

4-26-2019

Methods in Homology Inference

Nicholas J. Cavanna

University of Connecticut - Storrs, nicholas.j.cavanna@gmail.com

Follow this and additional works at: <https://opencommons.uconn.edu/dissertations>

Recommended Citation

Cavanna, Nicholas J., "Methods in Homology Inference" (2019). *Doctoral Dissertations*. 2118.
<https://opencommons.uconn.edu/dissertations/2118>

Methods in Homology Inference

Nicholas Cavanna, Ph.D.

University of Connecticut, 2019

In this dissertation we introduce novel techniques to infer the shape of a geometric space from local data. The major contribution of this research is that the techniques aim to minimize the hypotheses of the sample and the space compared to existing results in the field.

We first prove an equivalence between the two standard geometric sampling methods: adaptive and uniform. From this we create a collection of topological interleavings derived from this equivalence, which we use to compute the homology of an arbitrary Euclidean compact set only assuming a bound on the weak feature size of the domain and a sample of an arbitrary reference set for sampling.

Next we provide an algorithm to check for k -coverage of sensor network from a collection of finite sensors using computable structures built upon from sensors' local data. This algorithm generalizes the original Topological Coverage Criterion by De Silva and Ghrist to domains with non-smooth boundaries among other relaxed assumptions. We then prove that if one has coverage of the sensor network, there exists subsampling parameters such that one can infer the homology of the entire domain.

We culminate the dissertation by proving an approximation of the Persistent Nerve Theorem assuming a relaxation of the standard topological cover assumption. Namely, given such a cover exists, there is a tight bottleneck distance between the persistence diagrams of the nerve filtration and the space filtration, generalizing the Persistent Nerve Theorem. This results provides information about the shape of a subdivisions of triangulations with no global geometric data provided, as well as broadens the applications of nerves to cases there are not nice sampling guarantees or there exists approximation errors during triangulation.

Methods in Homology Inference

Nicholas J. Cavanna

B.S., University of Connecticut, 2014

A Dissertation

Submitted in Partial Fulfillment of the

Requirements for the Degree of

Doctor of Philosophy

at the

University of Connecticut

2019

Copyright by

Nicholas J. Cavanna

2019

APPROVAL PAGE

Doctor of Philosophy Dissertation

Methods in Homology Inference

Presented by

Nicholas J. Cavanna, B.S.

Major Advisor _____

Donald R. Sheehy

Associate Advisor _____

Parasara Sridhar Duggirala

Associate Advisor _____

Thomas J. Peters

Associate Advisor _____

Alexander Russell

University of Connecticut

2019

Contents

1	Introduction	1
2	Mathematical Preliminaries	6
2.1	Geometry	6
2.1.1	Metric Spaces	6
2.1.2	Geometric Sampling	9
2.1.3	Simplicial Complexes	12
2.2	Topology	16
2.2.1	Homotopy and the Nerve Theorem	16
2.2.2	Chain Complexes and Homology	19
3	Unifying Adaptive and Uniform Sampling towards Homology Inference	24
3.1	Introduction	24
3.2	Definitions	26
3.3	The Equivalence of Adaptive and Uniform Sampling	32
3.4	Filtrations and Interleavings	34

3.4.1	Approximating X with $\widehat{X}$	36
3.4.2	Approximating the Adaptive Offsets	36
3.4.3	Approximating L with $\widehat{L}$	37
3.4.4	Composing the Interleavings	38
3.5	Smoothing the Metric and Homology Inference	39
4	Homological Sensor Network Coverage and Inference	45
4.1	Introduction	45
4.1.1	Related Work	47
4.2	Definitions	48
4.2.1	K-Nearest Neighbor and Coverage Regions	48
4.2.2	The k -Barycentric Decomposition and Clique Complexes	50
4.2.3	Relative (Co)Homology and Duality	51
4.3	The Geometric Topological Coverage Criterion	55
4.4	The Algorithmic Topological Coverage Criterion	61
4.5	Certified Homology Inference	66
5	The Generalized Persistent Nerve Theorem	70
5.1	Introduction	70
5.1.1	Related Work	72
5.2	Definitions	73
5.2.1	CW-Complexes	73

5.2.2	Chain Homotopy and Cellular Homology	74
5.2.3	Persistent Homology	76
5.2.4	Nerve and Cover Filtrations	78
5.3	ε -Good Covers	79
5.4	Building the Interleaving	82
5.4.1	Homotopy Colimits	83
5.4.2	Nerve Diagrams	83
5.4.3	The Blowup Filtration	85
5.4.4	The Fundamental Chain Maps	87
5.4.5	Chain Map Lifting	93
5.5	The Main Theorem	99
5.5.1	The Bound is Tight	100

Chapter 1

Introduction

High-dimensional data analysis techniques are increasingly necessary in academic and industrial settings such as statistics, machine learning, genetics, and engineering disciplines. This data can be realized geometrically in Euclidean space or another metric space. For example, given a collection of data where each object of interest has k real-valued characteristics each object can be represented as a k -dimensional point; collectively these points form a geometric space. Similarly, when analyzing a geometric space or surface, e.g. tumor tissue, one takes a sample and the sampled space's shape must be reconstructed from only finitely many points. Topological data analysis, at the intersection of computational geometry and algebraic topology, has arisen as an approach to interpreting and categorizing this geometric data using tools from topology.

Topology is the study of the properties of spaces that are invariant under continuous bijective transformations of the space. The classic example of a topological equivalence is the solid torus and a coffee mug. Both of these can be continuously transformed into solid spheres with a handle attached. The contrapositive of this tells us that two geometric spaces can be distinguished topologically by showing that they do not share one of these topological invariants. In turn, one can consider a more manageable representation of a geometric space, compute the desired topological properties, and these are mirrored by the original space.

In this dissertation, we focus on a topological property that encodes coarse, but sufficiently differentiating topological information of spaces. The algebraic topology construction we will compute is homology — loosely speaking, a measure of the number of holes of various dimensions of a space. Although in the algebraic topology literature, the notion of “holes” often must be interpreted abstractly, in sufficiently nice geometric settings the formal definition aligns with what one would visualize. Homology’s strength is in the fact is invariant under bijective continuous changes to the space, so for example it the homology of a space and a triangulation of it are identical, and it is efficiently computable via linear algebra, in contrast to other topological invariants.

Given some finite geometric sample, we want to know from what is the homology of the underlying space from which it was sampled,. The finiteness of the sMPLW gives one access to computational methods to infer the homology of the sampled space. By considering metric balls of various radii around the sampled points we have a representation of the space to determine the homology of the underlying space. One way to compute the shape of the metric balls is to compute their intersection graph, with edges between the points centered at the balls if they intersect. From this graph comes a higher-dimensional discrete representation, a simplicial complex, which under some conditions has the same shape as the metric balls. As mentioned, the homology of such spaces is efficiently computable so this derived structure is a natural choice to represent the shape of the space. Figure 1.1 depicts this construction over a series of radii. This process shows how one goes from local metric information to global topological information, similar to how in principal component analysis and manifold learning particular data points are extracted as representative of the fundamental characteristics of the whole data set . This pipeline is known as homology inference and will be the central theme of this dissertation. We focus on weakening sampling hypotheses and, more broadly, generalizing existing methods for computing the homology of geometric spaces from local information.

The first problem we solve is unifying two standard geometric sampling methods, adaptive and uniform sampling. Adaptive sampling is where one samples more densely on prominent geometric

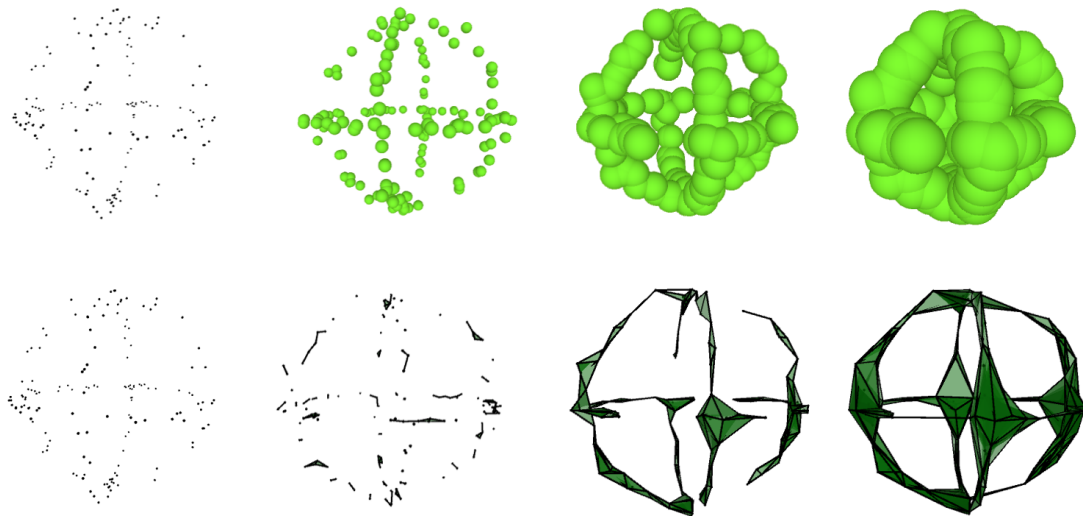


Figure 1.1: The simplicial complexes corresponding to a growing collection of metric balls around a point sample reveal the space from which they were covered — three great circles on a sphere.

features of a space while uniform sampling is where each point of the space is sufficiently close to some part of the sample. Adaptive sampling was introduced by Amenta and Bern [4] as a means to perform manifold reconstruction, in which the goal is to reconstruct a low-dimensional manifold via a triangulation. Uniform sampling is the standard assumption in many geometric/homology inference papers, e.g. in [11, 15], due to the global scaling of the corresponding metric balls. We prove that these sampling methods are equivalent when done with respect to two different metrics, while making minimal assumptions about the space and the reference set from which we view geometric features as prominent. Using this equivalency we are able to compute the homology of the domain by only looking at the homology of Euclidean metric balls derived from sufficiently close finite samples of the space and the reference set. Though the sampling parameters are complicated and narrow, the result is novel to its generality and the limited assumptions of the samples.

Next, we broaden the existing literature on verifying coverage in homological sensor networks. Sensor networks, which we model geometrically as collections of points in a space, are collections of sensors that can check local data and possibly communicate related information to those close

by, are prevalent in engineering, e.g. motion detectors, light sensors and robotic systems. In most cases the sensors should collectively cover the region in which they lie in order for them to gather all information for which they are designed. This is known as sensor network coverage. Of course, if the locations of all the sensors are known as well as the area in which each can detect, it is trivial to check for coverage of the domain.

In the setting of homological sensor networks, the location of the sensors in the domain is unknown, but some local information is available. Sensors are able to detect other sensor within some distance, as well as the domain itself, and the boundary of the domain. The pre-existing state of the art in this field is due to De Silva and Ghrist [22] who proved a sufficient condition for sensor networks in a smoothly bounded Euclidean domain — a result known as the Topological Coverage Criterion. We extend this result by providing a discrete algorithm that sufficiently checks for k -wise coverage by sensors of more general metric space, while also not required the boundary to be smooth. Furthermore, we don't require the outer region the sensors detect to be the domain's boundary, but just a set that surrounds the interior of the domain. In the process, we streamline previous results by separating the topological from the geometric. Next, we address the problem of scale selection — the question of at which scales to analyze the sample to perform homology inference. We create subsampling parameters of the sensors which we use to compute the homology of the domain given it is covered by computing the homology of two simplicial complexes built on the sensors. This establishes an algorithmic pipeline for verifying a sample is sufficient to perform homology inference, and then computing the homology if it indeed is.

The last problem we address is that of approximating parametrized homology, formally known as persistent homology. In topological data analysis, data is often examined at a series of growing radius around the points. This yields an increasing sequence of spaces, which lacks a “true” homology in the traditional sense. Instead, persistent homology arose as a way to quantify all the homological data of these sequences as scales change. So-called persistence diagrams that contain these changes in persistent homology have a proper distance metric between them which

measures the topological similarity of two growing spaces. This provides a novel means of objectively comparing two data sets. Note that there are methods to compute the persistent homology precisely, but in practice they are restrictive as they require deep knowledge of the sequence beforehand.

Our goal is to approximate the (persistent) homology of a growing sequence of triangulations by constructing a finite representation. Given geometric point data, introducing each of the points at a particular time and triangulating the resulting point set yields a sequence of growing triangulations. The Persistent Nerve Lemma [17] states that the topology of a growing sequence of spaces covered by a nice collection of sets can be identical to that of the simplicial complexes built from the intersections of the sets. Although this is a strong theoretical result, it requires the cover's sets have contractible intersections. In practice minor computation errors occur when computing triangulations make application of this theorem impossible, as well as when the metric balls of a space are not convex.

To rectify this shortcoming in the theory, we introduce a parametrized version of the contractibility assumption in order to provide a tight bound on the metric between the persistence diagrams of a sequence of simplicial complexes resulting from the covers to the sequence of triangulations itself. Though this result is aimed towards applications in topological data analysis, the constructions in the proof are of theoretical interest to the community as well. We compute at the level of the chain complexes of these sequence of spaces, rather than at the level of the homology of each of the spaces or the persistent homology in and of itself, which leads to new insights on how the structure of the nerves and the spaces are related.

Chapter 2

Mathematical Preliminaries

2.1 Geometry

2.1.1 Metric Spaces

The primary topological spaces we will consider are **metric spaces**. These are sets with a function defined on them that generalizes the traditional notion of the distance between two points.

