficient to distinguish nonisometric objects. (Nonisometric objects can have distance matrices with the same eigenvalues.) A different study, based on direct comparisons of distance matrices, is used here and detailed in Appendix E.

TABLE 1  
*Information about the models used in our recognition experiments.*

Dataset	Number of points in the cloud ( $n$ )	Grid size used
Bunny	15862	$80 \times 80 \times 70$
MAN2	26186	$120 \times 90 \times 200$
MAN3	26186	$120 \times 90 \times 165$
MAN5	26186	$120 \times 85 \times 160$
WOMAN2	29624	$120 \times 105 \times 175$
WOMAN3	29624	$120 \times 100 \times 180$

Our task then is to compare two manifolds in an intrinsic way; i.e., we want to check whether they are isometric or not. We want to check this condition by using point clouds representing each one of the manifolds. Let  $\mathcal{S}_1$  and  $\mathcal{S}_2$  be two submanifolds of  $\mathbb{R}^d$  and sample on each of them the two point clouds  $\mathcal{P}_n^{(1)} \subset \mathcal{S}_1$  and  $\mathcal{P}_n^{(2)} \subset \mathcal{S}_2$ . Then, following our theory, we compute the corresponding distances in the offset bands for these two sets of points,  $d_{\Omega_{\mathcal{P}_n^{(1)}}^{h_1}}$  and  $d_{\Omega_{\mathcal{P}_n^{(2)}}^{h_2}}$ , and for point subsets  $\{q_1^{(1)}, \dots, q_m^{(1)}\} = \mathcal{Q}_m^{(1)} \subseteq \mathcal{P}_n^{(1)}$ ,  $\{q_1^{(2)}, \dots, q_m^{(2)}\} = \mathcal{Q}_m^{(2)} \subseteq \mathcal{P}_n^{(2)}$  we compute the corresponding  $m \times m$  pairwise distance matrices (as defined above)

$$D_1 = \left( \left( d_{\Omega_{\mathcal{P}_n^{(1)}}^{h_1}}(q_i^{(1)}, q_j^{(1)}) \right) \right) \quad \text{and} \quad D_2 = \left( \left( d_{\Omega_{\mathcal{P}_n^{(2)}}^{h_2}}(q_i^{(2)}, q_j^{(2)}) \right) \right).$$

Let  $\mathcal{PM}_m$  be the set of  $m \times m$  permutation matrices and  $\|\cdot\|$  a unitary transformation invariant norm<sup>24</sup> (fix the Frobenius norm:  $\|A\| = \sqrt{\sum_i \sum_j a_{ij}^2}$ ). Then we define the  $\mathcal{J}$ -distance between (distance) matrices  $D_1$  and  $D_2$  as

$$d_{\mathcal{J}}(D_1, D_2) \triangleq \min_{P \in \mathcal{PM}_m} \|D_1 - PD_2P^T\|.$$

Clearly, if  $d_{\mathcal{J}}(D_1, D_2) = 0$ , then we have an isometry between the discrete metric sets  $(\mathcal{Q}_n^{(1)}, d_{\Omega_{\mathcal{P}_n^{(1)}}^{h_1}})$  and  $(\mathcal{Q}_n^{(2)}, d_{\Omega_{\mathcal{P}_n^{(2)}}^{h_2}})$ . This should allow us to establish a *rough isometry* (see [14, section 4.4]) between  $\mathcal{S}_1$  and  $\mathcal{S}_2$  with interesting constants.

The exact details on how this metric is approximated and how the subsets of points  $\mathcal{Q}$  are selected is presented in Appendix E. For the experiments regarding recognition of shapes we used the datasets listed in Table 1.

In Figure 5 we present the histogram of the error  $e(100)/100$  for 20 different  $100 \times 100$  distance matrices corresponding to the full Bunny model, with the 100 points chosen as in the “packing procedure” described in Appendix E, where the exact definition of  $e(\cdot)$  is also given (see (28)). We computed the mean of  $e(100)/100$  over the  $19 \times 18 \times \dots \times 1 = 190$  comparison experiments to be 0.4774 with standard deviation 0.0189. This can be interpreted as indicating that when one considers a large enough set of points, the information contained in the packing set is representative of the metric information of the manifold, independently of the particular choice of the packing set. This claim needs some further theoretical justification, which could come if a result of the following fashion were proved:<sup>25</sup>

<sup>24</sup> $\|AU\| = \|A\|$  for any matrix  $A$  and any unitary matrix  $U$ .

<sup>25</sup>Note added in proof: After this paper was submitted for publication, we proved that a properly modified version of the above claim holds *in probability*; see [50] for details.

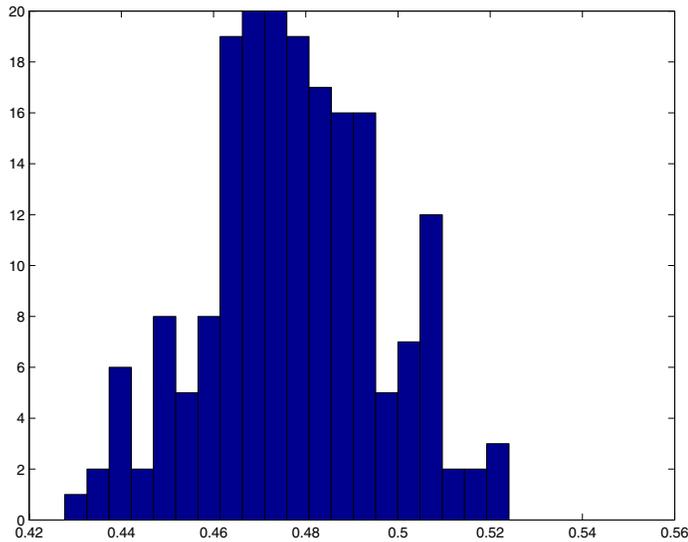


FIG. 5. Histogram showing the errors for different selections of point clouds on the bunny model.

Let  $\mathcal{S}$  be a smooth compact  $k$ -dimensional submanifold of  $\mathbb{R}^k$  such that its Ricci curvature is bounded below by  $\kappa(n-1)$  with  $\kappa \leq 0$ . Let  $\mathcal{Q}_m^{(r)} \subset \mathcal{S}$ ,  $r = 1, 2$ , be such that  $d_{\mathcal{S}}(q_i^{(r)}, q_j^{(r)}) \geq \varepsilon$  and  $B_{\mathcal{S}}(\mathcal{Q}_m^{(r)}, R)$  covers  $\mathcal{S}$  for some  $R > \varepsilon > 0$ . Then, with  $D_1$  and  $D_2$  defined as before,

$$d_{\mathcal{J}}(D_1, D_2) \leq 2mC_{\mathcal{S}}\sqrt{h} + C(R, \varepsilon, m),$$

where the exact form of  $C(R, \varepsilon, m)$  is to be determined, leading to an optimal choice of  $m$  (the size of the subset).

Using the same procedure, described in Appendix E, to choose the sets  $\mathcal{Q}_m^{(i)}$ , we computed the errors (according to  $e(D_1, D_2)$ ) for five artificial human models; three of them are bendings of a man and two are bendings of a woman; see Figures 6 and 7. Details on these models are also given in Table 1. The results of this cross-comparison are presented in Table 2 below.

TABLE 2

Cross-comparisons for the human models using the error measure  $e(300)/300$  normalized by the maximum of the errors.

MODEL	Man2	Man3	Man5	Woman2	Woman3
Man2	*	0.0514	0.0570	0.4690	0.4853
Man3	*	*	0.0206	0.4701	0.4859
Man5	*	*	*	0.4702	0.4862
Woman2	*	*	*	*	0.2639
Woman3	*	*	*	*	*

These examples show how our geodesic distance computation technique, when complemented with the matrix metric in Appendix E, can be used to compare manifolds given by point clouds, in a bending-invariant fashion and without explicit manifold reconstruction. More exhaustive experimentation and additional theoretical justification will be reported elsewhere.

