Adaptive Metrics for Adaptive Samples

Nicholas J. Cavanna [1],[†] [ID] **and Donald R. Sheehy** [1],[†] [ID] [2],*

[1]   University of Connecticut and Swift Health Systems, Inc.; nicholas.j.cavanna@gmail.com
[2]   North Carolina State University; don.r.sheehy@gmail.com
*   Correspondence: don.r.sheehy@gmail.com;
†   These authors contributed equally to this work.
‡   This work was partially supported by the National Science Foundation under grants CCF-1464379, CCF-1525978, and CCF-1652218.

Abstract: We generalize the local-feature size definition of adaptive sampling used in surface reconstruction to relate it to an alternative metric on Euclidean space. In the new metric, adaptive samples become uniform samples, making it simpler both to give adaptive sampling versions of homological inference results and to prove topological guarantees using the critical points theory of distance functions. This ultimately leads to an algorithm for homology inference from samples whose spacing depends on their distance to a discrete representation of the complement space.

Keywords: surface reconstruction; homology inference; adaptive sampling; topological data analysis

1. Introduction

1.1. From Points to Topology

Both surface reconstruction and homology inference are algorithmic problems that take points as input and produce a topological representation of the underlying space from which the points were drawn. In surface reconstruction, one often wants a homeomorphic reconstruction in the form of a triangulation, whereas in homology inference, it suffices to compute the homology groups. Although similar in many respects the general trend is that with weaker conditions on the input (i.e. noisier samples) one can only hope for weaker guarantees on the output (homology rather than homeomorphism). There is one aspect of these theories that directly contradicts this trend: **Many surface reconstruction algorithms are able to work with an adaptive sample, while most homology inference algorithms require a uniform[1] sample.** An adaptive sample has a density that adapts to some local sizing function. Thus, areas that require higher fidelity will have higher density (and smaller scale) while areas that can get by with less fidelity, will have lower density (and larger scale).

There have been some notable works that have bridged this gap between surface reconstruction and homology inference for adaptive samples. Most theoretically guaranteed surface reconstruction algorithms assume an input that is sufficiently dense with respect to the distance to the medial axis, a kind of skeleton describing the complement of the underlying shape. Cazals et al. [1] introduced the conformal alpha shape filtration as a way to build triangulations at different scales that have local connectivity related to the local feature size. Although their stated goal was surface reconstruction, the work employs many of the methods of homology inference. Chazal and Lieutier [2,3] gave a more direct generalization of methods in surface reconstruction with adaptive samples to homology inference, achieving some guarantees for smooth manifolds assuming both upper *and lower* bounds on the density. Dey et al. [4] give a homology inference algorithm algorithm for manifold data that attempts to sample a subset of the medial axis in order to approximate the local feature size. This

---

[1]   Here and throughout, we use "uniform" in the Hausdorff sense, not the statistical sense.

work was the main motivation for the current paper, and we have adopted their notation of X for the space and L for the approximation to the complement space. *We extend all these works by providing guaranteed homology inference for a much more general class of samples and spaces; we don't require the space to be a manifold or the sample to adapt to the medial axis.*

1.2. From Surface Reconstruction to Homology Inference

To reconstruct a surface from a point set, one needs the sample to be sufficiently dense with respect to not just the local curvature of the surface, but also the distance to parts of the surface that are close in the embedding but far in geodesic distance. Otherwise, algorithms have no way of identifying which geometrically close sample points correspond to local neighborhoods in the surface. Adaptive sampling with respect to the so-called *local feature size* as introduced by Amenta and Bern [5] neatly characterizes such "good" samples and was then used in many later works on surface reconstruction with topological guarantees [6]. There is an extensive literature on the problem in high dimensions (see [7,8] for recent examples) and the problem remains an active research area. Such *adaptive samples* are in contrast to *uniform samples* for which a single parameter determines the density. That parameter is usually driven by the minimum of the local feature size and results in a much larger sample.

Later work on generalizations of surface reconstruction and homology inference related the topology of unions of balls centered at a sample $\widehat{X}$ near the unknown set X to the topology of X itself. The most well-known such results were by Niyogi et al. [9,10]. A union of balls with a fixed radius can be viewed as a sublevel set of the distance function to $\widehat{X}$. If we have an adaptive sample, then we would like to scale the radii of the balls as well. However, if the sample is adaptive with respect to a local feature size defined as the distance to an unknown set L, another approximation $\widehat{L}$ near L is necessary. Indeed, one interpretation of some Voronoi-based surface reconstruction algorithms is that an approximation $\widehat{L}$ to the medial axis L is computed from the Voronoi diagram of the sample $\widehat{X}$ of the unknown surface X.

We present a new perspective on adaptive samples. For *any* pair of disjoint, compact sets X and L, we define a metric on $\mathbb{R}^d \setminus L$ with the property that a uniform sample of X in the new metric corresponds to an adaptive sample in the Euclidean metric. We call this the metric induced by L or simply the *induced metric* for short. This new metric can also be extended to an arbitrarily close Riemannian metric over the same domain. Our main motivation is to connect adaptive sampling theory to the critical point theory of distance functions used extensively to prove topological guarantees in topological data analysis [2,11,12]. That theory gives natural topological equivalences between sublevel sets of distance functions to compact sets in Riemannian metrics. Thus, we propose to use the induced metric as the underlying ideal object and then relate it to a union of Euclidean balls constructed from approximations of X and L. Our metric can be viewed as a smoothed version of a metric used by Clarkson [13]. Our new formulation reveals connections with work on path planning [14,15] and density-based distances [16,17]. These are all constructions where one looks at conformal change of metrics induced by subsets of Euclidean space.

1.3. Overview

We lay out the main objects of study in Section 2. This includes the induced metric and a discrete approximation. Throughout the paper, we will relate these two objects or variations thereof for different purposes. In Section 3.1, we prove the relationship between the adaptive samples used in surface reconstruction and uniform samples in the induced metric. The definition of the induced metric does not lend itself to direct computation. So, in Section 3.2, we bound the interleaving distance between the induced metric and its discrete approximation. This interleaving is then used in Section 3.3 to give a homology inference algorithm, that is guaranteed to recover the homology of a sublevel set of the induced metric under certain sampling conditions.

2. Methods

Let L and X be compact subsets of $\mathbb{R}^d$ with respect to the Euclidean metric. For $x, y \in \mathbb{R}^d$, define $\mathrm{Path}(x, y)$ to be the set of bounded piecewise-C_1 paths from x to y, parametrized by Euclidean arc-length. Similarly, $\mathrm{Path}(x, S) := \bigcup_{s \in S} \mathrm{Path}(x, s)$ denotes all paths from x to a set S.