Definition 2.1. *A metric space (X, d) consists of a set X and a non-negative real-valued function called a metric, $d : X \times X \rightarrow \mathbb{R}_{\geq 0}$ with the following properties for all x, y and $z \in X$.*

- $d(x, y) \geq 0$
- $d(x, y) = 0$ if and only if $x = y$
- $d(x, y) = d(y, x)$
- $d(x, z) \leq d(x, y) + d(y, z)$

When it is clear with what metric a space is equipped, we will simply write X rather than (X, d) . Any non-empty subset $X' \subset X$ of a metric space is a metric space as well, with the metric

$d_{X' \times X'} := d|_{X' \times X'}$. The space of all metric spaces may be partitioned into equivalence classes of *isometric* metric spaces. An **isometry** between two metric spaces X and Y is a function $f : X \rightarrow Y$ such that for all $x, y \in X$, $d_Y(f(x), f(y)) = d_X(x, y)$. Isometries are injective continuous functions with images homeomorphic to their domains. The relaxation of isometries are Lipschitz functions, which are also by definition continuous.

Definition 2.2. A function $f : X \rightarrow Y$ between metric spaces (X, d_X) and (Y, d_Y) is called *K-Lipschitz*, for $K > 0$, if $d_Y(f(x), f(x')) \leq K d_X(x, x')$ for all $x, x' \in X$.

Following are some examples metrics. The Euclidean distance between two points $x = (x_1, \dots, x_n)$ and $y = (y_1, \dots, y_n)$ in $\mathbb{R}^n$ is

$$d_2(x, y) := ((x_1 - y_1)^2 + \dots + (x_n - y_n)^2)^{\frac{1}{2}}.$$

This is also known as the 2-norm distance, or the *straight-line* distance. This metric generalizes to a family of metrics on $\mathbb{R}^n$ known as the *p-norm metrics* or *ℓ_p metrics*, which are defined for $p \geq 1$ on two points by

$$d_p(x, y) := (|x_1 - y_1|^p + \dots + |x_n - y_n|^p)^{\frac{1}{p}}.$$

Further generalizing one gets the supremum metric, $d_\infty := \lim_{p \rightarrow \infty} d_p$. An example of a n -Lipschitz function is the identity function $(\mathbb{R}^n, d_\infty) \rightarrow (\mathbb{R}^n, d_1)$ as $d_1(x, y) = \sum_i |x_i - y_i| \leq n \max_i |x_i - y_i| = n d_\infty(x, y)$.

A pseudo-metric d has all the same properties as a metric except it is not required that $d(x, y) = 0$ implies $x = y$, i.e. the points are not necessarily distinguished by the pseudo-metric. By identifying these undistinguishable points with respect to the pseudo-metric, a metric space can be constructed via quotienting by the resulting equivalence relation.

Given a metric space (X, d) and a point $x \in X$, the **metric ball** of x with radius α is

$$\text{ball}(x, \alpha) := \{y \in X \mid d(x, y) \leq \alpha\}.$$

As an example, the 2-norm metric ball around a point in $\mathbb{R}^n$ is a geometric $(n - 1)$ -sphere of radius α .

A subset of a metric space is **compact** if it is totally bounded, meaning the subset can be covered by finitely many open balls of any given radius, and is complete, i.e. every sequence of points has a limit point within the subset.

Given a metric space and a subspace the distance-to-set function which computes the distance of a point to the subspace.

Definition 2.3. For a non-empty subset A of a metric space X , the **distance function** with respect to A is $f_A : X \rightarrow \mathbb{R}_{\geq 0}$ and defined as follows for $x \in X$,

$$f_A(x) := d(x, A) := \inf\{d(x, a) \mid a \in A\}.$$

Alternatively, we may express the distance function to A as $d(\cdot, A) := f_A(\cdot)$. Note that if A is compact, e.g. finite subsets of metric spaces, then by the Extreme Value Theorem there is an $a' \in A$ such that $d(x, A) = d(x, a')$. This distance-to-set function is notably 1-Lipschitz, and thus a continuous function from the metric space X to $\mathbb{R}$.

Proposition 2.4. For a metric space X and a non-empty subset A , the distance function f_A is 1-Lipschitz to $\mathbb{R}$.

Proof. Given x and $y \in X$, for all $a \in A$, we have that $f_A(x) \leq d(x, a) \leq d(x, y) + d(y, a)$, so $f_A(x) - d(x, y) \leq d(y, a)$ for all $a \in A$. This implies that $f_A(x) - d(x, y) \leq f_A(y)$ by definition of f_A . By symmetry, we have that $f_A(y) - d(x, y) \leq f_A(x)$, so $|f_A(x) - f_A(y)| \leq d(x, y)$. $\square$

Given a distance-to-set function f_A , the pre-images of the intervals $[0, \alpha]$ filter the metric space around the set A . Each of these preimages is known as an offset of A .

Definition 2.5. For any subset A of a metric space (X, d) , the (closed) α -**offsets** with respect to d are

$$A^\alpha := \{x \in X \mid f_A(x) \leq \alpha\} = f_A^{-1}[0, \alpha].$$

An alternative way to define the offsets is the equality $A^\alpha = \bigcup_{a \in A} \text{ball}(a, \alpha)$.

The offsets and the distance function itself can be used to define the Hausdorff distance between two compact subsets A and B of a metric space with respect to the metric topology. Hausdorff distance measures the global metric similarity of two subsets. This is a proper metric on the space of compact subspaces of a metric space.

Definition 2.6. For two compact subsets A and B of a metric space (X, d) , the **Hausdorff distance** d_H between them is defined as follows.

$$d_H(A, B) := \max\left\{\sup_{a \in A} d_B(a), \sup_{b \in B} d_A(b)\right\} = \inf\{\varepsilon \geq 0 \mid A \subseteq B^\varepsilon \text{ and } B \subseteq A^\varepsilon\}$$

For example consider the two images of closed disks centered at the origin of Euclidean plane D_r and $D_{r'}$ with radii r and r' respectively, where $r < r'$. The Hausdorff distance $d_H(D_r, D_{r'}) = r' - r$.

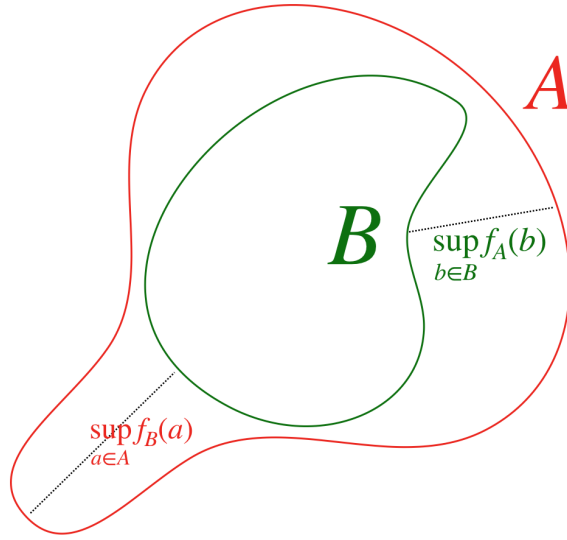


Figure 2.1: Two planar curves and the maximum of the distance-to-set functions between them.

2.1.2 Geometric Sampling

The Hausdorff distance between a sample and a set is used as a measure of how uniformly the sample is distributed on the space. Explicitly, given a sample point set $P \subset X$, if for all $x \in X$, there exists a close p , i.e. for some $\varepsilon > 0$, there is $p \in P$ such that $d(x, p) \leq \varepsilon$, then $X \subseteq P^\varepsilon$, so $d_H(X, P) \leq \varepsilon$,

and vice versa. While Hausdorff distance defines a good sample at a global level, it says little about whether the sample approximates regions of the space dense in geometric features. The following definition is a way to represent how well a sample models this local geometric information. Given a metric space X and a subset L , then a set P is an ε -**sample** of X with respect to L if for all $x \in X$ there exists $p \in P$ such that $d(x, p) \leq \varepsilon d(x, L) = \varepsilon f_L(x)$. This type of sample is also known as an adaptive sample, named so because the requirements of the sample P with respect to X are variable based on that points closeness to the reference set L in question.

Now consider a compact space A , for example a manifold, in a metric space X . Its **medial axis** $\mathcal{M}$ is the collection of points in X that have at least two closest points in A with respect to the underlying metric. As an example, the medial axis of a circle centered in the plane is the circle's center point. The medial axis of a set provides a skeleton that carries geometric information about the set.

Given a closed set A with medial axis $\mathcal{M}$, the **local-feature size** at a point $x \in S$ is defined as $\text{lfs}(x) := f_{\mathcal{M}}(x)$. Local-feature size is a measure of the size of the space near a point — as the size of the geometric feature increases, the local-feature size decreases, so points with low local-feature size are regions where denser sampling is needed to approximate the shape. This was introduced in a series of papers by Amenta and Bern [3, 4, 5] as a sampling condition for performing manifold reconstruction. They also introduced ε -sampling with respect to the medial axis, as defined above, except with $L = \mathcal{M}$. By its definition, the medial axis was the natural choice for the reference set as it is the points in the space that are closest to the set locally. As the medial axis cannot always be computed, in practice there is a need to sample with respect to a close approximation of or a finite sample of it, leading to the generic definition we choose for adaptive sampling within this dissertation.

Although given an arbitrary compact set A , the distance function d_A is not necessarily smooth or even continuous, there still exists a sound definition of its critical point. Given a compact set $A \subset \mathbb{R}^d$, define the set Γ as the set of closest points of A to x , i.e. $\Gamma_A(x) := \{a \in A \mid d(x, a) = f_A(x)\}$.

Note that if $x \in A$, then $\Gamma_A(x) = \{x\}$ and in general, the set is non-empty and compact. Denote the center of the minimal enclosing ball of $\Gamma_A(x)$ by $\theta_A(x)$. The gradient of the distance function d_A is defined as the following, for points $x \in \mathbb{R}^d \setminus A$,

$$\nabla_A(x) := \frac{x - \theta_A(x)}{f_A(x)},$$

and $\nabla_A(x) = 0$ for all $x \in A$. The **critical points** of the distance function d_A are the points $x \in \mathbb{R}^d$ such that $\nabla_A(x) = 0$. Equivalently, $x \in \mathbb{R}^d$ is a critical point of f_A if and only if x lies on the convex hull of $\Gamma_A(x)$. The **critical value** associated to a critical point x^* is $f_A(x^*)$.

A fundamental result in critical point theory of distance functions [15, 33] is the following theorem.

Theorem 2.7. *Let $A \subset \mathbb{R}^d$ be a compact set with distance function d_A . If an interval $[\alpha, \beta]$ contains no critical values of d_A , then the inclusion $A^\alpha \hookrightarrow A^\beta$ is a homotopy equivalence.*

The homotopy equivalence above roughly means that path-based topological properties of the offsets are preserved by the inclusion map in these intervals. The theorem allows for a decomposition of a set's offsets into collections of topologically equivalent spaces over the intervals where there are no critical points present. Homotopy will be explained rigorously in Subsection 2.2.1.

Weak-feature size is a condition on compact sets that was introduced in [11, 15] as a global version of local feature size from adaptive sampling. Given a compact set A , consider the set C of critical points of f_A on $\mathbb{R}^d \setminus A$. Note that this set of critical points is a subset of the medial axis. The **weak-feature size** of A is defined as $\text{wfs}(A) = \min_{x \in C} f_C(x)$, or alternatively, the infimum of the positive critical values of f_A . Weak-feature size coupled with Theorem 2.7 gives an assumption for preserving the topology of a space when it is grown by some distance around its points.

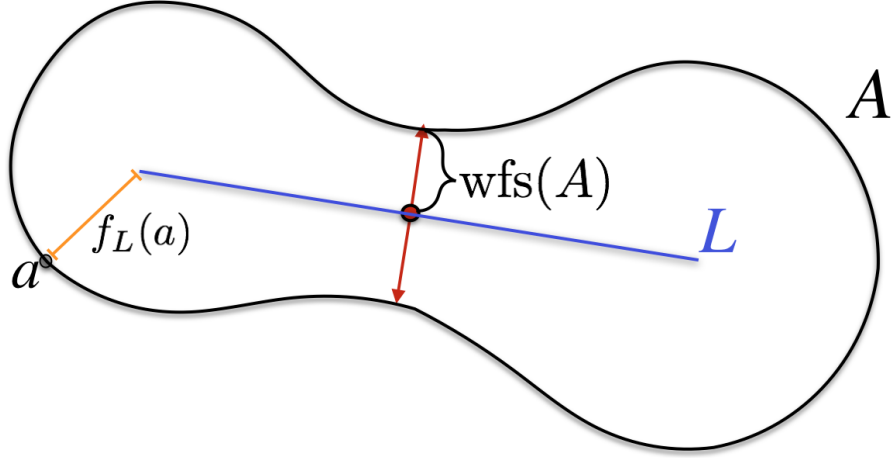


Figure 2.2: A planar curve A and its weak-feature size, where L is the curve's medial axis.

2.1.3 Simplicial Complexes

A **geometric simplex** is the convex hull of affinely independent points in $\mathbb{R}^d$, also known as polyhedrons. The convex hull of k points corresponds to a $(k - 1)$ -simplex. A simplex's faces are its subsimplices. A **geometric simplicial complex** is the union of geometric simplices over the same collection of vertices, which can also be viewed as gluing simplices along their faces, or alternatively, requiring that the intersection of two simplices is also a simplex.