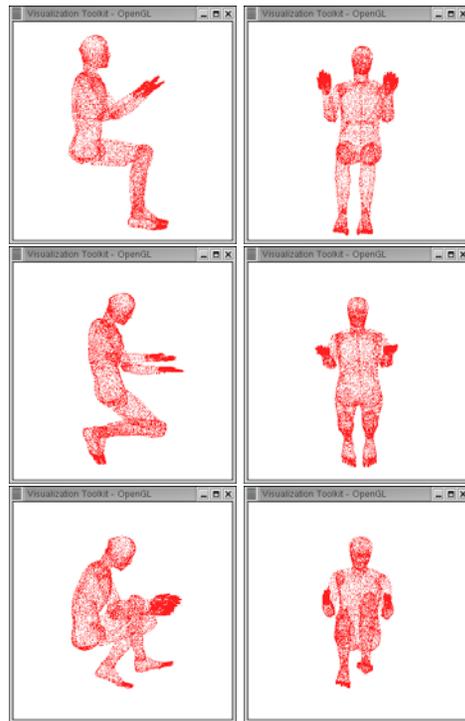


FIG. 6. *MAN* models. From top to bottom (two views of each model): *MAN2*, *MAN3*, and *MAN5*.

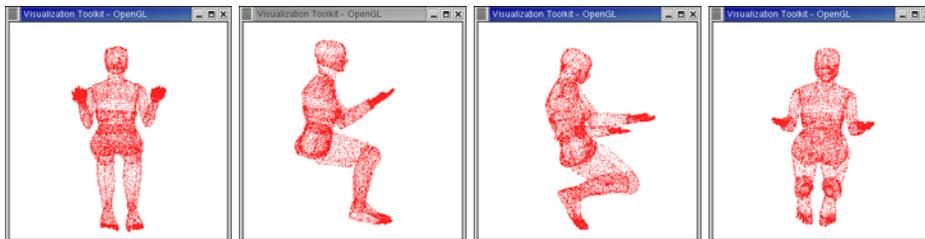


FIG. 7. *WOMAN* models. From left to right (two views of each model): *WOMAN2* and *WOMAN3*.

Before concluding, we should comment that, as frequently done in the literature, we could normalize the geodesic distances if scale invariance were also required. Moreover, we could also consider in the distance matrix only nonzero entries for local neighborhoods. In addition, the use of techniques for computing eigenvalues and eigenvectors such as those in the work of Coifman and colleagues [17], on high-dimensional geometric multiscale analysis should be explored.

**8. Concluding remarks.** In this paper, we have extended our previous work [49] to deal with (smooth) submanifolds of  $\mathbb{R}^d$  (of any codimension) and possibly with boundary, and using these extensions, we have also shown how to compute intrinsic distance functions on a generic manifold defined by a point cloud, without the intermediate step of manifold reconstruction. The basic idea is to use well-developed computational algorithms for computing Euclidean distances in an offset band sur-

rounding the manifold, to approximate the intrinsic distance. The underlying theoretical results were complemented by experimental illustrations.

As mentioned in the introduction, an alternative technique for computing geodesic distances was introduced in [9, 67] (see also [31]). In contrast with our work, the effects of noise were not addressed in [9, 31]. Moreover, as one can see from considerations in Appendix A, our framework seems to be more robust to noise. We should note that the memory requirements of the current way of implementing our framework are large, and this needs to be addressed for very high dimensions (the framework is, of course, still valid). In particular, we are interested in direct ways of computing distances inside regions defined by union of balls, without the need to use the Hamilton–Jacobi approach. Several classical computer science implementation tricks can be applied to avoid this memory allocation problem, and this is part of our current implementation efforts.

We are currently working on the use of this framework to create multiresolution representations of point clouds (in collaboration with C. Moenning and N. Dyn; see [55] and also [11, 18, 20, 58]), to further perform object recognition for larger libraries, and to compute basic geometric characteristics of the underlying manifold—all this, of course, without reconstructing the manifold. (See [54] for recent results on normal computations for 2D and 3D noisy point clouds.) Some results in these directions are reported in [50, 55]. Further applications of our framework for high-dimensional data are also currently being addressed, beyond the preliminary (toy) results reported in section 7. Of particular interest in this direction is the combination of this work with the one developed by Coifman and colleagues and the recent one in [31].

**Appendix A. Comparison with mesh-based strategies for distance calculation in the presence of noise.** We now make some very basic comparisons between our approach to geodesic distance computations and those based on graph approximations to the manifold, such as the one in Isomap [67, 31].<sup>26</sup> (Comparisons of the band framework with the one reported in [40] for 3D triangulated surfaces are reported in [49].) The goal is to show that such graph-based techniques are more sensitive to noise in the point cloud sample (and the error can even increase to infinity with the increase in the number of points). This is expected, since the geodesic in such techniques goes through the noisy samples, while in our approach, they just go through the union of balls. We make our argument only for the 1D case, while the high-dimensional cases can be similarly studied.

**A.1. 1D theoretical case.** Let us consider a rectilinear segment of length  $L$  and  $n + 1$  equispaced points  $p_1, \dots, p_{n+1}$  in that segment. Consider the *noisy* points  $q_i = p_i + \zeta_i \vec{n}$ , where  $\vec{n}$  is the normal to the segment and  $\zeta_i$   $1 \leq i \leq n$  are independent R.V. uniformly distributed in  $[-\Delta, \Delta]$ . Let  $l = L/n$  denote the distance between adjacent  $p_i$ 's. Let  $d_g^\Delta$  denote the length of the polygonal path  $\overline{q_1 q_2 \dots q_{n+1}}$  and  $d_0 = L$ . Then obviously  $d_g^\Delta \geq d_0$  for any realization of the R.V.'s  $\zeta_i$ . Let  $d_i = \|p_i - p_{i+1}\|$ ; then by Pythagoras theorem  $d_i = \sqrt{l^2 + z_i^2}$ , where  $z_i = \zeta_i - \zeta_{i+1}$  are R.V.'s with triangular density in  $[-2\Delta, 2\Delta]$ .

Next we compute  $\mathbb{E}(d_i) = \frac{1}{2\Delta} \int_{-2\Delta}^{2\Delta} \sqrt{l^2 + z^2} (1 - \frac{|z|}{2\Delta}) dz$ . The result is

$$\mathbb{E}(d_i) = \sqrt{l^2 + 4\Delta^2} + \frac{l^2}{2\Delta} \log \left( \frac{2\Delta + \sqrt{l^2 + 4\Delta^2}}{l} \right) - \frac{1}{6\Delta^2} \left( (l^2 + 4\Delta^2)^{3/2} - l^3 \right).$$

<sup>26</sup>Isomap builds a mesh by locally connecting the (noisy) samples.

TABLE 3  
Results of simulations with the Swiss Roll dataset.