For any compact set $L \subseteq \mathbb{R}^d$, define $f_L(\cdot) : \mathbb{R}^d \to \mathbb{R}$ by

$$f_L(x) := \min_{\ell \in L} \|x - \ell\|.$$

Define

$$\mathrm{d}^L(x, y) := \min_{\gamma \in \mathrm{Path}(x, y)} \int_\gamma \frac{dz}{f_L(z)}.$$

The length of a unit-speed path $\gamma : [0, a] \to \mathbb{R}^d$ is denoted as

$$|\gamma| := \int_\gamma dz = \int_0^a dt.$$

For $y \in \mathbb{R}^d$, define

$$f_X^L(y) := d^L(y, X) = \min_{x \in X} \mathrm{d}^L(y, x),$$

and

$$\widehat{f_X^L}(y) := \min_{x \in X} \frac{\|y - x\|}{f_L(x)}.$$

Note that $f_X^L(\cdot)$ is a distance function, while $\widehat{f_X^L}(\cdot)$ is not. The latter function can be interpreted as a first-order approximation of the former.

Definition 1. *For any compact set $X \subset \mathbb{R}^d \setminus L$, for some compact set $L \subset \mathbb{R}^d$, the α-offsets with respect to d^L are*

$$A_X^L(\alpha) := \{x \in \mathbb{R}^d \mid f_X^L(x) \leq \alpha\}.$$

The distance function $f_L(\cdot)$ can be transformed into an arbitrarily close smooth function $\tilde{f}_L(\cdot)$ [18], yielding a Riemannian metric $\tilde{\mathrm{d}}_L$ defined in an identical manner as d_L. From this, one has corresponding α-offsets $\tilde{A}_X^L(\alpha)$ that are arbitrarily close to $A_L^X(\alpha)$. We will encounter this smoother version in Section 3.3.

We will approximate the offsets $A_X^L(\alpha)$ by a union of balls as follows.

Definition 2. *For any compact set $X \subset \mathbb{R}^d \setminus L$, for some compact set $L \subset \mathbb{R}^d$, the* approximate α-offsets *with respect to d^L are*

$$B_X^L(\alpha) := (\widehat{f_X^L})^{-1}[0, \alpha] = \bigcup_{x \in X} \mathrm{ball}(x, \alpha f_L(x)).$$

A useful property of $f_X^L(\cdot)$ is that it a 1-Lipschitz function. In general, a function f between two metric spaces (X, d_X) and (Y, d_Y) is said to be k-Lipschitz if for all $x, y \in X$, $\mathrm{d}_Y(f(x), f(y)) \leq k\mathrm{d}_X(x, y)$.

Lemma 1. *The function f_X^L is 1-Lipschitz from the metric space $(\mathbb{R}^d, \mathrm{d}^L)$ to $\mathbb{R}$.*

Proof. Fix any $a, b \in \mathbb{R}^d$. There exists point $x \in X$ and a path $\gamma_1 \in \mathrm{Path}(a, x)$ such that $f_X^L(a) = \int_{\gamma_1} \frac{dz}{f_L(z)}$. Likewise, there exists $\gamma_2 \in \mathrm{Path}(a, b)$ such that $\mathrm{d}^L(a, b) = \int_{\gamma_2} \frac{dz}{f_L(z)}$.

This implies that the concatenation of γ_1 and γ_2 is a path γ_3 in $\mathrm{Path}(b, X)$. Thus $f_X^L(b) \leq \int_{\gamma_3} \frac{dz}{f_L(z)} \leq f_X^L(a) + \mathrm{d}^L(a, b)$. As this holds for all a, b, we conclude that $|f_X^L(a) - f_L^X(b)| \leq \mathrm{d}^L(a, b)$ as desired. $\square$

We can use f_X^L to the define the Hausdorff distance, which is a metric between compact sets. This metric is useful for stating bounds on the quality, or uniformity, of a sample near a set.

Definition 3. *The* Hausdorff distance *between two compact sets $X, Y \in (\mathbb{R}^d, d^L)$ is defined as*

$$d_H^L(X, Y) = \max\{\max_{x \in X} f_Y^L(x), \max_{y \in Y} f_X^L(y)\}$$

If the Hausdorff distance between a compact set and a sample is bounded, Lemma 2 shows their α-offsets are interleaved at particular scales.

Lemma 2. *Let $\widehat{X}, X \subseteq \mathbb{R}^d \setminus L$ be such that $d_H^L(\widehat{X}, X) \leq \delta$. Then, for all $\alpha \geq 0$, $A_X^L(\alpha) \subseteq A_{\widehat{X}}^L(\alpha + \delta)$ and $A_{\widehat{X}}^L(\alpha) \subseteq A_X^L(\alpha + \delta)$.*

Proof. Let $y \in A_X^L(\alpha)$ be any point. By the definition of A_X^L, we have $f_X^L(y) \leq \alpha$. So, there exists $x \in X$ such that $d^L(x, y) \leq \alpha$. The Hausdorff assumption that $d_H^L(\widehat{X}, X) \leq \delta$ implies that for all $x \in X$, we have $f_{\widehat{X}}^L(x) \leq \delta$. By Lemma 1, $f_{\widehat{X}}^L(y) \leq f_{\widehat{X}}^L(x) + d^L(x, y) \leq \delta + \alpha$, implying $y \in A_{\widehat{X}}^L(\alpha + \delta)$. The second inclusion is proven by a symmetric argument. $\square$

The following is the definition of an adaptive sample we will use throughout. For the special case when X is a manifold and L is its medial axis, it corresponds to the ε-sample used in surface reconstruction.

Definition 4. *Given a compact set $L \subset \mathbb{R}^d$ and compact sets $X, \widehat{X} \subset \mathbb{R}^d \setminus L$ such that $\widehat{X} \subseteq X$, we say that $\widehat{X}$ is an ε-sample of X, for $\varepsilon \in [0, 1)$, if for all $x \in X$, there exists $p \in \widehat{X}$ such that $\|x - p\| \leq \varepsilon f_L(x)$.*

This definition is closely related to that of the approximate α-offsets, because if $\widehat{X}$ is an ε-sample of X, then for all $x \in X$, $\text{ball}(x, \varepsilon f_L(x)) \cap \widehat{X} \neq \emptyset$.

3. Results

3.1. Adaptive Sampling

In this section, we prove that a uniform sample in the induced metric corresponds to an adaptive sample in the Euclidean metric and vice versa. The key to this proof is the following lemma, which will also be used for the more elaborate interleaving results of Section 3.2.

Lemma 3. *Let $L \subset \mathbb{R}^d$ be a compact set and let $a, b \in \mathbb{R}^d \setminus L$. Then, the following two statements hold for all $\delta \in [0, 1)$.*

(i) If $d^L(a, b) \leq \delta$ then $\frac{\|a-b\|}{f_L(a)} \leq \frac{\delta}{1-\delta}$.

(ii) If $\frac{\|a-b\|}{f_L(a)} \leq \delta$ then $d^L(a, b) \leq \frac{\delta}{1-\delta}$.