Simplicial complexes have combinatorial counterparts. An **abstract simplicial complex** K over a vertex set V is a non-empty subset of the powerset $\mathcal{P}(V)$ such that for each $v \in V$, $\{v\} \in K$ and if $\sigma \in K$, then for all $\tau \subset \sigma$, $\tau \in K$. A k -simplex $\sigma \in K$ is a simplex such that $k = \#(\sigma) - 1$. A simplex is called maximal if no other simplex properly contains it. The dimension of an abstract simplicial complex K , $\dim K$, is the maximum dimension over all its simplices. Given an abstract simplicial complex K , $K^{(n)}$ denotes the n -**skeleton**, the sub-complex of K consisting of all simplices of dimension less than or equal to n . With this notation, we have that $V = K^{(0)} \subseteq K^{(1)} \dots \subseteq K^{(\dim K)} = K$. For example, the 1-skeleton is a graph, and the 1-skeleton of

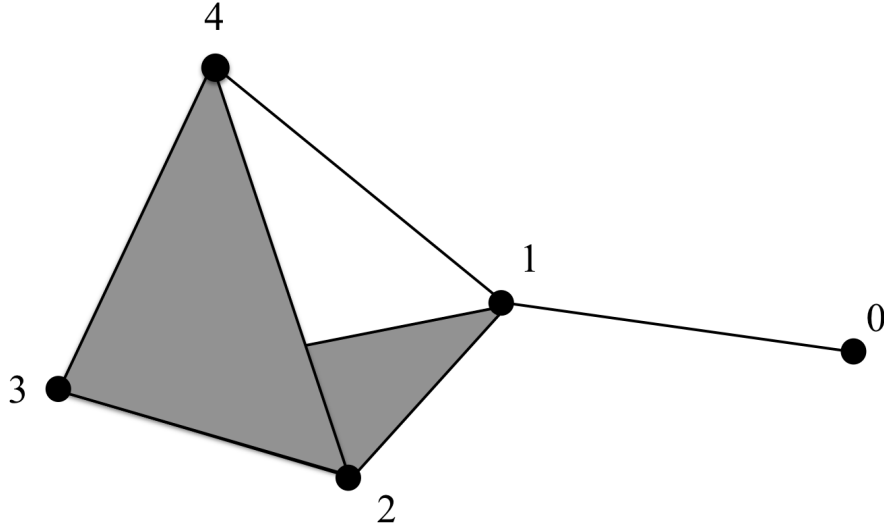


Figure 2.3: A geometric simplicial complex in $\mathbb{R}^3$ defined over the vertex set $V = \{0, \dots, 4\}$.

a geometric simplicial complex is a geometric graph.

A **simplicial map** between abstract simplicial complexes K and L , defined over vertex sets V_K and V_L respectively, is a function $\phi : V_K \rightarrow V_L$ such that if $\sigma = \{v_0, \dots, v_n\}$ is a simplex of K , then $\phi(\sigma) = \{\phi(v_0), \dots, \phi(v_n)\}$ is a simplex of L . The image $\phi(K)$ is a subcomplex of L . A simplicial map between geometric simplicial complexes is similarly defined — it is the linear extension of a function between the vertices.

A finite abstract simplicial complex K has a canonical geometric simplicial complex called the **geometric realization**, $|K| \subset \mathbb{R}^{n+1}$, where each vertex $v_i \in V = \{v_0, \dots, v_n\}$ gets mapped to $e_{i+1} = (0, \dots, 1, \dots, 0)$. For each k -simplex σ of K , where $\sigma = \{u_0, \dots, u_k\}$, the standard k -simplex is

$$|\sigma| = \left\{ \sum_{i=0}^k t_i |u_i| : t_i \geq 0, \sum_{i=0}^k t_i = 1 \right\},$$

i.e. the convex hull of the realization of σ 's vertices. The standard k -simplex is often written as Δ^k . The geometric realization is equipped with the subspace Euclidean topology. We will use the geometric realization of an abstract simplicial complex without mention of the coordinates of the

particular embedding of the vertices.

Given a simplicial map between two simplicial complexes K and L induced by a map ϕ on their vertex sets, there is a realization $|\phi|$ that maps a vertex $|v| \in |K|$ to $\phi(|v|) := |\phi(v)|$. This extends to the simplices in the natural way — namely a geometric simplex $|\sigma|$ gets mapped to $\phi(|\sigma|) := |\phi(\sigma)|$. This commutativity of simplicial maps and the geometric realization operator implies that the geometric realization $|\cdot|$ is an example of a functor (see MacLane [39]).

Given a finite metric space P , one can build two standard families of simplicial complexes, the Vietoris-Rips complexes and the Čech complexes. Although the following definitions define the aforementioned complexes abstractly, when P is a subset of Euclidean space one can consider them as the corresponding geometric simplicial complex with vertex set P . The Vietoris-Rips complex was defined originally by Leopold Vietoris as a means to extend homology theory from simplicial complexes to metric spaces, a connection which will become more evident after the introduction of the nerve of a cover, Definition 2.11.

Definition 2.8. *The **Vietoris-Rips complex** of a finite sample P in a metric space (X, d) at some scale $\alpha \geq 0$ is the following abstract simplicial complex*

$$\mathcal{R}^\alpha(P) := \{\sigma \subseteq P \mid \text{ball}(p, \alpha) \cap \text{ball}(q, \alpha) \neq \emptyset, \forall p, q \in \sigma\}.$$

Definition 2.9. *The **Čech complex** of a finite point set P in a metric space (X, d) scale $\alpha \geq 0$ is the following abstract simplicial complex*

$$\mathcal{C}^\alpha(P) := \{\sigma \subseteq P \mid \bigcap_{p \in \sigma} \text{ball}(p, \alpha) \neq \emptyset\}.$$

By definition, at a fixed scale α and finite point set P , the Čech complex is contained within the Rips complex at that same scale, i.e. $\mathcal{C}^\alpha(P) \subseteq \mathcal{R}^\alpha(P)$. For an arbitrary metric space X and finite point set $P \subset X$, the two definitions are related by the following containments.

Lemma 2.10. *Given a non-empty finite point set P in a metric space X , the following holds for all $\alpha \geq 0$,*

$$\mathcal{C}^\alpha(P) \subseteq \mathcal{R}^\alpha(P) \subseteq \mathcal{C}^{2\alpha}(P). \quad (2.1)$$

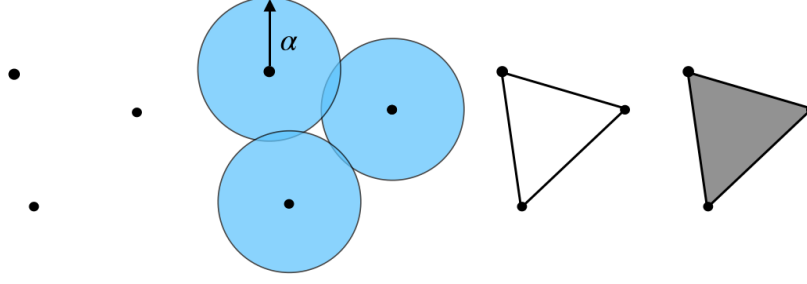


Figure 2.4: A point set P and the associated offsets P^α , and complexes $\mathcal{C}^\alpha(P)$ and $\mathcal{R}^\alpha(P)$ for a scale α , from left to right.

Proof. The first containment is true by definition. The second containment holds true because for $\sigma \in \mathcal{R}^\alpha(P)$, then for each $p, q \in \sigma$, there exists $x \in \text{ball}(p, \alpha) \cap \text{ball}(q, \alpha)$, so $d(p, q) \leq d(p, x) + d(q, x) \leq 2\alpha$, so without loss of generality $p \in \text{ball}(q, 2\alpha)$ and thus $\sigma \in \mathcal{C}^{2\alpha}(P)$. $\square$

Lemma 2.10 holds true for any metric space, but in Euclidean space the containment parameters are tighter. Explicitly, for a finite point set P in $\mathbb{R}^d$ with the Euclidean metric, Jung's Theorem [37] implies that

$$\mathcal{C}^\alpha(P) \subseteq \mathcal{R}^\alpha(P) \subseteq \mathcal{C}^{\vartheta_d \alpha}(P), \quad (2.2)$$

where $\vartheta_d = \sqrt{\frac{2d}{d+1}}$. It is easy to see there is a generic upper-bound of $\sqrt{2}$ for all dimensions of Euclidean space, by considering $\lim_{d \rightarrow \infty} \vartheta_d$.

The geometric realization of the Čech complex gives a more comprehensive summary of the geometry of the α -offsets of a point set than the Rips complex, but it is more computationally intensive to compute whether k metric balls intersect in Euclidean space, and in more generic metric spaces it is impossible. In contrast, the Rips complex at a scale α requires computing for each pair of points $p, q \in P$ if $d(p, q) \leq 2\alpha$. The sandwiching of the Čech complex by Rips complexes provides an approximated inference of the geometric structure of the Čech complex by computing the Rips complexes at two scales.

2.2 Topology

We preface this section by stating that although the topological notions provided are collectively sufficient to follow the results in this dissertation, more comprehensive coverage of the fundamentals of algebraic topology can be found in Hatcher's Algebraic Topology [35] and Munkres' Elements of Algebraic Topology [41].

2.2.1 Homotopy and the Nerve Theorem

The primary setting of this work is metric spaces, and accordingly we will with the metric topology for the most part. The metric topology has as a basis the open metric balls of all positive radii centered at the points in the metric space. When we say a metric space, it is implicit that is being considered as a topological space equipped with the metric topology unless otherwise stated. Any subspace of a topological space is also assumed to be equipped with the subspace topology which consists of the open sets of that space intersected with the subspace.

Given a set X , a collection of subsets τ of X is a **topology** on X if $\emptyset$ and $X \in \tau$, τ is closed under arbitrary union of its elements, and τ is closed under finite intersections of its elements. If τ forms a topology, then its elements are known as *open* sets, while their complements are called closed sets. Note that a topology can be defined from closed sets, by swapping the cardinality considerations of unions and intersections. A set along with its topology is called a topological space. For a metric space, there is always a topology generated by the metric balls in the space called the metric topology. We will not mention specifics about the topology chosen in this work as they will always be the metric topology, or the Euclidean topology, which is the metric topology on $\mathbb{R}^d$.

Given two topological spaces X and Y , a function $f : X \rightarrow Y$ is called *continuous* if for any open set $U \subseteq Y$, its preimage $f^{-1}(U)$ is open in X . For metric spaces, this is equivalent to the standard $\varepsilon - \delta$ definition of continuity used in the field of analysis. Another term we will use for a

continuous function is a map.

A **homeomorphism** between two topological spaces X and Y is a bijective function $f : X \rightarrow Y$ such that f and f^{-1} are continuous, and two spaces with a homeomorphism between them are called homeomorphic. This is an equivalence relation. For example, consider the cube $X := [-1, 1] \times [-1, 1] \times [-1, 1] \subset \mathbb{R}^3$ and the unit sphere $\mathbb{S}^2$ centered at the origin – these are homeomorphic via the continuous function $f : X \rightarrow \mathbb{S}^2$ where $f(\mathbf{x}) := \frac{\mathbf{x}}{\|\mathbf{x}\|}$. A **triangulation** of a topological space X is a (geometric) simplicial complex K and a homeomorphism $h : X \rightarrow K$. Triangulable spaces are those with a triangulation. The geometric 2-simplex is a triangulation of $\mathbb{S}^2$ with homeomorphism being the mapping of each point in the simplex to the intersection of the ray from the origin through the point and the sphere.

A **homotopy** between two continuous functions $f, g : X \rightarrow Y$ is a continuous function $H : X \times I \rightarrow Y$ such that $H(x, 0) = f(x)$ and $H(x, 1) = g(x)$, where I is the closed unit interval. One can view H as continuous change from f to g , where each $t \in I$ represents a change in time. If there is a homotopy between two functions f and g they are called homotopic, denoted by $f \simeq g$.

Two spaces X and Y have the same **homotopy type** or alternatively, are **homotopy equivalent**, if there exists continuous functions $f : X \rightarrow Y$ and $g : Y \rightarrow X$ such that $g \circ f \simeq \text{id}_X$ and $f \circ g \simeq \text{id}_Y$, where f and g are known as homotopy equivalences. Homotopy equivalency is an equivalence relation on topological spaces and is denoted by $X \simeq Y$. All homeomorphisms, and thus their inverses, are homotopy equivalences. A space is **contractible** if it has the homotopy type of a point, e.g. $\mathbb{R}^n$ is contractible for all n , as $\mathbb{R}^n \simeq \{*\}$ by the map pulling all points toward the origin.

The Čech complex of a collection of points P at some scale in a metric space is a particular example of a construction called the nerve of a cover. A **cover** of a topological space X is a collection of open sets $\mathcal{U} = \{U_a\}_{a \in A}$ indexed over an arbitrary set A such that $X \subseteq \bigcup_{a \in A} U_a$. A closed cover consists of closed sets and an open cover consists of open sets. We will often assume that the covered space is exactly the union of the elements of the cover set.

Definition 2.11. The *nerve* of a cover $\mathcal{U}$ of a topological space X indexed by A is the abstract simplicial complex

$$\text{Nrv } \mathcal{U} := \{ \sigma \subseteq A \mid \bigcap_{a \in \sigma} U_a \neq \emptyset \text{ and } \#(\sigma) < \infty \}.$$

This naturally gives an alternate formulation of the Čech complex, namely $\mathcal{C}^\alpha(P) = \text{Nrv } \{ \text{ball}(p, \alpha) \}_{p \in P}$, where the collection of metric balls are considered as the cover of P^α .

A **good cover** is a cover where all finite intersections of cover elements are contractible. Any convex set in Euclidean space is contractible and thus their intersections are as well, so any collection of convex sets in Euclidean space is a good cover of the underlying space. Specifically, this applies to ℓ_p metric balls for, which are convex for $p \geq 1$ and $p = \infty$. The Nerve Theorem is a classical result in algebraic topology that has rose to prominence in computational topology and topological data analysis due to its ability to switch from the continuous to the discrete and computable. This is a modern formulation. The classic formulation can be found in Hatcher [35], but it is originally due to Borsuk [6].