Noise power ( $n_k^2$ )	$\max_{ij}  D_{ij}^{g,n_k} - D_{ij}^{g,0} $	$k$	$\max_{ij}  D_{ij}^{h,n_k} - D_{ij}^{h,0} $	$h$
0.0001	2.5222	7	0.5266	1.8
0.01	4.6409	7	0.9430	1.8
0.04	5.1737	7	1.2489	1.8
0.09	5.3292	7	1.4682	1.8
0.16	5.4651	7	1.7965	1.8

Now assuming  $\frac{\Delta}{l} \ll 1$ , we find that up to first order  $\mathbb{E}(d_i) \simeq l + \Delta$  and

$$\mathbb{E}(d_g^\Delta - d_0) \simeq n\Delta.$$

From this we also get<sup>27</sup>

$$p_g \stackrel{\Delta}{=} \mathbb{P}(d_g^\Delta - d_0 > \varepsilon) \lesssim \frac{n\Delta}{\varepsilon}.$$

On the other hand, for our approximation  $d_h^\Delta$ , if the segment is contained in the union of the balls centered at the sampling points,  $d_h^\Delta = d_0$ . The probability of covering the segment by the band can be made arbitrarily close to 1 by increasing  $n$ . More precisely, one can prove that if  $p$  stands for the value of the probability of *not covering* the segment, then  $p \leq k \frac{L}{\Delta} (1 - k' \frac{\Delta}{L})^n$ , for some positive constants  $k$  and  $k'$ . Then we can write

$$p_h \stackrel{\Delta}{=} \mathbb{P}(d_h^\Delta - d_0 > \varepsilon) \leq \frac{k''}{\varepsilon} \frac{L}{\Delta} \left(1 - k' \frac{\Delta}{L}\right)^{n+1}.$$

The comparison is now easy. We see that in order to have  $p_g$  vanish as  $n \uparrow \infty$ ,  $\Delta$  must go to zero *faster* than  $\frac{1}{n}$ . However, we know that by requiring  $\Delta \simeq \frac{\log n}{n} \gtrsim \frac{1}{n}$  we have  $p_h \downarrow 0$  as  $n \uparrow \infty$ . This means that the graph approximation of the distance is more sensitive to noise than ours.<sup>28</sup> This gives some evidence about why our approach is more robust than popular mesh-based ones. Next we present results of some simulations carried out in order to further verify our claim.

**A.2. Simulations.** In Table 3 we present results of simulations carried out for the SwissRoll dataset [67]; see Figure 3. We used 10,000 points to define the manifold. We then generated 10,000 noise vectors, each component being uniform with power one and zero mean. Then we generated noisy datasets from the noiseless SwissRoll dataset by adding the noise vector times a constant  $n_k$  to each vector of the noiseless initial dataset. We then chose 1000 corresponding points in each dataset and computed the intrinsic pairwise distance approximation, obtaining the matrices  $\{(D_{ij}^{g,n_k})\}$  and  $\{(D_{ij}^{h,n_k})\}$  for the graph-based and our approach, respectively, where  $k = 1, 2, \dots, 5$ ,  $i, j \in [1, 1000]$ , and  $n_k$  denotes the noise level. We then computed the values of  $\max_{ij} |D_{ij}^{g,n_k} - D_{ij}^{g,0}|$  and  $\max_{ij} |D_{ij}^{h,n_k} - D_{ij}^{h,0}|$  for each  $k$ , where  $D_{ij}^{g,0}$  and  $D_{ij}^{h,0}$  stand for noiseless intrinsic distance approximations. In Table 3,  $h$  indicates the radii and  $k$  the size of the neighborhood for Isomap. The graph approximation shows

<sup>27</sup>Also, with similar arguments we can prove that  $\max_{\zeta_1, \dots, \zeta_{n+1}} (d_g^\Delta - d_0) \simeq \frac{2n^2\Delta^2}{L}$ .

<sup>28</sup>Another way of seeing this is by noting that, for a fixed noise level  $\Delta$ , by increasing  $n$  we actually worsen the graph approximation, whereas we are making our approximation better.

less robustness to noise than our method, as was argued above. This is also true for the sensitivity,<sup>29</sup> where our approach outperforms the graph-based one by at least one order of magnitude. Note that the sensitivity for our approach can be formally studied from Theorem 3.

**Appendix B. Properties of Euclidean distance functions.** The references for this section are [2, pp. 12–16], and [26].

**THEOREM 9** (see [2]). *Let  $\Gamma \subset \mathbb{R}^d$  be a compact, smooth manifold without boundary. Then  $\eta(x) \triangleq \frac{1}{2}d^2(\Gamma, x)$  is smooth in a tubular neighborhood  $U$  of  $\Gamma$ . Also, in  $U$  it satisfies  $\|D\eta\|^2 = 2\eta$ .*

**COROLLARY 5.** *The projection operator  $\Pi : U \rightarrow \Gamma$ , for a given  $x \in U$ , can be written as  $\Pi(x) = x - D\eta(x)$ . Moreover, this operator is smooth.*

**Remark 8.** Differentiation of the relation  $\langle D\eta, D\eta \rangle = 2\eta$  gives us  $D^2\eta D\eta = D\eta$ . Differentiating once more, we also find  $D^3\eta D\eta = D^2\eta$ .

**THEOREM 10** (see [2]). *Let  $\Gamma$  and  $U$  be as in Theorem 9, and let  $y \in U$  and  $x = y - D\eta(x) \in \Gamma$ ,  $k = \dim(\Gamma)$ . Then, denoting by  $\lambda_1, \dots, \lambda_n$  the eigenvalues of  $D^2\eta(y)$ ,*

$$\lambda_i(y) = \begin{cases} \frac{d(\Gamma, y)\kappa_i(x)}{1+d(\Gamma, y)\kappa_i(x)} & \text{if } 1 \leq i \leq k, \\ 1 & \text{if } k < i \leq n, \end{cases}$$

where  $\kappa_i(x)$  are the principal curvatures of  $\Gamma$  at  $x$  along  $Dd(\Gamma, y) \in N_x\Gamma$ , where  $N_x\Gamma$  is the normal space to  $\Gamma$  at  $x$ .

**Appendix C. Deferred proofs.**

*Proof of Corollary 3.* We present only a sketch of the proof. Let  $\mathcal{M}$  be an extension of  $\mathcal{S}$  such that  $\mathcal{S}$  is still strongly convex in  $\mathcal{M}$ , and let  $0 < \delta \triangleq \min_{x \in \mathcal{S}} \min_{z \in \mathcal{M}} \|x - z\|$ . Then,  $\overline{B(x, \alpha)} \cap \overline{B(z, \beta)} = \emptyset$  for all  $x \in \mathcal{S}$ ,  $z \in \partial\mathcal{M}$ , and  $\alpha, \beta < \frac{\delta}{3}$ . Hence,  $\Omega_{\mathcal{S}}^\alpha \cap \Omega_{\partial\mathcal{M}}^\beta = \emptyset$  for  $\alpha, \beta \leq \frac{\delta}{3}$ .

For any  $x, y \in \mathcal{S}$  consider  $\gamma_h$  the  $\Omega_{\mathcal{M}}^h$ -minimizing geodesic,  $\mathbf{L}(\gamma_h) = d_{\Omega_{\mathcal{M}}^h}(x, y)$ .

By the convexity of  $\mathcal{S}$  there exists a unique  $\mathcal{M}$ -minimizing geodesic  $\gamma_0 \subset \mathcal{S}$  joining  $x, y$ , and then, by Theorem 4,  $\gamma_h$  uniformly converges to  $\gamma_0$ . In particular, for any  $\epsilon > 0$  there exists  $h_\epsilon > 0$  such that  $\gamma_h \subset \Omega_{\gamma_0}^\epsilon$  for all  $h < h_\epsilon$ . Choose  $\epsilon \leq \frac{\delta}{3}$ ; then  $\gamma_h \subset \Omega_{\gamma_0}^\epsilon \subset \Omega_{\mathcal{S}}^\epsilon$ . Furthermore, if  $h \leq \frac{\delta}{3}$ , then  $\Omega_{\gamma_0}^\epsilon \cap \Omega_{\mathcal{M}}^h = \emptyset$ , and therefore  $\gamma_h$  does not touch  $\partial\Omega_{\mathcal{M}}^h \cap \partial\Omega_{\partial\mathcal{M}}^h$ . Thus,  $\gamma_h$  is  $C^{1,1}$  for  $h \leq \frac{\delta}{3}$ . Note that with this choice of  $h$  we have  $\Omega_{\mathcal{S}}^h \cap \mathcal{M} \subset \text{int}(\mathcal{M})$ , and therefore we also have a smooth orthogonal projection operator  $\Pi : \Omega_{\mathcal{S}}^h \rightarrow \mathcal{M}$ .