Proof. To prove (i), we assume $d^L(a, b) \leq \delta$. Let γ be the path in $\text{Path}(a, b)$ such that $d^L(a, b) = \int_\gamma \frac{dz}{f_L(z)} < \delta$. Then we have the following inequalities following from the Lipschitz property of f_L.

$$\begin{aligned}
|\gamma| = \int_\gamma dz &= (f_L(a) + |\gamma|) \int_\gamma \frac{dz}{f_L(a) + |\gamma|} \\
&\leq (f_L(a) + |\gamma|) \int_\gamma \frac{dz}{f_L(z)} \\
&\leq (f_L(x) + |\gamma|)\delta
\end{aligned}$$

It follows that $|\gamma| \leq \frac{\delta}{1-\delta} f_L(x)$. Because $\|a - b\|$ is the length of the shortest path between a and b in the Euclidean metric, we conclude that $\|a - b\| \leq |\gamma| \leq \frac{\delta}{1-\delta} f_L(x)$.

Next we prove (ii). Assume $\frac{\|a-b\|}{f_L(a)} \leq \delta$. For all points z in the straight line segment $\overline{ab}$,

$$f_L(z) \geq f_L(a) - \|a - z\| \geq f_L(a) - \|a - b\| \geq (1 - \delta) f_L(a).$$

This implies the following inequality.

$$
\begin{aligned}
\mathrm{d}^L(a, b) &= \inf_{\gamma \in \mathrm{Path}(a,b)} \int_\gamma \frac{dz}{f_L(z)} \\
&\leq \int_{\overline{ab}} \frac{dz}{f_L(z)} \\
&\leq \frac{1}{(1-\delta) f_L(a)} \int_{\overline{ab}} dz \\
&= \frac{\|a - b\|}{(1-\delta) f_L(a)} \\
&\leq \frac{\delta}{1-\delta}.
\end{aligned}
$$

$\square$

We can now state the main theorem relating adaptive samples in the Euclidean metric to uniform samples in the metric induced by a set L.

Theorem 1. *Let L and X be compact sets, let $\widehat{X} \subset X$ be a sample, and let $\varepsilon \in [0, 1)$ be a constant. If $\widehat{X}$ is an ε-sample of X with respect to the distance to L, then $\mathrm{d}_H^L(X, \widehat{X}) \leq \frac{\varepsilon}{1-\varepsilon}$. Also, if $\mathrm{d}_H^L(X, \widehat{X}) \leq \varepsilon < \frac{1}{2}$, then $\widehat{X}$ is an $\frac{\varepsilon}{1-\varepsilon}$-sample of X with respect to the distance to L.*

Proof. Given $x \in X$, there exists $p \in \widehat{X}$ such that $\|x - p\| \leq \varepsilon f_L(x)$. By Lemma 3, $\mathrm{d}^L(x, p) \leq \frac{\varepsilon}{1-\varepsilon}$, so for all $x \in X$, $f_{\widehat{X}}^L(x) \leq \frac{\varepsilon}{1-\varepsilon}$. As $\widehat{X} \subseteq X$, this proves $\mathrm{d}_H^L(\widehat{X}, X) \leq \frac{\varepsilon}{1-\varepsilon}$.

Also, $\mathrm{d}_H^L(\widehat{X}, X) \leq \varepsilon < \frac{1}{2}$ implies that for all $x \in X$, $f_{\widehat{X}}^L(x) \leq \varepsilon$, thus there exists $p \in \widehat{X}$ such that $\mathrm{d}^L(x, p) \leq \varepsilon$. Thus by Lemma 3 $\|x - p\| \leq \frac{\varepsilon}{1-\varepsilon} f_L(x)$. Since $\varepsilon < \frac{1}{2}$, then $\frac{\varepsilon}{1-\varepsilon} < 1$, so $\widehat{X}$ is an $\frac{\varepsilon}{1-\varepsilon}$-sample of X. $\square$

3.2. Interleaving

A *filtration* is a nested family of sets. In this paper, we consider filtrations F parameterized by a real number $\alpha \geq 0$ so that $F(\alpha) \subset \mathbb{R}^d$ and whenever $\alpha < \beta$ we have $F(\alpha) \subseteq F(\beta)$. Often, our filtrations are *sublevel filtrations* of a real valued function $f : \mathbb{R}^d \to \mathbb{R}$. The sublevel filtration F corresponding to the function f is the defined as

$$F(\alpha) := \{x \in \mathbb{R}^d \mid f(x) \leq \alpha\}.$$

Definition 5. *A pair of filtrations (F, G) is (h_1, h_2)-interleaved in an interval (s, t) if $F(r) \subseteq G(h_1(r))$ whenever $r, h_1(r) \in (s, t)$ and $G(r) \subseteq F(h_2(r))$ whenever $r, h_2(r) \in (s, t)$. We require that the functions h_1, h_2 be nondecreasing in (s, t).*

The following lemma gives us an easy way to combine interleavings.

Lemma 4. *If (F, G) is (h_1, h_2)-interleaved in (s_1, t_1), and (G, H) is (h_3, h_4)-interleaved in (s_2, t_2), then (F, H) is $(h_3 \circ h_1, h_2 \circ h_4)$-interleaved in (s_3, t_3), where $s_3 = \max\{s_1, s_2\}$ and $t_3 = \min\{t_1, t_2\}$.*

Proof. If $r, h_3(h_1(r)) \in (s_3, t_3)$, then we have $F(r) \subseteq G(h_1(r)) \subseteq H(h_3(h_1(r)))$. Similarly, if $r, h_2(h_4(r)) \in (s_3, t_3)$, then $H(r) \subseteq G(h_4(r)) \subseteq F(h_2(h_4(r)))$. $\square$

3.2.1. Approximating X with $\widehat{X}$

Ultimately, the goal is to relate A_X^L, the offsets in the induced metric, to $B_{\widehat{X}}^{\widehat{L}}$, the approximate offsets computed from approximations (or samples) to both X and L. This relationship will be given by an interleaving that is built up from an interleaving for each approximation step. For each of the following lemmas, let $L, \widehat{L} \subset \mathbb{R}^d$ and $X, \widehat{X} \subset \mathbb{R}^d \setminus (L \cup \widehat{L})$ be compact sets.

Lemma 5. *If* $d_H^L(\widehat{X}, X) \leq \varepsilon$, *then* $(A_X^L, A_{\widehat{X}}^L)$ *are* (h_1, h_1)-*interleaved in* $(0, \infty)$, *where* $h_1(r) = r + \varepsilon$.

Proof. This lemma is a reinterpretation of Lemma 2 in the interleaving notation. □

3.2.2. Approximating the Induced Metric

It is much easier to use a union of Euclidean balls to model the sublevel sets of the distance function f_X^L. Below, we show that this is a reasonable approximation. The following results may also be viewed as a strengthening of the adaptive sampling result of the previous section (Theorem 1).

Lemma 6. *The pair* $(A_{\widehat{X}}^L, B_{\widehat{X}}^L)$ *are* (h_2, h_2)-*interleaved in* $(0, 1)$, *where* $h_2(r) = \frac{r}{1-r}$.

Proof. It will suffice to show that for $r \in [0, 1)$, $A_{\widehat{X}}^L(r) \subseteq B_{\widehat{X}}^L(\frac{r}{1-r})$, and for $r \in [0, \frac{1}{2})$, $B_{\widehat{X}}^L(r) \subseteq A_{\widehat{X}}^L(\frac{r}{1-r})$.