Theorem 2.12 (The Nerve Theorem). *Given a finite good closed cover $\mathcal{U}$ of a triangulable space X ,*

$$|\text{Nrv } \mathcal{U}| \simeq X.$$

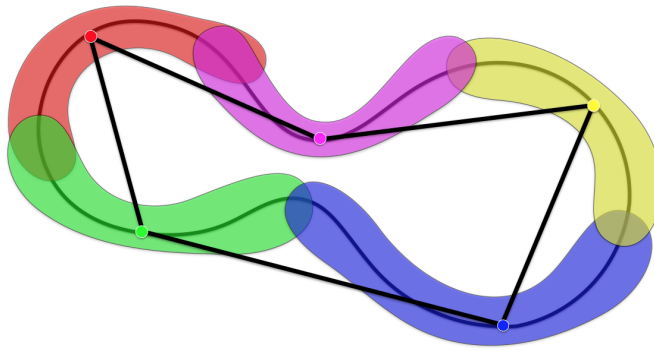


Figure 2.5: A planar curve covered by five sets. The cover's nerve is homeomorphic to the curve.

As the union of finitely many closed Euclidean balls is triangulable, then for a finite point set $P \subset \mathbb{R}^d$ and a scale α , The Nerve Theorem implies that $|\mathcal{C}^\alpha(P)| \simeq P^\alpha$. This homotopy equivalence in conjunction with Equations 2.1 or 2.2 lends itself to topological inference. The nerve complex can be sandwiched between two Rips complexes and looking at the homology map between provides insight to the topology of the covered space.

2.2.2 Chain Complexes and Homology

Homology is an algebraic method of measuring the holes in a topological space. It is coarser than homotopy groups, but it has the benefit of being computable via linear algebra, namely matrix reductions. Many homology theories exist, but in this dissertation we will focus on simplicial homology, the most readily computable homology theory. We will consider homology over the field of two elements, $\mathbb{F}_2 := \{0, 1\}$, where $1 + 1 = 0 \in \mathbb{F}_2$, and equivalently $-1 = 1 \in \mathbb{F}_2$, a fact we will use without mention. This is also written as $\mathbb{Z}/2\mathbb{Z}$ in other mathematics literature. This choice is both for proof simplicity and computational simplicity as computing the homology as a simplicial complex over the field of two elements consists of binary matrix reductions.

First we introduce simplicial chain complexes and simplicial homology theory. Consider a simplicial complex X over a vertex set V . An n -**chain** of X is a formal sum of the form $\sum_i c_i \sigma_i$, where $c_i \in \mathbb{F}_2$ and each σ_i is a unique n -simplex in X .

The k -chains of X over $\mathbb{F}_2$ form an abelian group called the n -**dimensional chain group** $C_n(X; \mathbb{F}_2)$. We will shorten this notation to $C_n(X)$ for the duration as we will not work over any other fields. Formally, $C_{-1}(X) = 0$. Not only this is an abelian group but also a vector space over $\mathbb{F}_2$ with basis consisting of the n -simplices of X .

The n -**th boundary map** $\partial_n : C_n(X) \rightarrow C_{n-1}(X)$, for all $n \geq 0$, is defined on an n -simplex $\sigma = \{v_0, \dots, v_n\}$, for $v_i \in V$.

$$\partial_n(\sigma) = \sum_{i=0}^n \{v_0, \dots, \hat{v}_i, \dots, v_n\},$$

where $\hat{v}_i$ denotes the removal of that vertex from σ .

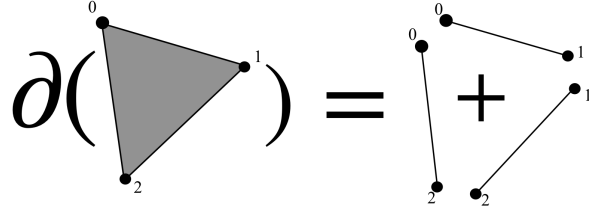


Figure 2.6: The boundary of a 2-simplex is the formal sum of the 1-simplices making up its geometric boundary.

This extends linearly to n -chains in the following way:

$$\partial\left(\sum_i c_i \sigma_i\right) = \sum_i c_i \partial(\sigma_i).$$

Definition 2.13. The *simplicial chain complex* of X is $C_*(X) := (C_n(X))_{n \geq 0}$ along with the boundary maps $\partial = (\partial_n)_{n \geq 0}$, depicted as follows.

$$\dots \longrightarrow C_n(X) \xrightarrow{\partial_n} C_{n-1}(X) \longrightarrow \dots \longrightarrow C_1(X) \xrightarrow{\partial_1} C_0(X) \xrightarrow{\partial_0} 0. \quad (2.3)$$

In the same way that topological spaces have structure-preserving morphisms between them called continuous maps, two chain complexes can be related via the following structure-preserving maps between them.

Definition 2.14. Given two simplicial complexes X and Y , a sequence of maps $f : C_*(X) \rightarrow C_*(Y)$, $f := (f_n)_{n \geq 0}$, where $f_n : C_n(X) \rightarrow C_n(Y)$, is a simplicial **chain map** if for all $n \geq 0$, $\partial_n^Y f_n = f_{n-1} \partial_n^X$, i.e. the following diagram commutes for all $n \geq 0$.

$$\begin{array}{ccccc} \dots & \rightarrow & C_n(X) & \xrightarrow{\partial_n^X} & C_{n-1}(X) & \rightarrow & \dots \\ & & \downarrow f_n & \circlearrowleft & \downarrow f_{n-1} & & \\ \dots & \rightarrow & C_n(Y) & \xrightarrow{\partial_n^Y} & C_{n-1}(Y) & \rightarrow & \dots \end{array} \quad (2.4)$$

Given a simplicial map $\phi : X \rightarrow Y$, there is a corresponding chain map $\phi_* : C_*(X) \rightarrow C_*(Y)$, where for an n -simplex $\sigma = \{v_0, \dots, v_n\} \in C_n(X)$,

$$\phi_n(\sigma) = \begin{cases} \{\phi(v_0), \dots, \phi(v_n)\} & \text{if } \phi(v_i) \neq \phi(v_j) \text{ for all } i, j, \\ 0 & \text{otherwise.} \end{cases}$$

The induced map ϕ_* is a chain map linearly extending to all n -chains in the natural way so it by Definition 2.14 it commutes with the boundary maps. We note this makes the chain complex operation a functor, as depicted in the following commutative diagram, for simplicial complexes X and Y .

Proposition 2.15. *Given a simplicial chain complex $C_*(X)$, $\partial_n \partial_{n+1} = 0$ for all $n \geq 0$.*

Proof. Given an arbitrary $(n+1)$ -simplex $\sigma \in C_{n+1}(X)$, the following equalities hold.

$$\begin{aligned} \partial_n \partial_{n+1}(\sigma) &= \partial_n \left(\sum_{i=0}^{n+1} \{v_0, \dots, \hat{v}_i, \dots, v_{n+1}\} \right) \\ &= \sum_{i=0}^{n+1} \partial_n (\{v_0, \dots, \hat{v}_i, \dots, v_{n+1}\}) \\ &= \sum_{i=0}^{n+1} \sum_{j=0}^{i-1} \{v_0, \dots, \hat{v}_j, \dots, \hat{v}_i, \dots, v_{n+1}\} \\ &\quad + \sum_{i=0}^{n+1} \sum_{j=i+1}^{n+1} \{v_0, \dots, \hat{v}_i, \dots, \hat{v}_j, \dots, v_{n+1}\} = 0. \end{aligned}$$

By symmetry, the two sums are equal and thus add to 0 as our chain complexes are defined over $\mathbb{F}_2$. □

Proposition 2.15 implies that for any simplicial chain complex $C_*(X)$, $\text{im } \partial_{n+1} \subseteq \ker \partial_n \subseteq C_n(X)$. Denote $B_n(X) = \text{im } \partial_{n+1}$ as the n -boundaries, and $Z_n(X) = \ker \partial_n$ as the n -cycles, then we have that the quotient group $Z_n(X)/B_n(X)$ is well defined as each chain group is an abelian group. The **n -th simplicial homology group** of a simplicial complex X is

$$H_n(X) := Z_n(X)/B_n(X).$$

Homology works nicely with respect to chain complex structures in the sense that homology sends chain maps to homological maps, as proven in the following lemma.

Lemma 2.16. *Given a chain map $f : C_*(X) \rightarrow C_*(Y)$, there is an induced homology map $f_n : H_n(X) \rightarrow H_n(Y)$ for all $n \geq 0$.*

Proof. We will prove this by a simple diagram chase. Fix $n \geq 0$. Given $c \in B_n^X$, there exists $d \in C_{n+1}(X)$ such that $\partial_{n+1}^X(d) = c$. As f is a chain map,

$$f_n \partial_{n+1}^X(d) = f_n(c) = \partial_{n+1}^Y f_{n+1}(d),$$

so $f_n(c) \in B_n^Y$. Given $k \in Z_n^X$, $\partial_n^X(k) = 0$, thus

$$f_{n-1} \partial_n^X(k) = 0 = \partial_n^Y f_n(k),$$

so $f_n(k) \in Z_n^Y$. We then have that $f_n(B_n^X) \subseteq B_n^Y$ and $f_n(Z_n^X) \subseteq Z_n^Y$, so there is a quotient map $f_n : H_n(X) \rightarrow H_n(Y)$. $\square$

If a simplicial complex X is finite, then $\dim C_n(X)$ is finite and is equal to the number of n -simplices in X , with the canonical basis of the n -simplices. This also means that the Betti numbers, defined as $\beta_n := \dim H_n(X)$, are finite as well. In fact, by basic linear algebra, $\beta_n = \dim Z_n - \dim B_n$ and $H_n(X)$ is the group generated by the n -dimensional holes in the simplicial complex. β_0 is the number of connected components in X .

The **reduced homology** of a simplicial chain complex is the homology of the chain complex where instead one considers $C_{-1}(X) = \mathbb{F}_2$. One can prove that the 0-dimensional reduced homology satisfies the equality $H_0(X) = \tilde{H}_0(X) \oplus \mathbb{F}_2$. This modification is done to ensure that for the one-point space $\{*\}$, $\tilde{H}_*(\{*\}) = 0$.

Lemma 2.16, along with the aforementioned fact that simplicial maps induce chain maps, implies we can transition from simplicial complexes to chain complexes to homology groups, while preserving each structures' respective morphisms.

Singular homology is the basic homology theory for topological spaces. Although not generally computable, it is necessary to define, as the goal in many parts of this dissertation is to compute the homology of a space that isn't a simplicial complex by computing that of a simplicial complex. Given a topological space X , an n -simplex of X is a map $\sigma : \Delta^n \rightarrow X$, where Δ^n is the standard n -simplex. The boundary is defined analogously to that of a the simplicial boundary in that we consider the summations of the restrictions of the function σ to each of the standard simplex's $n+1$ faces, and the homology groups are defined the same as well. We will simply denote the singular homology of a space as $H_*(X)$ just like simplicial homology, motivated by the following theorem about the homology simplicial complexes.

Similar to simplicial maps between simplicial complexes induce homology maps between their simplicial homology, continuous maps between topological spaces induce homology maps between their singular homology. If two topological spaces are homotopic, then their homology groups are identical, i.e. homology is homotopy invariant. As an example, if a space is contractible, then it is by definition homotopy equivalent to a point, and thus its reduced homology is also trivial.

An important fact is that for a simplicial complex X , the singular homology groups of its geometric realization $|X|$ and the simplicial homology groups of X are isomorphic — see Theorem 2.27 in Hatcher [35]. In our context, this implies that if a space is finitely triangulable, i.e. there exists a finite simplicial complex that is homeomorphic to it, then one can compute the simplicial homology of this triangulation instead of that of the singular homology of the original space. An example of this is computing the homology of the n -sphere by computing the homology of the canonical homeomorphic tetrahedron.

Chapter 3

Unifying Adaptive and Uniform Sampling towards Homology Inference

3.1 Introduction

Surface reconstruction addresses the problem of accurately computing the topology a manifold, or a triangulation of it, from a finite point sample on or near its surface. For a sample to be used for reconstruction, it must be sufficiently dense both locally with respect to salient geometric features and globally. The Hausdorff metric on the space of compact spaces is the standard measure quantifying the uniform density of a sample with respect to the space. This is a global assumption on the point sample. The medial axis of a compact space is the collection of points which have at least two closest points on the space and is intuitively a skeletal representation of it that captures local geometric features. This structure was the original reference for local sampling conditions, now known as adaptive samples. Adaptive sampling was introduced and used by Amenta and Bern in a series of papers [3, 4, 5] to reconstruct planar curves and 2-dimensional manifolds in Euclidean space up to homeomorphism from a point sample. In particular, they introduced the local-feature size with respect to the medial axis of the manifold to be reconstructed.

Though homeomorphic reconstruction is ideal, it is very difficult in three dimensions and generally impossible in higher dimensions, due the sparsity requirements of the sample along the manifold and the knowledge of the medial axis required. As a result, homology inference arose as the best one could hope for with regards to topological reconstruction in arbitrary dimensions. Nyogi, Smale, and Weinberger [42] first proved that one can infer the homology of smooth compact sub-manifold in Euclidean space with high probability from a uniformly randomly chosen sample of it by computing the homology of metric balls around the sample of some radius if the Hausdorff distance between the sample and the sub-manifold is less than $\sqrt{9} - \sqrt{8}$ times the infimum of the local feature size of the sub-manifold.

Motivated by the medial axis' homotopy type being unstable to small perturbations of a space's boundary, Chazal and Lieutier introduced the weak-feature size and the λ -medial axis, a more stable subset of the standard medial axis. The λ -medial axis is stable with respect to boundary perturbations of Hausdorff distance less than the weak feature size [11]. They then proved the homology of a bounded space was identical to that of a Hausdorff-close sample by assuming a lower bound on the weak feature size on space [15]. To do so, they proved that if the Hausdorff distance between a space and a sample is bounded above by a quantity related to the space's weak-feature size, then the tubular regions around each of them are homotopy equivalent for particular radii.