Proceeding as in the first steps of the proof of Theorem 5, we have  $\mathbf{L}(\gamma_h) = d_{\Omega_{\mathcal{M}}^h}(x, y) \leq d_{\mathcal{M}}(x, y) \leq \mathbf{L}(\Pi(\gamma_h))$ , since  $\Pi(\gamma_h) \subset \mathcal{M}$  but may not be a minimizing path. Then, using the convexity of  $\mathcal{S}$  in  $\mathcal{M}$ ,  $d_{\mathcal{M}}(x, y) = d_{\mathcal{S}}(x, y)$ , and therefore  $0 \leq d_{\mathcal{S}}(x, y) - d_{\Omega_{\mathcal{M}}^h}(x, y) \leq |\mathbf{L}(\Pi(\gamma_h)) - \mathbf{L}(\gamma_h)|$ , which can be bounded by a constant times  $\sqrt{h}$  just mimicking the proof of Theorem 5. We conclude by noting that  $\Omega_{\mathcal{S}}^h \subset \Omega_{\mathcal{M}}^h$ , and hence  $d_{\mathcal{S}}(x, y) - d_{\Omega_{\mathcal{M}}^h}(x, y) \geq d_{\mathcal{S}}(x, y) - d_{\Omega_{\mathcal{S}}^h}(x, y)$ .  $\square$

*Proof of Lemma 3.* We now estimate the covering number  $\mathcal{N}(\mathcal{S}, \delta)$ . The idea is constructive, very simple, and of course standard. We consider the following procedure (adopted from [9]): Let  $q_1$  be any point in  $\mathcal{S}$ , and choose  $q_2 \in \mathcal{S} \setminus B_{\mathcal{S}}(q_1, \delta)$ . Then choose  $q_3 \in \mathcal{S} \setminus \{B_{\mathcal{S}}(q_1, \delta) \cup B_{\mathcal{S}}(q_2, \delta)\}$ . Iterate this procedure until it is no longer possible

<sup>29</sup>Sensitivity is defined as  $\left| 1 - \frac{\text{distance for noisy points}}{\text{distance for clean points}} \right|$ .

to choose any point  $q \in \mathcal{S} \setminus \{\cup_{k=1}^{\mathcal{N}(\mathcal{S}, \delta)} B_{\mathcal{S}}(q_k, \delta)\}$ ; in such a case  $\mathcal{S} = \cup_{k=1}^{\mathcal{N}(\mathcal{S}, \delta)} B_{\mathcal{S}}(q_k, \delta)$ . Note that  $B_{\mathcal{S}}(q_k, \frac{\delta}{2}) \cap B_{\mathcal{S}}(q_l, \frac{\delta}{2}) = \emptyset$  if  $k \neq l$ , and therefore we can bound  $\mathcal{N}(\mathcal{S}, \delta) \leq \frac{\mu(\mathcal{S})}{\min_{x \in \mathcal{S}} \mu(B_{\mathcal{S}}(x, \delta/2))}$ . Therefore, using the Bishop–Günther inequalities in the same manner as in Lemma 1, we find (14).  $\square$

*Proof of Corollary 4.* Note first that the random variable  $d_{\mathcal{H}}(\mathcal{S}, \Omega_{\mathcal{P}_n}^h)$  is bounded by  $\max\{\text{diam}(\mathcal{S}) + h, h\}$ . By definition of the Hausdorff distance,  $d_{\mathcal{H}}(\mathcal{S}, \Omega_{\mathcal{P}_n}^h) = \max(\sup_{x \in \mathcal{S}} d(x, \Omega_{\mathcal{P}_n}^h), \sup_{y \in \Omega_{\mathcal{P}_n}^h} d(y, \mathcal{S}))$ . Then,  $\sup_{x \in \mathcal{S}} d(x, \Omega_{\mathcal{P}_n}^h) \leq \text{diam}(\mathcal{S}) + h$  by the triangle inequality, and  $\sup_{y \in \Omega_{\mathcal{P}_n}^h} d(y, \mathcal{S}) \leq h$ , trivially.

Now, we can write  $\mathbb{E}(d_{\mathcal{H}}(\mathcal{S}, \Omega_{\mathcal{P}_n}^h)) = \mathbb{E}(\mathbb{E}(d_{\mathcal{H}}(\mathcal{S}, \Omega_{\mathcal{P}_n}^h) | \mathcal{K}_{[\mathcal{S} \subseteq \Omega_{\mathcal{P}_n}^h]}))$ , but the inner expected value can be bounded by  $h$  when  $\mathcal{K}_{[\mathcal{S} \subseteq \Omega_{\mathcal{P}_n}^h]} = 1$ , and by  $\max\{\text{diam}(\mathcal{S}) + h, h\}$  when  $\mathcal{K}_{[\mathcal{S} \subseteq \Omega_{\mathcal{P}_n}^h]} = 0$ . Using Chebyshev’s inequality, we find

$$\begin{aligned} \mathbb{P}(d_{\mathcal{H}}(\mathcal{S}, \Omega_{\mathcal{P}_n}^h) > \delta) &\leq \frac{h}{\delta} \mathbb{P}(\{\mathcal{S} \subseteq \Omega_{\mathcal{P}_n}^h\}) + \frac{\max\{\text{diam}(\mathcal{S}) + h, h\}}{\delta} (1 - \mathbb{P}(\{\mathcal{S} \subseteq \Omega_{\mathcal{P}_n}^h\})) \\ &\leq \frac{h}{\delta} + \frac{\text{diam}(\mathcal{S}) + h}{\delta} (1 - \mathbb{P}(\{\mathcal{S} \subseteq \Omega_{\mathcal{P}_n}^h\})), \end{aligned}$$

a quantity that goes to zero for any fixed  $\delta > 0$  as  $h \downarrow 0$  and  $n \uparrow \infty$ , provided that (16) holds.  $\square$

*Proof of Theorem 8.* Since the proof is almost identical to that of Theorem 7, many steps will be skipped. Note that since  $\mathcal{S}$  is compact, there exists an upper bound  $K$  for all its sectional curvatures. This will allow us to use the volume comparison theorems as before.

We can start from the adequate version of (7). We must bound both  $\mathbb{P}(\{\mathcal{S} \subseteq \Omega_{\mathcal{P}_n}^h\}^c)$  and  $\mathbb{P}(\mathcal{L}_{\mathcal{S}}(\mathcal{P}_n; \Delta, h) > \varepsilon | \{\mathcal{S} \subseteq \Omega_{\mathcal{P}_n}^h\})$ . The second term can be bounded in an identical way as its  $\Delta = 0$  counterpart was, obtaining

$$(20) \quad \mathbb{P}(\mathcal{L}_{\mathcal{S}}(\mathcal{P}_n; \Delta, h) > \varepsilon | \{\mathcal{S} \subseteq \Omega_{\mathcal{P}_n}^h\}) \leq \frac{C_{\mathcal{S}} \sqrt{h + \Delta} + 2\Delta(2 + \sqrt{2}C_{\mathcal{S}}\sqrt{\Delta})}{\varepsilon},$$

which vanishes as  $n \uparrow \infty$ .