Take $y \in A_{\widehat{X}}^L(r)$ so that $f_{\widehat{X}}^L(y) \leq r$. Thus there exists $x \in X$ such that $d^L(x, y) \leq r$. By Lemma 3, this implies that $\|x - y\| \leq \frac{r}{1-r} f_L(x)$, which implies that $y \in B_{\widehat{X}}^L(\frac{r}{1-r})$.

Consider any point $y \in B_{\widehat{X}}^L(r)$. For some $x \in X$, we have $y \in \text{ball}(x, r f_L(x))$, so $\|x - y\| \leq r f_L(x)$. Applying Lemma 3, we have that $d^L(x, y) \leq \frac{r}{1-r}$. Finally, $y \in A_{\widehat{X}}^L(\frac{r}{1-r})$, because $f_{\widehat{X}}^L(y) \leq d^L(x, y)$. □

3.2.3. Approximating L with $\widehat{L}$

Usually, the set L is unknown at the start and must be estimated from the input. For example, if L is the medial axis of X, there are several known techniques for approximating L by taking some vertices of the Voronoi diagram [5,6]. We would like to give some sampling conditions that allow us to replace L with an approximation $\widehat{L}$. Interestingly, the sampling conditions for $\widehat{X}$ are dual to those used for $\widehat{L}$: we require $d_H^{\widehat{X}}(L, \widehat{L}) \leq \varepsilon$. In other words, $\widehat{L}$ must be an adaptive sample with respect to the distance to $\widehat{X}$.

Lemma 7. *If* $d_H^{\widehat{X}}(L, \widehat{L}) \leq \delta < 1$, *then* $(B_{\widehat{X}}^L, B_{\widehat{X}}^{\widehat{L}})$ *is* (h_3, h_3)-*interleaved in* $(0, \infty)$, *where* $h_3(r) = \frac{r}{1-\delta}$.

Proof. Fix any $x \in B_{\widehat{X}}^L(r)$. There is a point $p \in \widehat{X}$ such that $\frac{\|x - p\|}{f_L(p)} \leq r$. Moreover, there is a nearest point $z \in \widehat{L}$ to x such that $f_{\widehat{L}}(p) = \|p - z\|$. Lemma 3 and the assumption that $d_H^{\widehat{X}}(L, \widehat{L}) \leq \delta$ together imply that there exists $y \in L$ such that

$$\|y - z\| \leq \frac{\delta}{1 - \delta} f_{\widehat{X}}(z). \tag{1}$$

It then follows from the definitions that

$$f_{\widehat{X}}(z) = \min_{q \in \widehat{X}} \|z - q\| \leq \|z - p\| = f_{\widehat{L}}(p). \tag{2}$$

So, we can bound $f_L(p)$ in terms of $f_{\widehat{L}}(p)$ as follows.

$$\begin{aligned}
f_L(p) &\leq \|y - p\| & [y \in L] \\
&\leq \|y - z\| + \|z - p\| & [\text{triangle inequality}] \\
&\leq \frac{1}{1 - \delta} f_{\widehat{L}}(p) & [\text{by (1) and (2)}]
\end{aligned}$$

So,

$$\frac{\|x - p\|}{f_{\widehat{L}}(p)} \leq \frac{\|x - p\|}{(1 - \delta) f_L(p)} \leq \frac{r}{1 - \delta} = h_3(r).$$

Therefore, $x \in B_{\widehat{X}}^{\widehat{L}}(h_3(r))$ and so we conclude that $B_{\widehat{X}}^L(r) \subseteq B_{\widehat{X}}^{\widehat{L}}(h_3(r))$. The proof is symmetric to show that $B_{\widehat{X}}^{\widehat{L}}(r) \subseteq B_{\widehat{X}}^L(h_3(r))$ $\quad\square$

3.2.4. Putting it all together

We can now use Lemma 4 to combine the interleavings established in Lemmas 5, 6, and 7.

Theorem 2. *Let $L, \widehat{L} \subset \mathbb{R}^d$ and $X, \widehat{X} \subset \mathbb{R}^d \setminus (L \cup \widehat{L})$ be compact sets. If $d_H^{\widehat{X}}(L, \widehat{L}) \leq \delta < 1$ and $d_H^L(\widehat{X}, X) \leq \varepsilon < 1$, then $(A_X^L, B_{\widehat{X}}^{\widehat{L}})$ are (h_4, h_5)-interleaved in $(0, 1)$, where $h_4(r) = \frac{r + \varepsilon}{(1 - r - \varepsilon)(1 - \delta)}$ and $h_5(r) = \frac{r}{1 - \delta - r} + \varepsilon$.*

Proof. We use Lemma 4 to combine the interleavings from Lemmas 5, 6, and 7 to conclude that the pair $(A_X^L, B_{\widehat{X}}^{\widehat{L}})$ is $(h_3 \circ h_2 \circ h_1, h_1 \circ h_2 \circ h_3)$ interleaved in $(0, 1)$. To complete the proof, we expand $h_3 \circ h_2 \circ h_1$ and $h_1 \circ h_2 \circ h_3$ as follows.

$$\begin{aligned}
(h_3 \circ h_2 \circ h_1)(r) = (h_3 \circ h_2)(r + \delta) &= h_3\left(\frac{r + \delta}{1 - r - \delta}\right) \\
&= \frac{r + \delta}{(1 - r - \delta)(1 - \varepsilon)} \\
(h_1 \circ h_2 \circ h_3)(r) = (h_1 \circ h_2)\left(\frac{r}{1 - \varepsilon}\right) &= h_1\left(\frac{r}{(1 - \varepsilon)(1 - \frac{r}{1 - \varepsilon})}\right) \\
&= h_1\left(\frac{r}{1 - \varepsilon - r}\right) \\
&= \frac{r}{1 - \varepsilon - r} + \delta
\end{aligned}$$

So we have that $h_4(r) = \frac{r + \delta}{(1 - r - \delta)(1 - \varepsilon)}$ and $h_5(r) = \frac{r}{1 - \varepsilon - r} + \delta$. $\quad\square$

3.3. Smooth Adaptive Distance and Homology Inference

In the preceding sections, we have shown how to approximate (via interleaving) A_X^L, the sublevels of the distance to X in the induced metric, using a finite set of Euclidean balls, $B_{\widehat{X}}^{\widehat{L}}$. Now, we show how and when such an approximation gives a guarantee about the underlying space X itself. This is substantially more difficult, because it requires us to relate the sublevels of the induced metric to an object we don't have direct access to. As such, we will require some stronger hypotheses.

We will first review the critical point theory of distance functions. Then, we show how to smooth the induced metric to an arbitrarily close Riemannian metric, rendering the critical point theory applicable. Then, we put these together to prove the main inference result of the paper, Theorem 3.

3.3.1. Critical Points of Distance Functions

In this section, we give a minimal presentation of the critical point theory of distance functions to explain and motivate the results about interleaving offsets of distance functions in Riemannian manifolds. The main fact we use is that such interleavings lead immediately to results about homology inference (Lemma 9).