Chazal and Lieutier also proved a reconstruction method for a co-dimension 1 sub-manifold of Euclidean space from a compact sample of it, in addition to approximating the medial axis of the manifold, by using metric balls constructed from both uniform and adaptive samples [16]. Extending upon both the local-feature size and weak-feature size functions via interpolation, Chazal, Cohen-Steiner and Lieutier defined a dynamic feature size called μ -local weak feature size to provide sufficient sampling conditions of compact sets in Euclidean space to perform homology inference [13].

In this research, extended from research done with Sheehy [10], we relate the concepts of uniform samples with respect to the Euclidean metric and adaptive samples with respect to a path-based metric, proving that these sampling methods are equivalent under the correct lens. The adaptive

metric used is similar to a family of metrics studied to aid in path-planning and obstacle avoidance in the plane within the robotics community [2, 47], and can be interpreted as a smoothed version of the adaptive metric used by Clarkson to construct triangulations using so-called ε -nets [19].

Furthermore, conditions are provided under which one can compute the homology of an arbitrarily small expansion of the compact set in question by considering uniform samples, with respect to the adaptive metric and the Euclidean metric, of a disjoint generic reference set (which could be the medial axis, λ -medial axis, μ -medial axis, etc.) and the compact set. The homology inference method of the set in question is computable as well, as it is the homology of the inclusion between two simplicial complexes built from the known samples at different scales — the nerves of finite collections of Euclidean balls centered at the sample points. The generic assumption of the reference set differs from other homology inference research as knowledge of the medial axis is not assumed in our work, naturally leading to tighter sampling parameters, but allowing for the method to be applied to a much broader class of spaces, especially considering no smoothness assumption is made regarding the compact set. These results also provide a connection between the aforementioned adaptive sampling theory and the critical point theory of distance functions of Grove [33].

3.2 Definitions

For the duration of this chapter, let L and X denote disjoint subsets of $\mathbb{R}^d$ that are compact with respect to the Euclidean metric topology. The set X models the sampled space and we refer to L as the *reference set* because it will be the set from which the adaptivity of a sample is measured. In earlier geometric sampling research, L is often the medial axis or a subset of it, which by definition tells one about the local structure of the sampled space. We make no such assumptions here, although it is a valid candidate for L as it is compact and disjoint from X .

First we introduced the adaptive metric from which we will construct the adaptive and uniform sampling equivalence. For $x, y \in \mathbb{R}^d$, define $\Gamma(x, y)$ as the set of continuous paths from x to y

parametrized by Euclidean arc-length. In a similar fashion, define $\Gamma(x, Y) := \bigcup_{y \in Y} \Gamma(x, y)$, the set of continuous paths from x to Y . For any compact set $L \subset \mathbb{R}^d$, recall there is a distance function with respect to L defined on all $x \in \mathbb{R}^d$ by

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

From this distance function the adaptive metric is as follows.

Definition 3.1. For compact $L \subset \mathbb{R}^d$, and $x, y \in \mathbb{R}^d \setminus L$, the **adaptive metric** with respect to L is

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

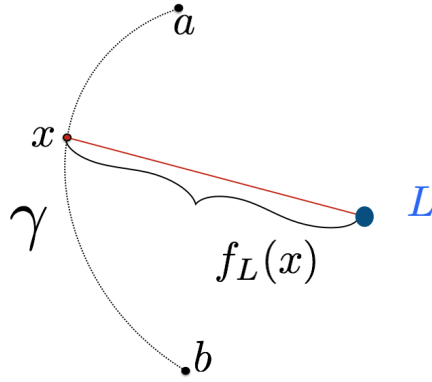


Figure 3.1: The minimal path with respect to the adaptive metric, where L is one point.

If two points x and y are not path-connected in $\mathbb{R}^d \setminus L$, then $d^L(x, y) = \infty$.

Lemma 3.2. For compact $L \subset \mathbb{R}^d$, the function $d^L(\cdot, \cdot)$ is a metric on $\mathbb{R}^d \setminus L$.

Proof. In the following we check each of the axioms of a metric space with respect to d^L . For any $x \in \mathbb{R}^d \setminus L$, consider the trivial one-point curve, $\{x\}$. As this curve has measure 0, $\int_{x \in \{x\}} \frac{dz}{f_L(x)} = 0$ so $d^L(x, x) = 0$. Conversely, assume $d^L(x, y) = 0$ for some $x \neq y \in \mathbb{R}^d \setminus L$. As $\frac{1}{f_L(z)} \in (0, \infty)$ for all $z \in \mathbb{R}^d \setminus L$, the integral is necessarily positive, a contradiction, so $x = y$.

For any x and $y \in \mathbb{R}^d \setminus L$, any $\gamma \in \Gamma(x, y)$ is equal to some $\gamma' \in \Gamma(y, x)$ as sets. The paths are the same with the direction reversed so we have that $d^L(x, y) = d^L(y, x)$ by this symmetry, thus the reflexive property holds.

To prove the triangle inequality property, consider arbitrary x, y and $z \in \mathbb{R}^d \setminus L$ that are path-connected (The case for not path-connected points is trivially true). For all $\varepsilon > 0$, there exists $\gamma_1 \in \Gamma(x, y)$ such that

$$d^L(x, y) \leq \int_{\gamma_1} \frac{dz}{f_L(z)} \leq d^L(x, y) + \frac{\varepsilon}{2}.$$

Likewise there exists $\gamma_2 \in \Gamma(y, z)$ such that

$$d^L(y, z) \leq \int_{\gamma_2} \frac{dz}{f_L(z)} \leq d^L(y, z) + \frac{\varepsilon}{2}.$$

Therefore we have the following

$$d^L(x, z) \leq \int_{\gamma_1 * \gamma_2} \frac{dz}{f_L(z)} \leq \int_{\gamma_1} \frac{dz}{f_L(z)} + \int_{\gamma_2} \frac{dz}{f_L(z)} \leq d^L(x, y) + d^L(y, z) + \varepsilon,$$

and by taking the limit $\varepsilon \rightarrow 0$, we have that $d^L(x, z) \leq d(x, y) + d(y, z)$. $\square$

This metric is well-defined only on $\mathbb{R}^d \setminus L$, as $f_L(l) = 0$ for all $l \in L$, so any path yielding a finite integral is entirely contained in the $\mathbb{R}^d \setminus L$. Notably this implies the paths are bounded away from L . We refer to this metric as *adaptive* because the definition naturally correlates with the concept of an *adaptive sample* as the metric defined as to be dependent on a given points distance to reference set L .

The adaptive metric coincides with the minimum cost of the paths between two given points as used in the robot motion-planning works of Wein et al. and later Agarwal et al. [2, 47]. In particular, the cost of a path is the line integral of the inverse of the “clearance” of each point along the path to the obstacle, where the obstacles in consideration are the polygons in the plane they consider for their obstacles. Visualizing (and computing) the optimal paths becomes complicated as L ceases to be finite. On this note, in Wein et al. [47] it is proven that when L is a single point

in $\mathbb{R}^2$ the minimal paths are logarithmic spirals, and when L is a line segment the minimal paths are circular arcs.

For $y \in \mathbb{R}^d \setminus L$, define the adaptive distance function as

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

and the approximate distance function as

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

Note that f_X^L is a proper distance function and $\widehat{f_X^L}$ is not, however the latter can be viewed as a first-order approximation of the former. The distance function f_X^L is not smooth generally, but a smooth Riemannian distance function can be constructed that approximates it arbitrarily well (see Section 3.5). Though not a distance function, it is a 1-Lipschitz function as shown in Lemma 3.3.

Lemma 3.3. *f_X^L is a 1-Lipschitz function from $(\mathbb{R}^d \setminus L, d^L)$ to $\mathbb{R}$.*

Proof. Consider any $a, b \in \mathbb{R}^d$ and $\varepsilon > 0$. By definition there exists $x \in X$ and $\gamma_1 \in \Gamma(a, x)$ such that

$$f_X^L(a) \leq \int_{\gamma_1} \frac{dz}{f_L(z)} \leq f_X^L(a) + \frac{\varepsilon}{2}.$$

Likewise, there exists $\gamma_2 \in \text{Path}(a, b)$ such that

$$d^L(a, b) \leq \int_{\gamma_2} \frac{dz}{f_L(z)} \leq d^L(a, b) + \frac{\varepsilon}{2}.$$

By construction the concatenation of the two paths $\gamma_1 * \gamma_2$ is in $\Gamma(b, X)$ with an appropriate reparameterization, so for all $\varepsilon > 0$,

$$f_X^L(b) \leq \int_{\gamma_1 * \gamma_2} \frac{dz}{f_L(z)} \leq f_X^L(a) + d^L(a, b) + \varepsilon,$$

so in fact $f_X^L(b) \leq f_X^L(a) + d^L(a, b)$. By symmetry we have $|f_X^L(b) - f_X^L(a)| \leq d^L(a, b)$. $\square$

As both f_X^L and $\widehat{f_X^L}$ are real-valued functions, they generate sublevel sets called offsets in topological data analysis as defined earlier. The sublevel sets of the former are the offsets generated by the adaptive metric d^L while the latter's are approximations thereof.

Definition 3.4. For any compact set $L \subset \mathbb{R}^d$ and compact $X \subset \mathbb{R}^d \setminus L$ the adaptive α -offsets with respect to d^L are

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

Definition 3.5. 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

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

Note the approximate offsets can also be expressed as the union of metric balls of varying radii. Our first goal is to relate the adaptive offsets with respect to X and L to the approximate offsets with respect to $\widehat{X}$ and $\widehat{L}$ via topological interleaving, where $\widehat{L}$ and $\widehat{X}$ are samples of L and X respectively. Once this achieved we can apply the Persistent Nerve Lemma to the interleaving as each offset $B_{\widehat{X}}^{\widehat{L}}(\alpha)$ is a collection of Euclidean balls, leading to a computational homology inference method.

The Hausdorff distance for the adaptive metric d^L is as follows, rewritten in the current notation.

Definition 3.6. The Hausdorff distance between two compact sets $X, Y \in (\mathbb{R}^d \setminus L, d^L)$ is defined as

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

The Hausdorff distance between a space and a sample is a measure of the quality of the sample, namely the uniformity of it with respect to the space it is sampled from. In this paper when we say a uniform sample, we mean one that is Hausdorff-close to the original space. By assuming a bound on the Hausdorff distance between a compact set X and a point sample $\widehat{X}$, we can prove containments between the offsets generated by $f_{\widehat{X}}^L$ and f_X^L for particular scales. The following symmetric relationship between the offsets over all scales is an example of a so-called filtration interleaving (see Definition 3.11.)

Lemma 3.7. Consider compact $\widehat{X}, X \subseteq \mathbb{R}^d \setminus L$ 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. Fix $y \in A_X^L(\alpha)$. By definition $f_X^L(y) \leq \alpha$, which implies that there exists $x \in X$ such that $d^L(x, y) \leq \alpha$. $d_H^L(\hat{X}, X) \leq \delta$ which implies that for all $x \in X$, $f_{\hat{X}}^L(x) \leq \delta$. Now by Lemma 3.3, $f_{\hat{X}}^L(y) \leq f_{\hat{X}}^L(x) + d^L(x, y) \leq \delta + \alpha$, implying $y \in A_{\hat{X}}^L(\alpha + \delta)$. By a symmetric argument, the latter assertion holds. $\square$

The following is our formal definition of an adaptive sample with respect to the Euclidean metric.

Definition 3.8. Given compact set $L \subset \mathbb{R}^d$ and compact sets $\hat{X} \subset X \subset \mathbb{R}^d \setminus L$, we say that $\hat{X}$ is an ε -sample of X , for $\varepsilon \in [0, 1)$, if for all $x \in X$, there exists $p \in \hat{X}$ such that $\|x - p\| \leq \varepsilon f_L(x)$.

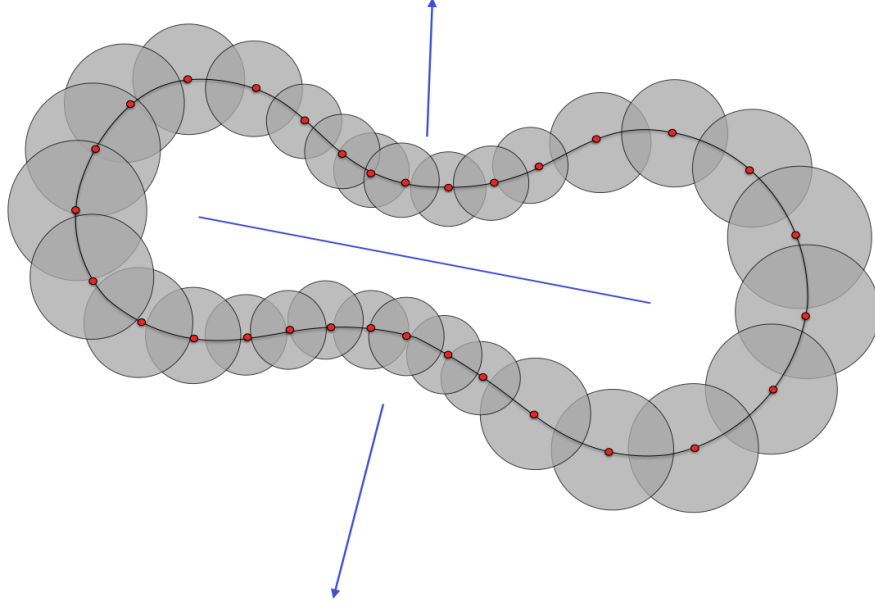


Figure 3.2: A planar curve and an adaptive sample with respect to the medial axis. Note the radius of the metric balls increasing as the points are further from the medial axis.