Now we upper bound  $\mathbb{P}(\{\mathcal{S} \subseteq \Omega_{\mathcal{P}_n}^h\}^c)$ . Everything carries over in the same fashion as in the proof of Lemma 1, except that now we must take into consideration that the  $p_i$ ’s are not necessarily on  $\mathcal{S}$  but inside  $\Omega_{\mathcal{S}}^{\Delta}$ . Following the described steps, we obtain

$$(21) \quad \mathbb{P}(\{x \notin \Omega_{\mathcal{P}_n}^h \cap \mathcal{S}\}) \leq \left(1 - \frac{\mu(B(x, h) \cap \Omega_{\mathcal{S}}^{\Delta})}{\mu(\Omega_{\mathcal{S}}^{\Delta})}\right)^n.$$

Notice that, since we are working with  $h \geq \Delta$ , we have  $B(x, \Delta) \subset B(x, h) \cap \Omega_{\mathcal{S}}^{\Delta}$  (see Figure 8), and we can rewrite the bound in (21) as

$$(22) \quad \mathbb{P}(\{x \notin \Omega_{\mathcal{P}_n}^h \cap \mathcal{S}\}) \leq \left(1 - \frac{\mu(B(x, \Delta))}{\mu(\Omega_{\mathcal{S}}^{\Delta})}\right)^n$$

$$(23) \quad = \left(1 - \frac{\mu(B(\cdot, \Delta))}{\mu(\Omega_{\mathcal{S}}^{\Delta})}\right)^n.$$

We can bound this quantity using formulas akin to Weyl’s tube theorem. More precisely, as explained in Appendix D, we can write

$$\mu(\Omega_{\mathcal{S}}^{\Delta}) = \mu(\mathcal{S}) v(d - k, \Delta) + \varphi_{\mathcal{S}}(\Delta),$$

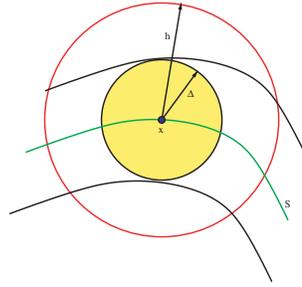


FIG. 8.  $B(x, \Delta) \subset B(x, h) \cap \Omega_S^\Delta$ .

where  $\frac{\varphi_S(\Delta)}{\Delta^{d-k+1}} \rightarrow 0$  as  $\Delta \rightarrow 0$  and  $v(D, R)$  is the volume of the ball of radius  $R$  in  $D$ -dimensional Euclidean space.

Now, for  $x \in S$  we must find a bound for  $\mathbb{P}(B_S(x, \delta) \not\subset \Omega_{\mathcal{P}_n}^h)$ , but as in the proof of Lemma 2,  $\mathbb{P}(B_S(x, \delta) \not\subset \Omega_{\mathcal{P}_n}^h) \leq \mathbb{P}(x \notin \Omega_{\mathcal{P}_n}^{h-\delta} \cap S)$ , which can be bounded by (22). Also the bound (14) for the covering number still works in this case, and thus we can write  $\mathbb{P}(S \not\subset \Omega_{\mathcal{P}_n}^h \cap S) \leq \frac{(1-y_\Delta)^n}{x_n}$ , where  $y_\Delta \triangleq \frac{\mu(B(\cdot, \Delta))}{\mu(\Omega_S^\Delta)}$ . Also since  $h \geq \Delta$ ,  $x_h \geq x_\Delta$ , then

$$\mathbb{P}(S \not\subset \Omega_{\mathcal{P}_n}^h) \leq \frac{(1 - y_\Delta)^n}{x_\Delta}.$$

But with  $\Delta$  small enough,  $y_\Delta \simeq \alpha \Delta^k$  and  $x_\Delta \simeq \beta \Delta^k$ , and then Lemma 4 and the hypotheses guarantee that  $\mathbb{P}(\{S \subset \Omega_{\mathcal{P}_n}^h\}^c) \rightarrow 0$  as  $n \uparrow \infty$ .  $\square$

**Appendix D. Basic differential geometry facts.** In this section we collect some facts that were used throughout the article, following [32].

**D.1. Measure of a  $d$ -dimensional ball.** Recall the definition of the  $\Gamma$  function:

$$\Gamma(\alpha) = \int_0^{+\infty} e^{-t} t^{\alpha-1} dt.$$

**THEOREM 11.** *The volume of  $d$ -dimensional ball of radius  $r$  is given by*

$$v(d, r) \triangleq \mu(B(\cdot, r)) = \omega_d r^d,$$

where  $\omega_d = \frac{2\pi^{d/2}}{d\Gamma(d/2)}$ .

**D.2. Bishop–Günther inequalities for the measure of a geodesic ball.**

**THEOREM 12.** *Let  $S$  be a complete  $k$ -dimensional Riemannian manifold, assume  $r$  to be smaller than the distance between  $m \in S$  and  $\text{Cut}(m, S)$  (cut locus of the points  $m$  in  $S$ ). Let  $K^S$  be the sectional curvatures of  $S$  and  $\gamma$  a constant. Then if  $\widehat{V}_\gamma(r) \triangleq \frac{2\pi^{k/2}}{\Gamma(k/2)} \int_0^r \left(\frac{\sin(t\sqrt{\gamma})}{\sqrt{\gamma}}\right)^{k-1} dt$ , then*

$$(24) \quad K^S \geq \gamma \text{ implies } \mu(B_S(m, r)) \leq \widehat{V}_\gamma(r),$$

$$(25) \quad K^S \leq \gamma \text{ implies } \mu(B_S(m, r)) \geq \widehat{V}_\gamma(r).$$

PROPOSITION 3. We have the following Taylor expansion for  $\widehat{V}_\gamma(r)$ , the volume of a geodesic ball in a space of constant sectional curvature  $\gamma$ :

$$\widehat{V}_\gamma(r) = \omega_k r^k \left( 1 - r^2 \frac{\gamma}{6} \frac{k(k-1)}{k+2} \right) + \phi(r),$$

where  $\frac{\phi(r)}{r^{k+2}} \rightarrow 0$  as  $r \downarrow 0$ .

**D.3. Weyl’s tube theorem.**

THEOREM 13. Let  $\mathcal{S}$  be a  $k$ -dimensional manifold topologically embedded in  $\mathbb{R}^d$ . Assume that  $\mathcal{S}$  is compact closure, and that every point in the tube  $T(\mathcal{S}, r) = \{x \in \mathbb{R}^d \text{ such that, } d(\mathcal{S}, x) \leq r\}$  has a unique shortest geodesic connecting it with  $\mathcal{S}$ ; then the volume  $\mu(T(\mathcal{S}, r))$  of the tube is given by

$$(26) \quad \mu(T(\mathcal{S}, r)) = r \frac{\sqrt{\pi}}{\Gamma(3/2)} \sum_{i=0}^{\lfloor \frac{d-1}{2} \rfloor} \frac{k_{2i}(\mathcal{S}) r^{2i}}{I(i)},$$

where  $I(i) = 1 \cdot 3 \cdot 5 \cdots (2i + 1)$  and the numbers  $k_{2i}$  depend on the curvature structure of  $\mathcal{S}$ . For our purposes we need know only that  $k_0 = \mu(\mathcal{S})$ .

COROLLARY 6. The volume of the tube  $T(\mathcal{S}, r)$  can be expanded as

$$\mu(T(\mathcal{S}, r)) = \mu(\mathcal{S}) v(d - k; r) + \phi_{\mathcal{S}}(r),$$

where  $\frac{\phi_{\mathcal{S}}(r)}{r^{d-k}} \xrightarrow{r \downarrow 0} 0$

**Appendix E. Details on object recognition.** The ideal objective is to actually compute the  $\mathcal{J}$ -distance between  $D_1$  and  $D_2$  as described in section 7.2; however, this is a very hard problem since there are  $m!$   $m \times m$  permutation matrices. The choice of  $m$  is subject to compromise: on one hand, we want it to be big enough so as to capture the metric structure of  $\mathcal{S}_i$  with the information given by  $(Q_n^{(i)}, d_{\Omega_{\mathcal{P}_n^{(i)}}^{h_i}})$ ; on the other hand, we want to be able to actually make the computations involved without too much processing cost. Therefore we should attempt to circumvent this  $m!$  search space by exploiting some other information we might have.