For a smooth Reimannian manifold M and a compact subset $X \subset M$, one can consider the function $f_X : M \to \mathbb{R}$ that maps each point in M to the distance to its nearest point in X as measured by the metric on the manifold. The gradient of f_X can be defined on M and the critical points are those points for which the gradient is zero. The critical values of f_X are those values of r such that $f_X^{-1}(r)$ contains a critical point. The critical point theory of distance functions developed by Grove and others [11] extends ideas from Morse theory to such distance functions. In particular, the theory gives the following result.

Lemma 8 (Grove [11]). *If $[r, r']$ contains no critical values then $f_X^{-1}([0, r]) \hookrightarrow f_X^{-1}([0, r'])$ is a homotopy equivalence.*

This means that for intervals that do not contain critical values, the inclusion maps in the filtration $F(r) := \{f_X^{-1}([0, r]) \mid r \geq 0\}$ are all homotopy equivalences and therefore induce isomorphisms in homology. This is used to give some information about the homology of filtrations that are interleaved with F.

We write H_* to denote homology over a field. So, for a set $X \subseteq \mathbb{R}^d$, we have a vector space $H_*(X)$ and for a continuous map $f : X \to Y$, we have a linear map $H_*(f)$. For the canonical inclusion map $X \hookrightarrow Y$ for a subset $X \subseteq Y$, we will denote the corresponding linear map in homology as $H_*(X \hookrightarrow Y)$. The image of this map is denoted im $H_*(X \hookrightarrow Y)$.

Lemma 9. *Let f_X be the distance function to a compact set in a Riemannian manifold such that $[r, r']$ contains no critical values of f_X. Let F be the sublevel filtration of f_X and let G be a filtration such that (F, G) are (h_1, h_2)-interleaved in (r, r'). If $r' < (h_2 \circ h_1 \circ h_2 \circ h_1)(r)$, then*

$$\text{im } H_*(G(h_1(r)) \hookrightarrow G((h_1 \circ h_2 \circ h_1)(r))) \cong H_*(F(r)).$$

Proof. The interleaving and the hypotheses imply that we have the following inclusions.

$$F(r) \subseteq G(h_1(r)) \subseteq F((h_2 \circ h_1)(r)) \subseteq G((h_1 \circ h_2 \circ h_1)(r)) \subseteq F((h_2 \circ h_1 \circ h_2 \circ h_1)(r))$$

The preceding lemma implies that the maps $F(r) \hookrightarrow F((h_2 \circ h_1)(r))$, $F((h_2 \circ h_1)(r)) \hookrightarrow F((h_2 \circ h_1 \circ h_2 \circ h_1)(r))$, and $F(r) \hookrightarrow F((h_2 \circ h_1 \circ h_2 \circ h_1)(r))$, all induce isomorphisms in homology. It follows that im $H_*(G(h_1(r)) \hookrightarrow G((h_1 \circ h_2 \circ h_1)(r))) \cong H_*(F(r))$, because the inclusion of spaces in G is factored through a space in F and it factors an inclusion of spaces, all of which are isomorphic in homology. ☐

3.3.2. Smoothing The Metric

To apply the critical point theory of distance functions to the induced metric directly, we would need it to be a smooth Riemannian manifold. Although, it is not smooth, we can smooth it with an arbitrarily small change. The process, though a little technical, is not surprising, nor very difficult. It proceeds in three steps.

1. We smooth the distance to L. This is the source of non-smoothness in the induced metric. This replaces f_L with a smooth approximation, $\widetilde{f_L}$.
2. The smoothed distance to L is used to define the smoothed induced metric $\widetilde{d^L}$ analogously to the original construction of d^L.

3. The induced distance function $f_{\widetilde{X}}^L$ can then be replaced by its smoothed version $\widetilde{f_X^L}$ and the corresponding smoothed offsets $\widetilde{A_X^L}$ are then well defined.

The complete construction of the smoothed offsets is presented in Appendix A. The end result is an interleaving between the induced offsets A_X^L and the smoothed version $\widetilde{A_X^L}$ as expressed in the following lemma.

Lemma 10. *Given $\alpha, \beta \in (0,1)$, consider compact sets $\widehat{L} \subseteq L \subset \mathbb{R}^d$ and compact sets $\widehat{X} \subseteq X \subset \mathbb{R}^d \setminus L^\beta$, such that $d_H^{\widehat{X}}(L, \widehat{L}) \leq \delta < 1$ and $d_H^L(\widehat{X}, X) \leq \varepsilon < 1$, then $(\widetilde{\mathcal{A}}_X^L, \mathcal{B}_{\widehat{X}}^{\widehat{L}})$ are (h_8, h_9)-interleaved on $(0,1)$, where $h_8(r) = \frac{r + \alpha r + \varepsilon}{(1 - r - r\alpha - \varepsilon)(1 - \delta)}$ and $h_9(r) = \frac{r}{(1-\alpha)(1-\delta-r)} + \frac{\varepsilon}{1-\alpha}$.*

Proof. The proof can be found in Section A.1 □

3.3.3. The weak feature size

Chazal and Leutier [19] introduced the *weak feature size (wfs)* as the least positive critical value of a Riemannian distance function. We denote the weak feature size with respect to $\widetilde{f_X^L}(\cdot)$ as $\text{wfs}^L(X)$. In light of the critical point theory of distance functions, a bound on the weak feature size gives a guaranteed interval with no critical points. This allows one to infer the homology from another filtration (usually one that is discrete and built from data) as long as the second filtration is interleaved in that critical point free interval.

Lemma 11 (Adapted from [19] Theorem 4.2, see also [20]). *Let S and $\widehat{S}$ be compact subsets of $\mathbb{R}^d$. If $d_H(S, \widehat{S}) > \varepsilon$ and $\text{wfs}(S) > 4\varepsilon$, then for all sufficiently small $\eta > 0$,*

$$H_*(A_S(\eta)) \cong \text{im } H_*(A_{\widehat{S}}(\varepsilon) \hookrightarrow A_{\widehat{S}}(3\varepsilon)).$$

The key idea in that proof is that the Hausdorff bound gives an interleaving while the weak feature size bound gives the interval without critical points. The technical condition regarding η is present to account for strange compact sets that may be homologically different from their arbitrarily small offsets. It is reasonable to assume that for some sufficiently small η, that $H_*(A_S(\eta)) \cong H_*(S)$ and thus one could "compute" the homology of S using only the sample $\widehat{S}$.

Most previous uses of the weak feature size have been applied in Euclidean spaces, but the critical point theory of distance functions can be applied more broadly to other smooth Riemannian manifolds. This is why we have introduced it as wfs^L (with the superscript) to indicate the underlying metric.

3.3.4. Homology Inference

We have now introduced all the necessary pieces to prove our main homology inference result.