Recall that the local feature size at a point on a manifold is the distance from it to the manifold's medial axis—the closure of the collection of points with more than one closest point to the manifold. Our definition of an ε -sample is a generalization of the original notion of an adaptive sample defined

by Amenta and Bern [5], which only considers L to be the medial axis of a *manifold* X . Note that there is a nice relationship between an ε -sample and the approximate offsets. If one has an ε -sample $\widehat{X}$ of X , then for all $x \in X$, $\text{ball}(x, \varepsilon f_L(x)) \cap \widehat{X} \neq \emptyset$ which is the same as saying $B_X^L(\varepsilon) \cap \widehat{X} \neq \emptyset$.

3.3 The Equivalence of Adaptive and Uniform Sampling

In this section we prove that a uniform sample with respect to the adaptive metric corresponds to an adaptive sample with respect to the Euclidean metric within certain parameter assumptions. The following lemma is a simpler correspondence for two points between a bound on the adaptive distance between them and their proximity with respect to the Euclidean metric. The choice of a in the two statements is arbitrary.

Lemma 3.9. *Let $L \subset \mathbb{R}^d$ be a compact set and consider $x, y \in \mathbb{R}^d \setminus L$. The following two statements hold for all $\varepsilon \in [0, 1)$.*

(i) *If $d^L(x, y) \leq \varepsilon$, then $\|x - y\| \leq \frac{\varepsilon}{1-\varepsilon} f_L(x)$.*

(ii) *If $\|x - y\| \leq \varepsilon f_L(x)$, then $d^L(x, y) \leq \frac{\varepsilon}{1-\varepsilon}$.*

Proof. (i) Assume $d^L(x, y) \leq \varepsilon$. Given some $\delta > 0$, there exists some $\gamma_\delta \in \Gamma(x, y)$ such that

$$d^L(x, y) \leq \int_{z \in \gamma_\delta} \frac{dz}{f_L(z)} \leq d^L(x, y) + \delta.$$

Note that $\|x - y\|$ is the length of the shortest path $\overline{xy}$ between x and y under the Euclidean metric, so $\|x - y\| \leq |\gamma_\delta|$. Also, for any $z \in \gamma_\delta$, $\|x - z\| \leq |\gamma_\delta|$, as the shortest path that goes from x to y through z contains $\overline{xz}$, which has length $\|x - z\|$. As f_L is 1-Lipschitz because it is a distance function, it follows that $f_L(x) - f_L(z) \leq \|x - z\| \leq |\gamma_\delta|$. From these facts and the main hypothesis

we have the following sequence of inequalities.

$$\begin{aligned}
|\gamma_\delta| &= \int_{z \in \gamma_\delta} dz = (f_L(x) + |\gamma_\delta|) \int_{z \in \gamma_\delta} \frac{dz}{f_L(x) + |\gamma|} \\
&\leq (f_L(x) + |\gamma_\delta|) \int_{z \in \gamma_\delta} \frac{dz}{f_L(z)} \\
&\leq (f_L(x) + |\gamma_\delta|)(d^L(x, y) + \delta) \\
&\leq (f_L(x) + |\gamma_\delta|)(\varepsilon + \delta).
\end{aligned}$$

By considering the limit as $\delta \rightarrow 0$ and rearranging the resulting inequality we have the inequality

$$|\gamma_\delta| \leq \frac{\varepsilon}{1-\varepsilon} f_L(x). \text{ We conclude that } \|x - y\| \leq |\gamma_\delta| \leq \frac{\varepsilon}{1-\varepsilon} f_L(x).$$

(ii) Assume $\|x - y\| \leq \varepsilon f_L(x)$. For all points z in the straight line segment $\overline{xy}$, $\|x - z\| \leq \|x - y\|$ so the following sequence of inequalities holds by f_L being 1-Lipschitz,

$$f_L(z) \geq f_L(x) - \|x - z\| \geq f_L(x) - \|x - y\| \geq (1 - \varepsilon) f_L(x). \quad (3.1)$$

and thus we have the following annotated sequence of inequalities proving the statement.

$$\begin{aligned}
d^L(x, y) &= \inf_{\gamma \in \Gamma(x, y)} \int_{z \in \gamma} \frac{dz}{f_L(z)} && [\text{definition of } d^L] \\
&\leq \int_{z \in \overline{xy}} \frac{dz}{f_L(z)} && [\overline{xy} \in \Gamma(x, y)] \\
&\leq \frac{1}{(1 - \varepsilon) f_L(x)} \int_{z \in \overline{xy}} dz && [\text{Equation 3.1}] \\
&= \frac{\|x - y\|}{(1 - \varepsilon) f_L(x)} && [\text{definition of distance of a curve}] \\
&\leq \frac{\varepsilon}{1 - \varepsilon} .
\end{aligned}$$

□

This lemma directly leads to our theorem which gives a relationship between uniform samples and adaptive samples. Specifically, Theorem 3.10 says that ε -adaptive samples in the Euclidean metric are dual to uniform samples in the adaptive metric d^L with respect to any arbitrary compact set L for $\varepsilon < \frac{1}{2}$.

Theorem 3.10. *Let L and X be compact sets, let $\hat{X} \subset X$ be a sample, and let $\varepsilon \in [0, 1)$ be a constant. If $\hat{X}$ is an ε -sample of X with respect to the distance to L , then $d_H^L(X, \hat{X}) \leq \frac{\varepsilon}{1-\varepsilon}$. Also, if $d_H^L(X, \hat{X}) \leq \varepsilon < \frac{1}{2}$, then $\hat{X}$ is an $\frac{\varepsilon}{1-\varepsilon}$ -sample of X with respect to the distance to L .*

Proof. If $\hat{X}$ is an ε -sample of X , given $x \in X$, there exists $p \in \hat{X}$ such that $\|x - p\| \leq \varepsilon f_L(x)$. By Lemma 3.9, $d^L(x, p) \leq \frac{\varepsilon}{1-\varepsilon}$, so for all $x \in X$, $f_{\hat{X}}^L(x) \leq \frac{\varepsilon}{1-\varepsilon}$. As $\hat{X} \subseteq X$, this proves $d_H^L(\hat{X}, X) \leq \frac{\varepsilon}{1-\varepsilon}$.

If $d_H^L(\hat{X}, X) \leq \varepsilon < \frac{1}{2}$ then for all $x \in X$, $f_{\hat{X}}^L(x) \leq \varepsilon$, thus there exists $p \in \hat{X}$ such that $d^L(x, p) \leq \varepsilon$, and thus by Lemma 3.9, $\|x - p\| \leq \frac{\varepsilon}{1-\varepsilon} f_L(x)$. As $\varepsilon < \frac{1}{2}$, $\frac{\varepsilon}{1-\varepsilon} < 1$, so $\hat{X}$ is an $\frac{\varepsilon}{1-\varepsilon}$ -sample of X . $\square$

Though noteworthy on its own, this theorem lends itself to a homology inference scheme of a space given an adaptive sample by way of the Hausdorff distance bound, which we will explore in Section 3.5 once we have introduced the necessary technical lemmas.

3.4 Filtrations and Interleavings

In this section the goal is to provide an interleaving between the adaptive offsets filtration and the approximate offsets filtration. This relationship between the filtrations will ultimately allow us to infer the homology of the sampled space from the approximate offsets which are generated by Euclidean distance.

A *filtration* is an increasing sequence of sets or topological spaces $\mathcal{F} = (F(\alpha))_{\alpha \geq 0}$ where for all α , $F(\alpha) \subset \mathbb{R}^d$ and $F(\alpha) \subseteq F(\beta)$ iff $\alpha \leq \beta$. Here we specifically consider filtrations that are generated by the sub-level sets of a real-valued function $f : \mathbb{R}^d \rightarrow \mathbb{R}$, i.e. the filtration $\mathcal{F}$ which at scale α is defined as

$$F(\alpha) := \{x \in \mathbb{R}^d \mid f(x) \leq \alpha\} = f^{-1}[0, \alpha].$$

Interleavings provide a constructive relationship between two filtrations' sets. The following definition is a generalization of the standard definition of a topological interleaving, where we allow

for asymmetry and restrictions of the intervals over which the filtrations' elements are interleaved (Asymmetric interleavings have also been addressed in [34].)

Definition 3.11. *A pair of filtrations $(\mathcal{F}, \mathcal{G})$ is (h_1, h_2) -interleaved on 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 $r, h_2(r) \in (s, t)$, where the functions h_1 and h_2 are non-decreasing on the interval (s, t) .*

In topological data analysis, researchers primarily focus on interleavings that are symmetric and infinite, i.e. when $h_1 = h_2$, and the filtrations are interleaved over either the interval $(-\infty, \infty)$, $(0, \infty)$ or $[0, \infty)$. By relaxing the standard definition we can define more specific sampling parameters and conditions on which we can infer the homology of the sampled space with our significantly generic sampling assumptions. Showing the existence of an interleaving between filtrations, explicitly or implicitly, provides insight into the topological and geometric differences (and similarities) between the filtrations. Furthermore, when one filtration is generated by a particularly nice function there is already implied information about the relative shape of the filtration elements. This idea will be expanded upon and utilized in Section 3.5 by considering a so-called smoothing of the adaptive distance function f_X^L , a *nice* function for these purposes.

The following lemma gives us an iterative generic way of combine two pairs of interleaved filtrations over their defined intervals and will ultimately be used to construct the desired relationship.

Lemma 3.12. *If $(\mathcal{F}, \mathcal{G})$ is (h_1, h_2) -interleaved on (s_1, t_1) , and $(\mathcal{G}, \mathcal{H})$ is (h_3, h_4) -interleaved on (s_2, t_2) , then $(\mathcal{F}, \mathcal{H})$ is $(h_3 \circ h_1, h_2 \circ h_4)$ -interleaved on (s_3, t_3) , where $s_3 = \max\{s_1, s_2\}$ and $t_3 = \min\{t_1, t_2\}$.*

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

For the rest of this section, let $\widehat{L} \subseteq L \subset \mathbb{R}^d$ and $\widehat{X} \subseteq X \subset \mathbb{R}^d \setminus L$ be compact sets, with $\widehat{L}$ and $\widehat{X}$ representing compact samples of the reference set L and the set X respectively. This is motivated

by the fact that in practice one only has a finite sample of each. The desired relationship between the adaptive offset filtration $\mathcal{A}_X^L := (A_X^L)_{\alpha \geq 0}$ and the approximate offset filtration $\mathcal{B}_{\hat{X}}^L := (B_{\hat{X}}^L)_{\alpha \geq 0}$ will be provided by an interleaving that is built up by multiple applications of Lemma 3.12 to the interleavings constructed in the remainder of this section.

3.4.1 Approximating X with $\hat{X}$

Lemma 3.13. *If $d_H^L(\hat{X}, X) \leq \varepsilon$, then $(\mathcal{A}_X^L, \mathcal{A}_{\hat{X}}^L)$ is (h_1, h_1) -interleaved on $(0, \infty)$, where $h_1(r) = r + \varepsilon$.*

Proof. This lemma is identical to Lemma 3.7 expressed in our interleaving notation. $\square$

3.4.2 Approximating the Adaptive Offsets

Next we prove that we can approximate the sublevel sets of f_X^L by Euclidean balls, which are more manageable than arbitrary sublevel sets, particularly for computing intersections. These results may be viewed as an extension of the adaptive sampling results of the previous section: Lemma 3.9 and Theorem 3.10.

Lemma 3.14. *Given compact set $L \subset \mathbb{R}^d$, and compact set $X \subset \mathbb{R}^d \setminus L$, for $r \in [0, 1)$, $A_X^L(r) \subseteq B_X^L(\frac{r}{1-r})$, and for $r \in [0, \frac{1}{2})$, $B_X^L(r) \subseteq A_X^L(\frac{r}{1-r})$.*

Proof. Consider $r \in [0, 1)$ and $y \in A_X^L(r)$ such that $f_X^L(y) \leq r$. By definition there exists $x \in X$ such that $d^L(x, y) \leq r$. By Lemma 3.9, this implies that $\|x - y\| \leq \frac{r}{1-r} f_L(x)$, which implies that $y \in B_X^L(\frac{r}{1-r})$.

Now consider $r \in [0, \frac{1}{2})$ and $y \in B_X^L(r)$. By definition, $y \in \text{ball}(x, r f_L(x))$ for some $x \in X$ so $\|x - y\| \leq r f_L(x)$. Applying Lemma 3.9, we have then have that $d^L(x, y) \leq \frac{r}{1-r}$, and as $f_X^L(y) \leq d^L(x, y)$, $y \in A_X^L(\frac{r}{1-r})$. $\square$

A direct corollary of this lemma is that there is a symmetric interleaving between $\mathcal{A}_{\widehat{X}}^L$ and $\mathcal{B}_{\widehat{X}}^L$, noting that the assumptions of Lemma 3.14 are valid for a finite sample $\widehat{X}$ of a compact set X and also for a finite sample $\widehat{L}$ of a compact reference set L .

Corollary 3.15. *The pair $(\mathcal{A}_{\widehat{X}}^L, \mathcal{B}_{\widehat{X}}^L)$ are (h_2, h_2) -interleaved on $(0, \frac{1}{2})$, where $h_2(r) = \frac{r}{1-r}$.*

Proof. This follows from considering Lemma 3.14 with respect to the interleaving notation. $\square$

3.4.3 Approximating L with $\widehat{L}$

The reference set L is often only able to be finitely approximated as it is usually dependent on the space X , whose shape is the object of interest to be discovered. For example, in the case when L is the medial axis of X there are several known techniques for approximating L , e.g. taking some vertices of the Voronoi diagram [5, 24]. By default this uncertainty prevents accurate evaluation of f_L and by extension d^L . We will provide some sampling conditions that allow us to reasonably infer information about L , and said functions, by only looking at a sample $\widehat{L}$. Interestingly, the sampling conditions we use for $\widehat{X}$ are dual to those used for $\widehat{L}$. Specifically we assume an upper-bound on $d_{\widehat{H}}^{\widehat{X}}(L, \widehat{L})$ or equivalently, by Theorem 3.10, $\widehat{L}$ must be an adaptive sample of L with respect to the distance to $\widehat{X}$.