One possibility for bypassing this difficulty is to try to upper bound the  $\mathcal{J}$ -distance by some difference between eigenvalues of the matrices. However, it turns out that one can easily find two distance matrices which have positive  $\mathcal{J}$ -distance (they are not *cogredient*) but have the same spectra. Then an upper bound should take into account also another term that measures our inability to really differentiate distance functions by looking only at their eigenvalues. Of course this information must then be contained in the eigenvectors.<sup>30</sup>

A way of dealing with this particular issue is working with the spectral factorization of each of the matrices. Let  $D_1 = QDQ^T$  and  $D_2 = \widehat{Q}\widehat{D}\widehat{Q}^T$ , where  $Q$  and  $\widehat{Q}$  are unitary matrices and  $D$  and  $\widehat{D}$  are diagonal matrices whose entries are the eigenvalues of  $D_1$  and  $D_2$ , respectively. Note that we are not saying anything about the order in which those eigenvalues are presented; for convenience, let

<sup>30</sup>Another idea, for example, is the following: We know that the searched-for isometry (if it exists) must be a Lipschitz continuous map, and therefore it makes no sense to consider the huge set of transformations spanned by  $\mathcal{PM}_m$ . We leave the exploitation of this idea for future work.

$D_{11} = |D_{11}| > |D_{22}| > \dots > |D_{mm}|$  and  $\widehat{D}_{11} = |\widehat{D}_{11}| > |\widehat{D}_{22}| > \dots > |\widehat{D}_{mm}|$ .<sup>31</sup> Then with little effort we can write

$$(27) \quad \min_{P \in \mathcal{PM}_m} \|D_1 - PD_2P^T\| \leq \|(Q - P\widehat{Q})D\| + \|(Q - P\widehat{Q})\widehat{D}\| + \|D - \widehat{D}\|.$$

Note that if  $W$  is any matrix and  $T$  is diagonal, then  $\|WT\|^2 = \sum_k \|W_{(:,k)}t_{kk}\|^2 = \sum_k \|W_{(:,k)}\|^2 t_{kk}^2$  where  $W_{(:,k)}$  is the  $k$ th column vector of  $W$ . Using this observation, we note that the first two terms in (27) can be bounded as follows (let  $Q = (q_1 | \dots | q_m)$ , and  $\widehat{Q} = (\widehat{q}_1 | \dots | \widehat{q}_m)$ ):

$$\|(Q - P\widehat{Q})D\| + \|(Q - P\widehat{Q})\widehat{D}\| = \sqrt{\sum_k D_{kk}^2 \|q_k - P\widehat{q}_k\|^2} + \sqrt{\sum_k \widehat{D}_{kk}^2 \|q_k - P\widehat{q}_k\|^2}.$$

Now, using the trivial inequality  $\frac{\sqrt{a} + \sqrt{b}}{2} \leq \sqrt{\frac{a+b}{2}}$  for all  $a, b \geq 0$ , we finally arrive at the expression

$$(28) \quad \min_{P \in \mathcal{PM}_m} \|D_1 - PD_2P^T\| \leq \sqrt{\sum_{k=1}^m (D_{kk} - \widehat{D}_{kk})^2} + \sqrt{2} \sqrt{\sum_{k=1}^m (D_{kk}^2 + \widehat{D}_{kk}^2) \|q_k - P\widehat{q}_k\|^2}.$$

This inequality holds for any  $P \in \mathcal{PM}_m$ . It is important to note that in case  $D_1$  and  $D_2$  are cogredient, all their eigenvectors will also be related through that same permutation; therefore this inequality is sharp.<sup>32</sup>

Note that in the second term of (28), the values of  $\|q_i - P\widehat{q}_i\|$  are weighted by  $(D_{ii}^2 + \widehat{D}_{ii}^2)$ , so one can think that since  $\|q_i - P\widehat{q}_i\| \leq 2$ , the most important terms of the sum will be those for which  $(D_{ii}^2 + \widehat{D}_{ii}^2)$  is large. This is not a rigorous consideration, but gives some guidelines on how to compute an approximate bound when the sizes of the distance matrices are prohibitively large.

In some situations, the choice of the subsampled set size  $m$  that guarantees a good metric approximation in the sense discussed above might be too large, making the computation of the full bound (28) onerous. But still a measure of similarity must be provided which does not require the computation of all of the eigenvalues and eigenvectors of each distance matrix. Therefore, in order to estimate  $d_J(D_1, D_2)$ , we use the following idea: Instead of computing all the eigenvalues and eigenvectors of the matrices  $D_1$  and  $D_2$ , compute the  $N \ll m$  more important ones, where important means, in the light of the expression for the bound, those with the largest moduli, at least for the part of the bound involving eigenvectors. Then, for a (computationally) reasonable  $N$  we define the approximate error bound (still letting  $P$  be any convenient choice of a permutation matrix)

$$(29) \quad e(N) \triangleq \sqrt{\sum_{k=1}^N (D_{kk} - \widehat{D}_{kk})^2} + \sqrt{2} \sqrt{\sum_{k=1}^N (D_{kk}^2 + \widehat{D}_{kk}^2) \|q_k - P\widehat{q}_k\|^2}.$$

<sup>31</sup>We have used Frobenius theorem [51], which asserts that nonnegative matrices have a positive largest absolute value eigenvalue. Note that we have also assumed that there are no repeated eigenvalues.

<sup>32</sup>Note that from (27) one can obtain  $d_J(D_1, D_2) \leq \|D - \widehat{D}\| + (\|D\| + \|\widehat{D}\|)\|Q\widehat{Q}^T - P\|$ ; then one further idea to be explored is how to best approximate a given unitary matrix by a permutation matrix. This would not only allow us to obtain an explicit bound for the  $\mathcal{J}$ -distance, but would also provide us with a low metric distortion way of mapping  $S_1$  ( $\mathcal{P}_n^{(1)}$ ) into  $S_2$  ( $\mathcal{P}_n^{(2)}$ ), with applications like texture mapping, brain warping, etc.

Now, we fix the permutation  $P$  as follows: Let  $S$  be the permutation matrix such that  $Sq_1$  is a column vector whose components are sorted from largest to smallest. Do the same with  $\widehat{q}_1$  to obtain  $\widehat{S}$ ; then compare  $Sq_1$  with  $\widehat{S}\widehat{q}_1$ , which amounts to comparing  $q_1$  with  $S^T\widehat{S}$ ; hence we let  $P = S^T\widehat{S}$ . We could again use a more sophisticated way of choosing  $P$ , but this one suffices for demonstration purposes and, of course, achieves equality in (28) when both matrices are cogredient.

Another possibility is to directly compare the distance matrices according to the expression  $\|D_1 - PD_2P^T\|$ , using a certain sensible choice for  $P$ . We first put both matrices in a “canonical” order. Let  $(i_1, j_1)$  be one position on the matrix  $D_1$  with the maximum value. We then order the rest of the points in the set according to their distances to either  $q_{i_1}^{(1)}$  or  $q_{j_1}^{(1)}$  from smallest to largest.<sup>33</sup> This induces an ordering for the matrix  $D_1$ , letting  $P_1$  be the underlying permutation matrix. We do the same with  $D_2$  and obtain  $P_2$ . Finally we let

$$e_G(D_1, D_2) \triangleq \|D_1 - P_1^T P_2 D_2 P_2^T P_1\|,$$

and note that obviously  $d_J(D_1, D_2) \leq e_G(D_1, D_2)$  and that the inequality is sharp.