Theorem 3. *Given $\alpha, \beta \in (0,1)$, consider compact sets $\widehat{L} \subseteq L \subset \mathbb{R}^d$ and compact sets $\widehat{X} \subseteq X \subset \mathbb{R}^d \setminus L^\beta$, such that $d_H^{\widehat{X}}(L, \widehat{L}) \leq \delta < 1$ and $d_H^L(\widehat{X}, X) \leq \varepsilon < 1$. Define the real-valued functions Ψ and Φ as*

$$\Psi(r) := \frac{r + \alpha r + \varepsilon}{(1 - r - \alpha r - \varepsilon)(1 - \delta)}$$

and

$$\Phi(r) := \frac{r}{(1-\alpha)(1-\delta-r)} + \frac{\varepsilon}{1-\alpha}.$$

Given any $\eta > 0$, such that $\Phi\Psi\Phi\Psi(\eta) < 1$, if $\text{wfs}_L(X) > \Phi\Psi\Phi\Psi(\eta)$, then

$$H_*(\widetilde{A}_X^L(\eta)) \cong \text{im } (H_*(B_{\widehat{X}}^{\widehat{L}}(\Psi(\eta))) \hookrightarrow B_{\widehat{X}}^{\widehat{L}}(\Psi\Phi\Psi(\eta))).$$

Proof. Given $\eta > 0$ such that $\Phi\Psi\Phi\Psi(\eta) < 1$, we have the following sequence of inclusions as a result of Lemma 10.

$$\widetilde{A}_X^L(\eta) \xhookrightarrow{\quad a \quad} B_{\widehat{X}}^{\widehat{L}}(\Psi(\eta)) \xhookrightarrow{\quad b \quad} \widetilde{A}_X^L(\Phi\Psi(\eta)) \xhookrightarrow{\quad c \quad} \ldots \tag{3}$$

$$\ldots \xhookrightarrow{\quad c \quad} B_{\widehat{X}}^{\widehat{L}}(\Psi\Phi\Psi(\eta)) \xhookrightarrow{\quad d \quad} \widetilde{A}_X^L(\Phi\Psi\Phi\Psi(\eta)).$$

As we assume that $\mathrm{wfs}_L(X) > \Phi\Psi\Phi\Psi(\eta)$, by the definition of weak feature size, Lemma 9 implies that the inclusions $b \circ a$ and $d \circ c$ are homotopy equivalences. We remind the reader that if two spaces are homotopy equivalent, all the induced homology maps between the spaces are isomorphisms. By applying homology to each space and inclusion in the previous sequence, we have the following sequence of homology groups, where $b_* \circ a_*$ and $d_* \circ c_*$ are isomorphisms.

$$H_*(\widetilde{A}_X^L(\eta)) \xrightarrow{\quad a_* \quad} H_*(B_{\widehat{X}}^{\widehat{L}}(\Psi(\eta))) \xrightarrow{\quad b_* \quad} H_*(\widetilde{A}_X^L(\Phi\Psi(\eta))) \xrightarrow{\quad c_* \quad} \tag{4}$$

$$H_*(B_{\widehat{X}}^{\widehat{L}}(\Psi\Phi\Psi(\eta))) \xrightarrow{\quad d_* \quad} H_*(\widetilde{A}_X^L(\Phi\Psi\Phi\Psi(\eta))).$$

The aforementioned isomorphisms $b_* \circ a_*$ and $d_* \circ c_*$ factor through $H_*(B_{\widehat{X}}^{\widehat{L}}(\Psi(\eta)))$ and $H_*(B_{\widehat{X}}^{\widehat{L}}(\Psi\Phi\Psi(\eta)))$ respectively, proving that b_* is surjective and c_* is injective. We then have that $H_*(\widetilde{A}_X^L(\eta)) \cong H_*(\widetilde{A}_X^L(\Phi\Psi(\eta))) \cong \mathrm{im}\, b_* \cong \mathrm{im}\, (c_* \circ b_*)$. $\square$

3.3.5. Computing the Homology

The last step is to relate the smoothed offsets to something that can be computed. It will generally be the case that the approximation $\widehat{X}$ of X is not just compact but also finite. Then, for any scale $\alpha \geq 0$, we have $B_{\widehat{X}}^{\widehat{L}}(\alpha)$ is the union of a finite set of Euclidean balls.

The Nerve Theorem provides a natural way to compute the homology of a union of Euclidean balls. The *nerve* of a collection U of sets is the sets of all subsets of U that have a nonempty intersection. It has the structure of a simplicial complex, whose homology can be directly computed by standard matrix reduction algorithms. When all nonempty intersections are contractible, the cover is said to be *good*. A cover by Euclidean balls (or any convex shape) is always good. For good covers, the Nerve Theorem, a standard result in algebraic topology [21] implies that

$$H_*(\mathrm{Nrv}(\{\mathrm{ball}(x, \alpha f_{\widehat{L}}(x))\}_{x \in \widehat{X}})) \cong H_*(B_{\widehat{X}}^{\widehat{L}}(\alpha)).$$

This is the most basic way to compute the homology of union of balls and is used throughout topological data analysis.

In our case, we are not just computing the homology of the union, but also the homology of the inclusion map. This computation will require a slightly stronger result. The Persistent Nerve Lemma [20], applies to Diagram 4 when combined with the above isomorphisms yields the following.

$$H_*(\widetilde{A}_X^L(\eta)) \cong \mathrm{im}\, H_*(\mathrm{Nrv}(\{\mathrm{ball}(x, \Psi(\eta)f_{\widehat{L}}(x))\}_{x \in \widehat{X}}) \hookrightarrow \mathrm{Nrv}(\{\mathrm{ball}(x, \Psi\Phi\Psi(\eta)f_{\widehat{L}}(x))\}_{x \in \widehat{X}})).$$

This last statement turns the isomorphism into an algorithm, because standard algorithms [22] can compute the homology of the inclusion of the nerves.

4. Conclusion

We have presented an alternative metric in Euclidean space that connects adaptive sampling and uniform sampling. We show how to apply classical results from the critical point theory of distance

functions to infer topological properties of the underlying space from such samples. This provides a connection between methods in surface reconstruction (based on adaptive sampling) and homology inference (based on uniform sampling).

We show in Theorem 1 that there is a precise relationship between samples that are uniformly taken with respect to d^L at some scale, to those same samples being adaptive in the Euclidean metric. In Theorem 2, we show that we can interleave the sublevel sets of our distance function under this alternative metric with the metric balls resulting from our approximation of the metric, assuming that both $\widehat{X}$ and $\widehat{L}$ are uniformly well-sampled with respect to the Hausdorff distance of d^L and $d^{\widehat{X}}$. Finally, we showed how to fully extend the critical point theory of distance functions and the weak feature size to give theoretical guarantees on homology inference from finite samples of X and L using the induced metric (Theorem 3).

The main limitation of adaptive metrics is that they require two sets as input, one to define the set and one to define the metric. In many instances, this is not available. However, we expect that the approach could find wider use in problems with labelled data. For example, data with binary labels may be viewed as the two sets X and L. Then, each set defines a metric on the other, where the metric is scaled according to how close it is to the other set. This is the subject of ongoing and future work.

Conflicts of Interest: Declare conflicts of interest or state "The authors declare no conflict of interest." Authors must identify and declare any personal circumstances or interest that may be perceived as inappropriately influencing the representation or interpretation of reported research results. Any role of the funders in the design of the study; in the collection, analyses or interpretation of data; in the writing of the manuscript, or in the decision to publish the results must be declared in this section. If there is no role, please state "The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results".