Lemma 3.16. *If $d_{\widehat{H}}^{\widehat{X}}(L, \widehat{L}) \leq \delta < 1$, then $(\mathcal{B}_{\widehat{X}}^L, \mathcal{B}_{\widehat{X}}^{\widehat{L}})$ is (h_3, h_3) -interleaved on $(0, \infty)$, where $h_3(r) = \frac{r}{1-\delta}$.*

Proof. Begin with arbitrary $r \in (0, \infty)$ and $x \in B_{\widehat{X}}^L(r)$. There is a point $p \in \widehat{X}$ such that $\frac{\|x-p\|}{f_L(p)} \leq r$ and there is also a closest point $z \in \widehat{L}$ to p , because $\widehat{L}$ is compact, such that $f_{\widehat{L}}(p) = \|p - z\|$. Together, Theorem 3.10 and the assumption that $d_{\widehat{H}}^{\widehat{X}}(L, \widehat{L}) \leq \delta$ imply that there exists $y \in L$ such that

$$\|y - z\| \leq \frac{\delta}{1 - \delta} f_{\widehat{X}}(z). \quad (3.2)$$

We also know by the definition of the distance-to-set function that

$$f_{\widehat{X}}(z) = \min_{q \in \widehat{X}} \|z - q\| \leq \|z - p\| = f_{\widehat{L}}(p), \quad (3.3)$$

so we can relate $f_L(p)$ to $f_{\widehat{L}}(p)$ as follows

$$\begin{aligned} f_L(p) &\leq \|y - p\| && [y \in L] \\ &\leq \|y - z\| + \|z - p\| && [\text{triangle inequality}] \\ &\leq \frac{1}{1 - \delta} f_{\widehat{L}}(p). && [\text{by (3.2) and (3.3)}]. \end{aligned}$$

Collectively the following is true,

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

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

3.4.4 Composing the Interleavings

We now combine all the previous interleaving results using Lemma 3.12 to arrive at our penultimate theorem which establishes an interleaving between the approximate offsets filtration for the approximate spaces to the adaptive metric offsets filtration for the true spaces.

Theorem 3.17. *Let $\widehat{L} \subseteq L \subset \mathbb{R}^d$ and $\widehat{X} \subseteq X \subset \mathbb{R}^d \setminus L$ 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 $(\mathcal{A}_X^L, \mathcal{B}_{\widehat{X}}^{\widehat{L}})$ are (h_4, h_5) -interleaved on $(0, 1)$, where $h_4(r) = \frac{r + \varepsilon}{(1 - r - \varepsilon)(1 - \delta)}$ and $h_5(r) = \frac{r}{1 - \delta - r} + \varepsilon$.*

Proof. Applying Lemma 3.12 to the interleavings from Lemma 3.13 and Corollary 3.15, we have that $(\mathcal{A}_X^L, \mathcal{B}_{\widehat{X}}^L)$ is $(h_2 \circ h_1, h_1 \circ h_2)$ -interleaved on $(0, 1)$. This interleaving combined with that from Lemma 3.16 implies that $(\mathcal{A}_X^L, \mathcal{B}_{\widehat{X}}^{\widehat{L}})$ is $(h_3 \circ h_2 \circ h_1, h_1 \circ h_2 \circ h_3)$ interleaved on $(0, 1)$. Now we simply

must compute $h_3 \circ h_2 \circ h_1$ and $h_1 \circ h_2 \circ h_3$ as follows.

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

This computation results in our interleaving functions being $h_4(r) = \frac{r + \delta}{(1 - r - \delta)(1 - \varepsilon)}$ and

$$h_5(r) = \frac{r}{1 - \varepsilon - r} + \delta. \quad \square$$

3.5 Smoothing the Metric and Homology Inference

We will now define the smoothed distance function $\widetilde{f}_X^L$ and use it in conjunction with the critical point homotopy equivalence result, Theorem 2.7, to provide a method to infer the homology of the so-called smooth adaptive offsets by looking solely at the approximate offsets.

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 , denoted $\widetilde{f}_L$, which is defined on an arbitrarily smaller subset of Euclidean space.

Lemma 3.18. *Consider a compact set $L \subset \mathbb{R}^d$. Given $\alpha \in (0, 1)$, for all $\beta \in (0, 1)$, there exists smooth function $\widetilde{f}_L : \mathbb{R}^d \setminus L^\beta \rightarrow \mathbb{R}$ 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 [32], for all $\varepsilon > 0$, there exists a smoothing $\widetilde{f}_L : \mathbb{R}^d \setminus L^\beta \rightarrow \mathbb{R}$ 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 3.18. 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 the metric

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

and the smooth distance function $\widetilde{f}_X^L(y) := \widetilde{d}^L(y, X)$.

Lemma 3.19. *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 $\gamma, \gamma' \in \text{Path}(x, y)$ such that $d^L(x, y) \leq \int_{\gamma} \frac{dz}{f_L(z)} \leq d^L(x, y) + \varepsilon$ and $\widetilde{d}^L(x, y) \leq \int_{\gamma'} \frac{dz}{\widetilde{f}_L(z)} \leq \widetilde{d}^L(x, y) + \varepsilon$. We then have the following inequalities resulting from inverting the inequalities in Lemma 3.18.

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

and

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

. 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 $x' := \operatorname{argmin}_{x \in X} d^L(y, x)$ and $x'' = \operatorname{argmin}_{x \in X} \widetilde{d}^L(y, x)$. 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

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

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

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

□

Define the Riemannian adaptive offsets of X as $\widetilde{A}_X^L(\alpha) := \{x \in \mathbb{R}^d \mid \widetilde{f}_X^L(x) \leq \alpha\}$, and denote the corresponding filtration by $\widetilde{\mathcal{A}}_X^L$. The following result reestablishes Lemma 3.19 in the language of filtrations and establishes an interleaving of the Riemannian adaptive offsets with the original adaptive offsets.

Corollary 3.20. *Consider a compact set $L \subset \mathbb{R}^d$. Given $\alpha, \beta \in (0, 1)$, for compact $X \subset \mathbb{R}^d \setminus L^\beta$, there exists a Riemannian distance function $\widetilde{f}_X^L : \mathbb{R}^d \rightarrow \mathbb{R}$, 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 3.19, there exists a Riemannian distance function $\widetilde{f}_X^L : \mathbb{R}^d \rightarrow \mathbb{R}$, 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 $y \in \widetilde{A}_X^L(r)$, $\widetilde{f}_X^L(y) \leq r$, and thus $f_X^L(y) \leq (1 + \alpha)r$, 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 $y \in A_X^L(r)$, $f_X^L(y) \leq r$, and thus $\widetilde{f}_X^L(y) \leq \frac{r}{1-\alpha}$, so $A_X^L(r) \subseteq \widetilde{A}_X^L(\frac{r}{1-\alpha})$. □

Combining the previous corollary with Theorem 3.17 in Subsection 3.4.4, we obtain an interleaving between the Riemannian adaptive offsets and the approximate offsets. This allows us to apply Theorem 2.7 and standard topological data analysis techniques to this interleaving yielding

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

Lemma 3.21. *Given $\alpha, \beta \in (0, 1)$, consider compact sets $\hat{L} \subseteq L \subset \mathbb{R}^d$ and compact sets $\hat{X} \subseteq X \subset \mathbb{R}^d \setminus L^\beta$, such that $d_H^{\hat{X}}(L, \hat{L}) \leq \delta < 1$ and $d_H^L(\hat{X}, X) \leq \varepsilon < 1$, then $(\tilde{\mathcal{A}}_X^L, \mathcal{B}_{\hat{X}}^L)$ are (h_8, h_9) -interleaved on $(0, 1)$, where $h_8(r) = \frac{r + \alpha r + \varepsilon}{(1 - r - \alpha r - \varepsilon)(1 - \delta)}$ and $h_9(r) = \frac{r}{(1 - \alpha)(1 - \delta - r)} + \frac{\varepsilon}{1 - \alpha}$.*

Proof. The hypotheses of the statement satisfy the hypotheses of both Theorem 3.17 and Corollary 3.20 so $(\mathcal{A}_X^L, \mathcal{B}_{\hat{X}}^L)$ are (h_4, h_5) -interleaved on $(0, 1)$, where $h_4(r) = \frac{r + \varepsilon}{(1 - r - \varepsilon)(1 - \delta)}$, and $h_5(r) = \frac{r}{1 - \delta - r} + \varepsilon$ and $(\tilde{\mathcal{A}}_X^L, \mathcal{A}_X^L)$ 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 3.12 and composing the necessary functions, we achieve the stated interleavings. $\square$

For clarity, we rewrite Theorem 2.7, due to Grove [33], for our current setting.

Theorem 3.22. *Let $X \subset \mathbb{R}^d$ be a compact set and let $\widetilde{f}_X^L(\cdot) = \widetilde{d}_L(\cdot, X)$ be the smooth Riemannian adaptive distance function. If an interval $[s, t]$ contains no critical values of $\widetilde{f}_X^L$, then the inclusion $\tilde{A}_X^L(s) \hookrightarrow \tilde{A}_X^L(t)$ is a homotopy equivalence.*

Recall that weak feature size (wfs), introduced by Chazal and Leutier in [15], is the least positive critical value of a Riemannian distance function, and we denote the weak feature size with respect to $\widetilde{f}_X^L$ as $\text{wfs}_L(X)$. Now we introduce our final theorem on homology inference of X and the resulting computational homology inference method. Together they provide a method to infer the homology in all dimensions of an arbitrarily small adaptive offset of our space X of interest by looking at the homology of the inclusion between the approximate offsets of the sample $\hat{X}$ with respect to $\hat{L}$.

Theorem 3.23. *Given $\alpha, \beta \in (0, 1)$, consider compact sets $\hat{L} \subseteq L \subset \mathbb{R}^d$ and compact sets $\hat{X} \subseteq X \subset \mathbb{R}^d \setminus L^\beta$, such that $d_H^{\hat{X}}(L, \hat{L}) \leq \delta < 1$ and $d_H^L(\hat{X}, X) \leq \varepsilon < 1$. 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

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

Proof. Given $\eta > 0$ such that $\Phi\Psi\Phi\Psi(\eta) < 1$, we have the following sequence of inclusions as a result of Lemma 3.21.

$$\tilde{A}_X^L(\eta) \xhookrightarrow{a} B_{\hat{X}}^{\hat{L}}(\Psi(\eta)) \xhookrightarrow{b} \tilde{A}_X^L(\Phi\Psi(\eta)) \xhookrightarrow{c} \dots \quad (3.4)$$

$$\dots \xhookrightarrow{c} B_{\hat{X}}^{\hat{L}}(\Psi\Phi\Psi(\eta)) \xhookrightarrow{d} \tilde{A}_X^L(\Phi\Psi\Phi\Psi(\eta)).$$

As we assume that $\text{wfs}_L(X) > \Phi\Psi\Phi\Psi(\eta)$, by the definition of weak feature size, Lemma 3.22 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.

$$H_*(\tilde{A}_X^L(\eta)) \xrightarrow{a_*} H_*(B_{\hat{X}}^{\hat{L}}(\Psi(\eta))) \xrightarrow{b_*} H_*(\tilde{A}_X^L(\Phi\Psi(\eta))) \xrightarrow{c_*} \dots \quad (3.5)$$

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

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

Furthermore, if we assume that our approximation $\hat{X}$ of X is not just compact but a finite sample, then $B_{\hat{X}}^{\hat{L}}(\alpha)$ at all scales is the union of a finite number of Euclidean balls. Thus, we have the good covers $\{\text{ball}(x, \alpha f_{\hat{L}}(x))\}_{x \in \hat{X}}$ of each $B_{\hat{X}}^{\hat{L}}(\alpha)$, and a corresponding good cover filtration of

$\mathcal{B}_{\hat{X}}^{\hat{L}}$. Note these are good covers. as each element is a Euclidean metric ball so their intersections are contractible. By the Nerve Theorem, Theorem 2.12,

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

As the interleaving maps are inclusions, the Persistent Nerve Lemma [17], Lemma 5.4, applies to Diagram 3.5 when combined with the above isomorphisms. This allows us to examine the nerves to compute the homology of the arbitrarily small adaptive offset of X . Explicitly it yields the following computable homology inference equality,

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

Chapter 4

Homological Sensor Network Coverage and Inference

4.1 Introduction

A geometric sensor network is a collection of sensors, realized as points, distributed in a domain in a metric space where each sensor can detect which portions of the domain are close. The primary goal addressed in sensor network research is accurately measuring some aspects of the domain through the data associated to the sensors, so their usefulness rely on the sensors collectively being able to access the entire domain. Coordinate-free sensor networks are sensor networks in which one does not know the exact location of the sensors in the domain, but each sensor is able to detect whether other sensors are nearby — the goal being to construct global information of the domain from the collective local information provided by each sensor. In the theory, this is interpreted as each point in the space being able to detect nearby points within some distance, or the boundary of the domain. Thus, Lack of coverage of a sensor network can be interpreted as the existence of a hole in the union of the metric balls centered at the sensors. Using homology to check for coverage in coordinate-free sensor networks, known as homological sensor networks, was studied by de Silva and

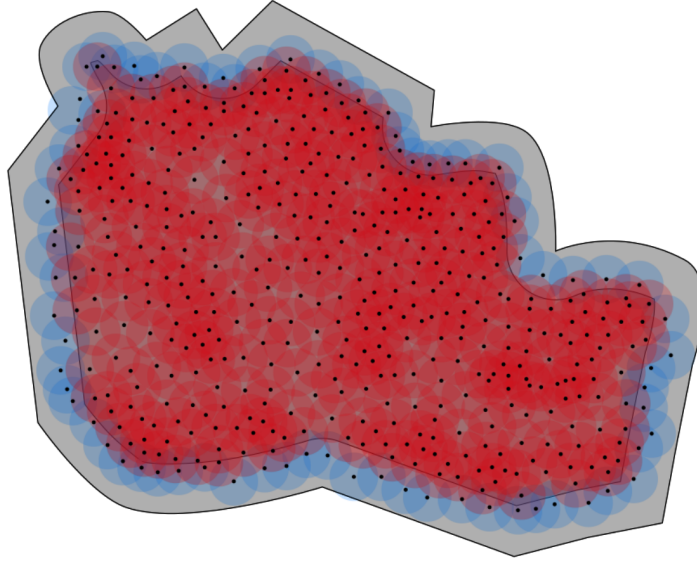


Figure 4.1: A sensor network that is covered by the metric balls of some radius centered at the sensors. Sensing regions that detect the surrounding set are depicted in blue.