**E.1. Choice of the point cloud subset  $\mathcal{Q}^{(i)}$ .** In general, the number of points in the cloud is too big. This means that the actual computation of the distance matrices, if done using all the points in the cloud, and subsequent eigenvalue and eigenvector computations (if needed) become onerous. Therefore we need a procedure which allows us to select a small cardinality subset  $\mathcal{Q}_m$  of  $\mathcal{P}_n$  for which we will actually compute the approximate distance matrix, but still using  $\mathcal{P}_n$  to define the offset  $\Omega_{\mathcal{P}_n}^h$  inside which the computations are performed. This subset  $\mathcal{C}_r \subset \mathcal{P}_n$  must be “representative” of the geometry of the underlying manifold. One way of selecting those points is by not allowing them to cluster inside any region of the manifold. This can be accomplished in practice by using the “packing idea” in [24]: Given  $m < n$ , choose the first point  $c_1 \in \mathcal{C}_m$  randomly, then proceed by always choosing a point as far as possible from the set of points that have already been chosen. End the process when  $m$  points have been chosen. This is the procedure used in the experiments.

**Acknowledgments.** We acknowledge useful conversations on the topic of this paper with L. Aspirot, P. Bermolén, R. Coifman, D. Donoho, N. Dyn, J. Giesen, O. Gil, R. Gulliver, R. Kimmel, A. Pardo, and O. Zeitouni. We thank M. Levoy and the Digital Michelangelo Project for data provided for this project. We thank the anonymous reviewers for their comments, which helped improve the presentation of the paper.

REFERENCES

[1] R. ALEXANDER AND S. ALEXANDER, *Geodesics in Riemannian manifolds with boundary*, Indiana Univ. Math. J., 30 (1981), pp. 481–488.  
 [2] L. AMBROSIO AND H. M. SONER, *Level set approach to mean curvature flow in arbitrary co-dimension*, J. Differential Geom., 43 (1996), pp. 693–737.  
 [3] N. AMENTA, S. CHOI, AND R. KOLLURI, *The power crust, unions of balls, and the medial axis transform*, Comput. Geom., 19 (2001), pp. 127–153.  
 [4] N. AMENTA, S. CHOI, AND R. KOLLURI, *The power crust*, in Proceedings of the 6th Annual ACM Symposium on Solid Modeling, Ann Arbor, MI, 2001, ACM, New York, pp. 249–260.

<sup>33</sup>In our current implementation we don’t worry about repeated distance values, since this can be easily handled. We will present more details and further refinements elsewhere.

- [5] N. AMENTA AND R. KOLLURI, *Accurate and efficient unions of balls*, in Proceedings of the ACM Symposium on Computational Geometry, Hong Kong, ACM, New York, 2000, pp. 119–128.
- [6] T. APOSTOL, *Mathematical Analysis*, Addison–Wesley Ser. Math., Addison–Wesley, Reading, MA, 1974.
- [7] A. BARTESAGHI AND F. MÉMOLI, *Flujos software*, <http://iie.fing.edu.uy/investigacion/grupos/gti/flujos/flujos.html>.
- [8] A. BARTESAGHI AND G. SAPIRO, *A system for the generation of curves on 3D brain images*, Human Brain Mapping, 14 (2001), pp. 1–15.
- [9] M. BERNSTEIN, V. DE SILVA, J. LANGFORD, AND J. TENENBAUM, *Graph approximations to geodesics on embedded manifolds*, <http://isomap.stanford.edu/BdSLT.ps>.
- [10] *Blitz++ website*, <http://www.oonumerics.org/blitz>.
- [11] J.-D. BOISSONNAT AND F. CAZALS, *Coarse-to-fine surface simplification with geometric guarantees*, in Proceedings of EUROGRAPHICS 2001, A. Chalmers and T.-M. Rhyne, eds., Manchester, UK, 2001.
- [12] M. BOTSCH, A. WIRATANAYA, AND L. KOBBELT, *Efficient high quality rendering of point sampled geometry*, in Proceedings of the 13th EUROGRAPHICS Workshop on Rendering, Pisa, Italy, 2002, pp. 53–64.
- [13] E. CALABI, P. J. OLVER, AND A. TANNENBAUM, *Affine geometry, curve flows, and invariant numerical approximations*, Adv. Math., 124 (1996), pp. 154–196.
- [14] I. CHAVEL, *Riemannian Geometry: A Modern Introduction*, Cambridge University Press, Cambridge, UK, 1993.
- [15] R. COIFMAN, *personal communication*, Yale University, New Haven, CT, 2002.
- [16] R. COIFMAN, *personal communication*, Yale University, New Haven, CT, 2003 (talk presented at IPAM-UCLA).
- [17] R. COIFMAN, *personal communication*, Yale University, New Haven, CT, 2003 (talk presented at University of Minnesota).
- [18] T. K. DEY, J. GIESEN, AND J. HUDSON, *Decimating samples for mesh simplification*, in Proceedings of the 13th Annual Canadian Conference on Computational Geometry, Waterloo, ON, 2001, pp. 85–88.
- [19] D. L. DONOHO AND C. GRIMES, *When Does ISOMAP Recover the Natural Parametrization of Families of Articulated Images?*, Technical Report 2002-27, Department of Statistics, Stanford University, Stanford, CA, 2002.
- [20] N. DYN, M. S. FLOATER, AND A. ISKE, *Adaptive thinning for bivariate scattered data*, J. Comput. Appl. Math., 145 (2002), pp. 505–517.
- [21] M. DOCARMO, *Riemannian Geometry*, Birkhäuser Boston, Cambridge, MA, 1992.
- [22] A. DVORETSKY, *On covering a circle by randomly placed arcs*, Proc. Natl. Acad. Sci. USA, 42, (1956), pp. 199–203.
- [23] H. EDELSBRUNNER, *The union of balls and its dual shape*, Discrete Comput. Geom., 13 (1995), pp. 415–440.
- [24] A. ELAD (ELBAZ) AND R. KIMMEL, *Bending invariant representations for surfaces*, in Proceedings of the Computer Vision and Pattern Recognition (CVPR’01), Kauai, HI, 2001, pp. I-168–I-174.
- [25] R. ELLIS, X. JIA, AND C. YAN, *On Random Points in the Unit Disk*, submitted; available online at <http://www.math.tamu.edu/~rellis/papers/5random.pdf>.
- [26] H. FEDERER, *Curvature measures*, Trans. Amer. Math. Soc., 93 (1959), pp. 418–491.
- [27] L. FLATTO AND D. J. NEWMAN, *Random coverings*, Acta Math., 138 (1977), pp. 241–64.
- [28] L. FLATTO, *A limit theorem for random coverings of a circle*, Israel J. Math., 15 (1973), pp. 167–184.
- [29] M. S. FLOATER AND A. ISKE, *Thinning algorithms for scattered data interpolation*, BIT, 38 (1998), pp. 705–720.
- [30] S. F. FRISKEN, R. N. PERRY, A. P. ROCKWOOD, AND T. R. JONES, *Adaptively sampled distance fields: A general representation of shape for computer graphics*, in Proceedings of SIGGRAPH 2000, New Orleans, LA, ACM, New York, 2000, pp. 249–254.
- [31] J. GIESEN AND U. WAGNER, *Shape dimension and intrinsic metric from samples of manifolds with high co-dimension*, in Proceedings of the 19th ACM Symposium on Computational Geometry, San Diego, CA, 2003, pp. 329–337.
- [32] A. GRAY, *Tubes*, Addison–Wesley, Reading, MA, 1990.
- [33] M. ALEXA, M. GROSS, M. PAULY, H. PFISTER, M. STAMMINGER, AND M. ZWICKER, *Point Based Computer Graphics*, EUROGRAPHICS Lecture Notes, 2002, available online at <http://graphics.stanford.edu/~niloy/research/papers/ETH/PointBasedComputerGraphics-TutorialNotes.pdf>.
- [34] P. HALL, *Introduction to the Theory of Coverage Processes*, Wiley Series in Probability and