Appendix A. Details on Metric Smoothing

This section includes the full construction and relevant lemmas about the smoothed version of the induced metric.

For a compact set $L \subset \mathbb{R}^d$ and $\beta \geq 0$ denote by $L^\beta := \{x \in \mathbb{R}^d \mid \min_{y \in L} \|x - y\| \leq \beta\}$ the offsets of L with respect to the Euclidean metric. The following lemmas gives upper and lower bounds on the value of a smoothing of the distance-to-set function f_L, $\widetilde{f}_L$, which is defined on an arbitrarily smaller subset of Euclidean space.

Lemma A1. *Consider a compact set $L \subset \mathbb{R}^d$. Given $\alpha \in (0,1)$, for all $\beta \in (0,1)$, there exists smooth function $\widetilde{f}_L : \mathbb{R}^d \setminus L^\beta \to \mathbb{R}$ such that for all $x \in \mathbb{R}^d \setminus L^\beta$, $(1 - \alpha)f_L(x) < \widetilde{f}_L(x) < (1 + \alpha)f_L(x)$.*

Proof. By a standard result from [18], for all $\varepsilon > 0$, there exists a smoothing $\widetilde{f}_L : \mathbb{R}^d \setminus L^\beta \to \mathbb{R}$ of the distance function f_L such that $\|f_L - \widetilde{f}_L\|_\infty < \varepsilon$. Choose $\varepsilon = \beta\alpha$, for the given $\alpha \in (0,1)$. By the approximation property of $\widetilde{f}_L$, for all $x \in \mathbb{R}^d \setminus L^\beta$ we have that $f_L(x) - \varepsilon < \widetilde{f}_L(x) < f_L(x) + \varepsilon$. Also note that for all $x \in \mathbb{R}^d \setminus L^\beta$, $f_L(x) > \beta = \frac{\varepsilon}{\alpha}$ and thus $\alpha f_L(x) > \varepsilon$. Combining the aforementioned we have that $f_L(x)(1 - \alpha) < f_L(x) - \varepsilon$ and $f_L(x) + \varepsilon < f_L(x)(1 + \alpha)$. $\square$

Consider $\widetilde{f}_L$ as defined in Lemma A1. Using this we can define a smooth adaptive distance function $\widetilde{f_X^L}$ and provide upper and lower bounds on its value with respect to the original adaptive distance function f_X^L. For $x, y \in \mathbb{R}^d \setminus L^\beta$, we define

$$\widetilde{d^L}(x,y) := \inf_{\gamma \in \text{Path}(x,y)} \int_\gamma \frac{dz}{\widetilde{f}_L(z)}$$

and $\widetilde{f_X^L}(y) := \widetilde{d^L}(y, X)$.

Lemma A2. *Given $\alpha, \beta \in (0,1)$ and a smooth function $\widetilde{f_L}$ defined on $\mathbb{R}^d \setminus L^\beta$ approximating f_L, consider a compact set $X \subset \mathbb{R}^d \setminus L^\beta$. The Riemannian distance function $\widetilde{f_X^L}(\cdot) := \widetilde{d^L}(\cdot, X)$ satisfies the following property for all $y \in \mathbb{R}^d \setminus L^\beta$,*

$$\frac{1}{1+\alpha} f_X^L(y) < \widetilde{f_X^L}(y) < \frac{1}{1-\alpha} f_X^L(y).$$

Proof. Given two points $x, y \in \mathbb{R}^d \setminus L^\beta$, and any $\varepsilon > 0$, consider $\gamma, \gamma' \in \mathrm{Path}(x,y)$ such that $d^L(x,y) \le \int_\gamma \frac{dz}{f_L(z)} \le d^L(x,y) + \varepsilon$ and $\widetilde{d^L}(x,y) \le \int_{\gamma'} \frac{dz}{\widetilde{f_L}(z)} \le \widetilde{d^L}(x,y) + \varepsilon$. We then have the following inequalities resulting from inverting the inequalities in Lemma A1.

$$\widetilde{d^L}(x,y) \le \int_\gamma \frac{dz}{\widetilde{f_L}(z)} < \frac{1}{1-\alpha} \int_\gamma \frac{dz}{f_L(z)} \le \frac{1}{1-\alpha} d^L(x,y) + \frac{\varepsilon}{1-\alpha},$$

and

$$\frac{1}{1+\alpha} d^L(x,y) \le \frac{1}{1+\alpha} \int_{\gamma'} \frac{dz}{f_L(z)} < \int_{\gamma'} \frac{dz}{\widetilde{f_L}(z)} \le \widetilde{d^L}(x,y) + \varepsilon.$$

. Since these equalities hold for all $\varepsilon > 0$, then we can conclude that for all pairs $x, y \in \mathbb{R}^d \setminus L^\beta$, $\frac{1}{1+\alpha} d^L(x,y) < \widetilde{d^L}(x,y) < \frac{1}{1-\alpha} d^L(x,y)$.

Now consider $y \in \mathbb{R}^d \setminus L^\beta$. Define $x' := \mathrm{argmin}_{x \in X} d^L(y,x)$ and $x'' = \mathrm{argmin}_{x \in X} \widetilde{d^L}(y,x)$. We remind the reader that these points' existences are guaranteed by the Extreme Value Theorem. By examining these variables with respect to the previous inequality we know that

$$\frac{1}{1+\alpha} d^L(y,x') \le \frac{1}{1+\alpha} d^L(y,x'') < \widetilde{d^L}(y,x'') \le \widetilde{d^L}(y,x') < \frac{1}{1-\alpha} d^L(y,x').$$

By applying the definitions of both adaptive distance functions to the previous expression we obtain the desired inequality,

$$\frac{1}{1+\alpha} f_X^L(y) < \widetilde{f_X^L}(y) < \frac{1}{1-\alpha} f_X^L(y).$$

$\square$

Define the Riemannian adaptive offsets of X as $\widetilde{A_X^L}(\alpha) := \{x \in \mathbb{R}^d \mid \widetilde{f_X^L}(x) \le \alpha\}$, and denote the corresponding filtration by $\widetilde{\mathcal{A}_X^L}$. The following result reestablishes Lemma A2 in the language of filtrations and establishes an interleaving of the Riemannian adaptive offsets with the original adaptive offsets.

Corollary A1. *Let $L \subset \mathbb{R}^d$ be a compact set. Given $\alpha, \beta \in (0,1)$, for compact $X \subset \mathbb{R}^d \setminus L^\beta$, there exists a Riemannian distance function $\widetilde{f_X^L} : \mathbb{R}^d \to \mathbb{R}$, such that $(\widetilde{\mathcal{A}_X^L}, \mathcal{A}_X^L)$ are (h_6, h_7)-interleaved on $(0, \infty)$, where $h_6(r) = (1+\alpha)r$ and $h_7(r) = \frac{r}{1-\alpha}$.*

Proof. By Lemma A2, there exists a Riemannian distance function $\widetilde{f_X^L} : \mathbb{R}^d \to \mathbb{R}$, such that for all $y \in \mathbb{R}^d \setminus L^\beta$,

$$\frac{1}{1+\alpha} f_X^L(y) < \widetilde{f_X^L}(y) < \frac{1}{1-\alpha} f_X^L(y),$$

so for $r \in (0, \infty)$ and $y \in \widetilde{A_X^L}(r)$, $\widetilde{f_X^L}(y) \le r$, and thus $f_X^L(y) \le (1+\alpha)r$, which implies that $y \in A_X^L((1+\alpha)r)$, so $\widetilde{A_X^L}(r) \subseteq A_X^L((1+\alpha)r)$.

On the other hand, for $r \in (0, \infty)$ and $y \in A_X^L(r)$, $f_X^L(y) \le r$, and thus $\widetilde{f_X^L}(r) \le \frac{r}{1-\alpha}$, so $A_X^L(r) \subseteq \widetilde{A_X^L}(\frac{r}{1-\alpha})$. $\square$

Combining the previous corollary with Theorem 2 in Section 3.2.4, we obtain an interleaving between the smoothed adaptive offsets and the approximate offsets. This will then allow us to apply Lemma 9 and standard topological data analysis techniques to this interleaving to give a method

of homology inference for arbitrary small offsets of X as we have a Riemannian distance function generating the smooth adaptive offsets filtration.

Appendix A.1. Proof of Lemma 10

Proof. The hypotheses of the statement satisfy the hypotheses of both Theorem 2 and Corollary A1 so one knows that $(\mathcal{A}_X^L, \mathcal{B}_{\widehat{X}}^{\widehat{L}})$ are (h_4, h_5)-interleaved on $(0, 1)$, where $h_4(r) = \frac{r + \varepsilon}{(1 - r - \varepsilon)(1 - \delta)}$, and $h_5(r) = \frac{r}{1 - \delta - r} + \varepsilon$ Also, $(\widetilde{\mathcal{A}}_X^L, \mathcal{A}_X^L)$ are (h_6, h_7)-interleaved on $(0, \infty)$, where $h_6(r) = (1 + \alpha)r$ and $h_7(r) = \frac{r}{1 - \alpha}$. By applying Lemma 4 and composing the necessary functions, we achieve the stated interleavings. $\quad\square$

References

1. Cazals, F.; Giesen, J.; Pauly, M.; Zomorodian, A. The conformal alpha shape filtration. *The Visual Computer* **2006**, *22*, 531–540.
2. Chazal, F.; Lieutier, A. Smooth Manifold Reconstruction from Noisy and Non-Uniform Approximation with Guarantees. *Computational Geometry: Theory and Applications* **2008**, *40*, 156–170.
3. Chazal, F.; Lieutier, A. Topology Guaranteeing Manifold Reconstruction using Distance Function to Noisy Data. Proceedings of the 22nd ACM Symposium on Computational Geometry, 2006.
4. Dey, T.K.; Dong, Z.; Wang, Y. Parameter-free topology inference and sparsification for data on manifolds. Proceedings of the Twenty-Eighth Annual ACM-SIAM Symposium on Discrete Algorithms. Society for Industrial and Applied Mathematics, 2017, pp. 2733–2747.
5. Amenta, N.; Bern, M.; Kamvysselis, M. A new Voronoi-based surface reconstruction algorithm. SIGGRAPH, 1998, pp. 415–421.
6. Dey, T.K. *Curve and Surface Reconstruction : Algorithms with Mathematical Analysis*; Cambridge University Press, 2007.
7. Boissonnat, J.D.; Dyer, R.; Ghosh, A. Delaunay Triangulation of Manifolds. *Foundations of Computational Mathematics* **2018**, *18*, 399–431.
8. Boissonnat, J.D.; Wintraecken, M. The Topological Correctness of PL-Approximations of Isomanifolds. 36th International Symposium on Computational Geometry (SoCG 2020); Cabello, S.; Chen, D.Z., Eds.; Schloss Dagstuhl–Leibniz-Zentrum für Informatik: Dagstuhl, Germany, 2020; Vol. 164, *Leibniz International Proceedings in Informatics (LIPIcs)*, pp. 20:1–20:18. doi:10.4230/LIPIcs.SoCG.2020.20.
9. Niyogi, P.; Smale, S.; Weinberger, S. Finding the Homology of Submanifolds with High Confidence from Random Samples. *Discrete & Computational Geometry* **2008**, *39*, 419–441.
10. Niyogi, P.; Smale, S.; Weinberger, S. A Topological View of Unsupervised Learning from Noisy Data. *SIAM J. Comput.* **2011**, *40*, 646–663.
11. Grove, K. Critical point theory for distance functions. Proceedings of the Symposia in Pure Mathematics, 1993, Vol. 54.
12. Chazal, F.; Cohen-Steiner, D.; Lieutier, A. A Sampling Theory for Compact Sets in Euclidean Space. *Discrete & Computational Geometry* **2009**, *41*, 461–479.
13. Clarkson, K.L. Building triangulations using ε-nets. Proceedings of the thirty-eighth annual ACM symposium on Theory of computing, 2006, pp. 326–335.
14. Wein, R.; van den Berg, J.; Halperin, D. Planning High-quality Paths and Corridors Amidst Obstacles. *The International Journal of Robotics Research* **2008**, *27*, 1213–1231.
15. Agarwal, P.K.; Fox, K.; Salzman, O. An efficient algorithm for computing high quality paths amid polygonal obstacles. Proceedings of the 27th Annual ACM-SIAM Symposium on Discrete Algorithms, 2016, pp. 1179–1192.
16. Cohen, M.B.; Fasy, B.T.; Miller, G.L.; Nayyeri, A.; Sheehy, D.R.; Velingker, A. Approximating Nearest Neighbor Distances. Proceedings of the Algorithms and Data Structures Symposium, 2015.
17. Chu, T.; Miller, G.L.; Sheehy, D.R. Exact computation of a manifold metric via Lipschitz Embeddings and Shortest Paths on a Graph. Proceedings of the ACM-SIAM Symposium on Discrete Algorithms, 2020.
18. Green, R.; Wu, H. C^∞ approximations of convex, subharmonic, and plurisubharmonic functions. *Ann. Sci. École Norm. Sup.* **1979**, *12*, 47–84.

19. Chazal, F.; Lieutier, A. Weak Feature Size and Persistent Homology: Computing Homology of Solids in R^n from Noisy Data Samples. Proceedings of the 21st ACM Symposium on Computational Geometry, 2005, pp. 255–262.
20. Chazal, F.; Oudot, S.Y. Towards Persistence-Based Reconstruction in Euclidean Spaces. Proceedings of the 24th ACM Symposium on Computational Geometry, 2008, pp. 232–241.
21. Hatcher, A. *Algebraic Topology*; Cambridge University Press, 2001.
22. Edelsbrunner, H.; Letscher, D.; Zomorodian, A. Topological Persistence and Simplification. *Discrete & Computational Geometry* **2002**, *4*, 511–533.