- Mathematical Statistics, John Wiley & Sons, New York, 1988.
- [35] J. HELMSEN, E. G. PUCKETT, P. COLLELA, AND M. DORR, *Two new methods for simulating photolithography development in 3D*, in Proceedings of the SPIE Microlithography, IX (1996), pp. 253.
- [36] P. W. JONES, *Rectifiable sets and the traveling salesman problem*, Invent. Math., 102 (1990), pp. 1–15.
- [37] S. JANSON, *Random coverings in several dimensions*, Acta Math., 156 (1986), pp. 83–118.
- [38] J. HOFFMANN-JØRGENSEN, *Coverings of metric spaces with randomly placed balls*, Math. Scand., 32 (1973), pp. 169–186.
- [39] M. G. KENDALL AND P. A. P. MORGAN, *Geometrical Probability*, Griffin, London, 1963.
- [40] R. KIMMEL AND J. A. SETHIAN, *Computing geodesic paths on manifolds*, Proc. Natl. Acad. Sci. USA, 95 (1998), pp. 8431–8435.
- [41] A. N. KOLMOGOROV AND S. V. FOMIN, *Elements of the Theory of Functions and Functional Analysis*, Dover, New York, 1999.
- [42] R. KUNZE, F. E. WOLTER, AND T. RAUSCH, *Geodesic Voronoi diagrams on parametric surfaces*, in Proceedings of the Computer Graphics International (CGI) '97, IEEE, Computer Society Press, Kinopolis, Belgium, 1997, Piscataway, NJ, 1997, pp. 230–237.
- [43] G. LEIBON AND D. LETSCHER, *Delaunay triangulations and Voronoi diagrams for Riemannian manifolds*, ACM Symposium on Computational Geometry 2000, Hong Kong, ACM, New York, 2000, pp. 341–349.
- [44] G. LERMAN, *How to Partition a Low-Dimensional Data Set into Disjoint Clusters of Different Geometric Structure*, preprint, 2000.
- [45] L. LINSEN, *Point Cloud Representation*, CS Technical Report, University of Karlsruhe, Karlsruhe, Germany, 2001.
- [46] L. LINSEN AND H. PRAUTZSCH, *Local versus global triangulations*, in Proceedings of EUROGRAPHICS, 2001, A. Chalmers and T-M. Rhyne, eds., Manchester, UK.
- [47] C. MANTEGAZZA AND A.C. MENUCCI, *Hamilton-Jacobi equations and distance functions on Riemannian manifolds*, Appl. Math. Optim., 47 (2003), pp. 1–25.
- [48] A. MARINO AND D. SCOLOZZI, *Geodetiche con ostacolo*, Boll. Un. Mat. Ital., 6 (1983), pp. 1–31.
- [49] F. MÉMOLI AND G. SAPIRO, *Fast computation of weighted distance functions and geodesics on implicit hyper-surfaces*, J. Comput. Phys., 173 (2001), pp. 730–764.
- [50] F. MÉMOLI AND G. SAPIRO, *Comparing point clouds*, in Proceedings of the 2nd Annual EUROGRAPHICS Symposium on Geometry Processing, Nice, France, 2004.
- [51] H. MINC, *Nonnegative Matrices*, Wiley Ser. Discrete Math. Optim., Wiley Interscience, New York, 1988.
- [52] J. S. B. MITCHELL, *An algorithmic approach to some problems in terrain navigation*, Artificial Intelligence, 37 (1988), pp. 171–201.
- [53] J. S. B. MITCHELL, D. PAYTON, AND D. KEIRSEY, *Planning and reasoning for autonomous vehicle control*, Internat. J. Intelligent Systems, 2 (1987), pp. 129–198.
- [54] N. J. MITRA AND A. NGUYEN, *Estimating surface normals in noisy point cloud data*, in Proceedings of the ACM Symposium on Computational Geometry, San Diego, 2003, ACM, New York, 2003, pp. 322–328.
- [55] C. MOENNING, F. MÉMOLI, G. SAPIRO, N. DYN, AND N. A. DODGSON, *Meshless Geometric Subdivision*, Technical report IMA TR 1977, 2004, available online from <http://www.ima.umn.edu/preprints/apr2004/apr2004.html>.
- [56] M. PAULY AND M. GROSS, *Spectral processing of point-sampled geometry*, in Proceedings of SIGGRAPH 2001, Los Angeles, 2001, ACM, New York, pp. 379–386.
- [57] M. PAULY, N. J. MITRA, AND L. GUIBAS, *Uncertainty and variability in point cloud surface data*, in Proceedings of the Symposium on Point-Based Graphics, Zurich, 2004.
- [58] M. PAULY, M. GROSS, AND L. KOBELT, *Efficient simplification of point-sampled surfaces*, in Proceedings of the IEEE Visualization meeting, 2002, pp. 163–170.
- [59] S. RUSINKIEWICZ AND M. LEVOY, *QSplat: A multiresolution point rendering system for large meshes*, in Proceedings of SIGGRAPH 2000, New Orleans, LA, ACM, New York, 2000, pp. 343–352.
- [60] T. SAKAI, *Riemannian Geometry*, AMS Translations of Mathematical Monographs 149, AMS, New York, 1996.
- [61] E. SCHWARTZ, A. SHAW, AND E. WOLFSON, *A numerical solution to the generalized mapmaker's problem: Flattening nonconvex polyhedral surfaces*, IEEE Trans. Pattern Anal. Machine Intelligence, 11 (1989), pp. 1005–1008.
- [62] J. SETHIAN, *Fast marching level set methods for three-dimensional photolithography development*, in Proceedings of the SPIE International Symposium on Microlithography, Santa Clara, CA, SPIE, Bellingham, WA, 1996.
- [63] J. A. SETHIAN, *A fast marching level-set method for monotonically advancing fronts*, Proc.

- Natl. Acad. Sci. USA, 93 (1996), pp. 1591–1595.
- [64] C. E. SHANNON, *Coding theorems for a discrete source with a fidelity criterion*, in *Information and Decision Processes*, McGraw-Hill, New York, 1960.
  - [65] L. A. SHEPP, *Covering the circle with random arcs*, *Israel J. Math.*, 11 (1972), pp. 328–345.
  - [66] H. SOLOMON, *Geometric Probability*, in *CBMS-NSF Reg. Conf. Ser. Appl. Math.*, 28, SIAM, Philadelphia, 1978.
  - [67] J. B. TENENBAUM, V. DE SILVA, AND J. C. LANGFORD, *A global geometric framework for nonlinear dimensionality reduction*, *Science*, 290 (2002), pp. 2319–2323.
  - [68] Y.-H. R. TSAI, L.-T. CHENG, S. OSHER, AND H.-K. ZHAO, *Fast sweeping algorithms for a class of Hamilton–Jacobi equations*, *SIAM J. Numer. Anal.*, 41 (2003), pp. 673–694.
  - [69] J. N. TSITSIKLIS, *Efficient algorithms for globally optimal trajectories*, *IEEE Trans. Automat. Control*, 40 (1995), pp. 1528–1538.
  - [70] *The Visualization Toolkit*, available online at <http://www.vtk.org>.
  - [71] F. E. WOLTER, *Cut Loci in Bordered and Unbordered Riemannian Manifolds*, Doctoral Dissertation, Technische Universität Berlin, Berlin, 1985.
  - [72] M. ZWICKER, M. PAULY, O. KNOLL, AND M. GROSS, *PointShop 3D: An interactive system for point-based surface editing*, *Proceedings of SIGGRAPH 2002*, San Antonio, TX, 2002, pp. 322–329.

Copyright of SIAM Journal on Applied Mathematics is the property of Society for Industrial and Applied Mathematics and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.