Ghrist in a series of papers [22, 23, 45], and in earlier work by Ghrist and Muhammed [29]. Their main result is a computable sufficient condition for coverage of a compact connected Euclidean domain with smooth boundary with the assumption that one can check whether two sensors are within either a small or large detection distance of each other and if they are within some distance of the domain's boundary. This condition is known as the Topological Coverage Criterion (TCC).

In this thesis we present results that relax the domain and input hypotheses of the Topological Coverage Criterion, applied to weighted point samples and k -coverage, and provide a subsampling scheme of the sensors is provided that allows for computable homology inference of the domain assuming there exists sensor network coverage. Checking k -coverage is analogous to ensuring that the sensor network is robust, allowing for a certain amount of sensor failures. Allowing for weighted point samples provides a coverage inference method for sensors with unique area and neighbor detection radii. This is accomplished by successfully separating the geometric and topological details of the problem.

The theorems, lemmas, and proofs in Sections 4.3 and 4.4, along with portions of Section 4.2, are quoted verbatim from a conference paper by Gardner, Sheehy, and I [9], first published in *Proceedings of the ACM-SIAM Symposium on Discrete Algorithms* in 2017, published by the Society for Industrial and Applied Mathematics (SIAM). Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

The main input is two clique graphs built upon the sensors in the domain, in contrast to the potentially difficult-to-compute Rips complexes used in the original certified coverage algorithm. The primary geometric assumption of the domain is that it is a compact length space that is homeomorphic to a subset of Euclidean space, with a so-called surrounding space generalizing the topological boundary used in the original results. Most notably we don't assume the boundary is smooth. The topological assumptions are that the connected components of the domain are neither too close nor too small relative to the detection distances. This results in a computable algorithm that takes a pair of graphs as input and tells that either there is indeed coverage, or better sampling needs to be performed/more sensors need to be added to the domain.

Lastly, we provide a method to compute the homology of a sensor-covered domain, as certified by the new coverage criterion, by examining the homology of the Vietoris-Rips complexes of two subsamples of the sensors in the domain assuming a bound on the weak feature size of the domain's proper boundary. This removes the assumption of having a good sample used in homology inference results, and is replaced with the algorithmic coverage criterion, which certifies the sample is good enough to perform homology inference.

4.1.1 Related Work

In the wake of the foundational work on homological sensor networks, the computability of the Topological Coverage Criterion was investigated. Munch et al. [40] extended the coverage problem to the setting in which each sensor has a probabilistic chance of failure. Under these conditions the authors proved that the problem of checking for coverage is $\#P$ -complete, but created an efficient

algorithm for the problem under the modification where sensors that will fail are flagged ahead of time. Dlotko et al. introduced parallel reduction algorithms for computation of the TCC and provided experimental results of their performance in practice [25].

Others have addressed the problem of computing coverage and holes in homological sensor networks with dynamic/variable sensor locations. Chintakunta, Gamble, and Krim [18, 28] examined networks in which sensors locations change over time by using so-called zigzag persistent homology to identify holes in coverage and the compute the most significant holes — those that persist for the longest periods of time. Similarly, Adams and Carlsson [1] took a topological approach to providing a necessary and sufficient condition for a solution to the evasion problem: finding a time-parametrized path avoiding sensors in a dynamic sensor network. This is a follow-up to the evasion problem introduced in [45], which gives a necessary condition for there to be such a path.

4.2 Definitions

4.2.1 K-Nearest Neighbor and Coverage Regions

One of the ways we generalize the original Topological Coverage Criterion is by addressing the problem of checking for coverage in metric spaces which aren't Euclidean, namely compact length spaces which are homeomorphic to a subset of $\mathbb{R}^d$. A length space is a metric space which has the property that the metric distance between two points in the space is realizable by a curve of the same length. For example, Euclidean space is a length space, where the curve is simply the straight line segment adjoining two points.

Denote by Γ the collection of all rectifiable curves $\gamma : [0, 1] \rightarrow X$. A rectifiable curve is one which has a convergent sequence of increasing piecewise linear curves that approximates its length in a metric space. A metric space (M, d) is a **length space** if for any two points $x, y \in M$,

$$d(x, y) = \inf_{\gamma \in \Gamma} \{L(\gamma) \mid \gamma(0) = x, \gamma(1) = y\}$$

where $L(\cdot)$ is the length of the curve. Note that as if the length space M is compact, then it is complete, so for two points x and $y \in M$, there exists some $\gamma' \in \Gamma$ such that $d(x, y) = L(\gamma')$.

Definition 4.1. For a pair of subsets (X, X') of a compact length space M , with $X' \subset X$, we say that X **surrounds** X' if there does not exist a path from each point $x \in X \setminus X'$ to a point in $\bar{X} := M \setminus X$, the complement of X in M , that does not intersect X' .

Having introduced this definition, we clarify that in the research segment of this chapter we will work directly on the compact homeomorphic subspace of the length space, while noting that paths and local topological characteristics are preserved by homeomorphisms and thus their inverses.

A compact subset $A \subset M$ of a metric space M is a **weighted set** if there exists a *weight* function $w : A \rightarrow \mathbb{R}_{\geq 0}$. For a given point $a \in A$, we denote by w_a the weight $w(a)$ of a . This can be extended to a function on all of M by setting $w_x = 0$ for all $x \in M \setminus A$. The **weighted distance** from a point $x \in M$ to a weighted point $a \in A$ is the power distance

$$\rho_a(x) := \sqrt{d(x, a)^2 + w_a^2}.$$

If a has a weight of 0 then $\rho_a(x) = d(x, a)$, and for points $x, y \in M \setminus A$, both power distances likewise agree with their metric distance. Weighting the points in a metric subset allows us to mathematically model varying sensor detection radii, with larger weights modeling smaller sensing regions.

Definition 4.2. Given a weighted compact set A in a metric space M , the corresponding **weighted k -nearest neighbor distance** is defined at a point $x \in M$ by

$$d_k(x, A) := \inf_{\binom{A}{k}} \max_{y \in K} \rho_y(x).$$

Note that if A is finite $\binom{A}{k}$ is finite so d_k is realized by a particular point $a \in A$.

As with all distance functions, the k -NN distance has associated offsets around A for all non-negative distance parameters, and is also 1-Lipschitz. Given $\alpha \geq 0$ and $k \geq 1$, the **weighted**

(k, α) -**offsets** of a weighted subset $A \subset M$ are defined by

$$A_k^\alpha := \{x \in M \mid d_k(x, A) \leq \alpha\},$$

while the unweighted (standard) α -offsets are

$$A^\alpha := \{x \in M \mid d(x, A) \leq \alpha\}.$$

The α -**coverage region** of a point $a \in A$ is defined as

$$\text{cov}(a, \alpha) := \{x \in M \mid \rho_a(x) \leq \alpha\}.$$

This is identical if to the metric ball of radius α if the point a has weight 0. It is called the coverage region because it signifies the region in which a sensor at point a can check for coverage in a domain.

4.2.2 The k -Barycentric Decomposition and Clique Complexes

Here will we introduce combinatorial definitions used for the algorithmic portion of this chapter. Given a simplicial complex K , a **flag** in K to be an ordered sequence of simplices $(\sigma_i)_{i=0}^t$ where $\sigma_i \in K$, such that $\sigma_0 \subset \dots \subset \sigma_t$. The **barycentric decomposition** of K is the abstract simplicial complex formed by the collection of flags of K and is defined as $\text{Bary}(K) := \{\Sigma \subset K \mid \Sigma \text{ is a flag of } K\}$. The vertices of the barycentric decomposition are the simplices of K . We define the **degree** of a flag $\Sigma = (\sigma_i)_{i=0}^t$ to be $|\Sigma| := \#(\sigma_0)$. The k -**barycentric decomposition** of a simplicial complex K is defined as

$$k\text{-Bary}(K) := \{\Sigma \subset K \mid \Sigma \text{ is a flag in } K, |\Sigma| \geq k\}.$$

Given a simple undirected graph $G = (V, E)$, the **clique complex** of G , $\text{Clq}(G)$, is the combinatorial simplicial complex whose simplices are the collections of vertices that form cliques in G . Recall a clique in a graph (V, E) is a collection of vertices $K \subset V$ such that for all $p \neq q \in K$, $\{p, q\} \in E$. In set notation,

$$\text{Clq}(G) := \{\sigma \subseteq V \mid \forall p \neq q \in \sigma, \{p, q\} \in E, |\sigma| < \infty\}.$$

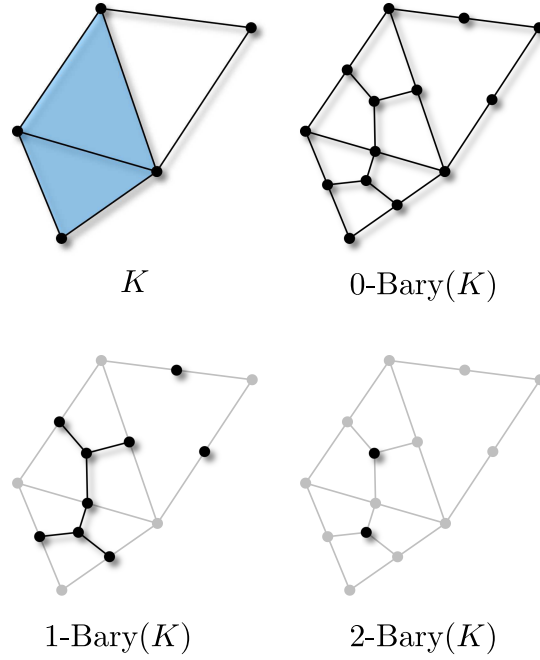


Figure 4.2: A simplicial complex K and its 0, 1, and 2-barycentric decompositions.

For example, the Čech complex of a compact weighted set A at scale α , $\mathcal{C}^\alpha(A)$, is the nerve of the individual α -coverage regions of the points $a \in A$ viewed as a cover of their underlying union. The clique complex of the intersection graph formed by the collection $\{\text{cov}(a, \alpha)\}_{a \in A}$ is equal to $\mathcal{R}^\alpha(A)$.

For a pair of graphs (G, H) such that H is a subgraph of G we denote by $\text{Clq}(G, H)$ the pair of clique complexes $(\text{Clq}(G), \text{Clq}(H))$. Similarly, for two compact subsets $B \subseteq A$ of M and $\varepsilon \geq 0$, $\mathcal{C}_\alpha(A, B) := (\mathcal{C}^\alpha(A), \mathcal{C}^\alpha(B))$. The k -barycentric decomposition of a clique complex $\text{Clq}_k(G)$ is denoted by $\text{Clq}_k(G)$, and for a Čech complex $\mathcal{C}^\alpha(A)$, its k -barycentric decomposition is denoted $\mathcal{C}_k^\alpha(A)$.

4.2.3 Relative (Co)Homology and Duality

Within this chapter we work with both singular homology for generic topological spaces and simplicial homology for simplicial complexes, liberally using the natural isomorphism between singular and simplicial homology for simplicial complexes.

A topological pair is two spaces A and X such that $A \subseteq X$ and is denoted by (X, A) . A map between topological pairs $f : (X, A) \rightarrow (X', A')$ is a map $f : X \rightarrow X'$ such that $f(A) \subseteq A'$. Given a topological pair (X, A) , consider the chain complexes $C_*(A) := (C_n(A))_n$ and $C_*(X) := (C_n(X))_n$. The n -dimensional relative chain group of the pair (X, A) is the quotient group

$$C_n(X, A) := C_n(X)/C_n(A).$$

These chain groups form a chain complex with boundary map $\partial_n^{(X,A)} : C_n(X, A) \rightarrow C_{n-1}(X, A)$ which is induced by ∂_n^X restricted to $C_*(A)$. The homology of this chain complex, computed in the same manner as singular homology, is the **relative homology** of the pair (X, A) and is denoted $H_*(X, A)$.

Relative homology will play a central role in this chapter as we wish to check for coverage of a domain $\mathcal{D}$ minus some offset of the surrounding region $\mathcal{B}$. This can be interpreted as checking for a homological hole corresponding to an uncovered region, but only holes in this shrunken portion of $\mathcal{D}$ are of interest, thus we want to ignore $\mathcal{B}$ when computing the homology.

For nice spaces relative homology can be interpreted intuitively (but not rigorously) as only considering the chains that remain after collapsing A to a point in X , i.e. in X/A . Formally, the following relationship exists between the relative simplicial homology of a pair of simplicial complexes and the singular homology of the quotient space of their geometric realizations. For finite simplicial complexes $A \subset X$ we have following isomorphisms for all dimensions:

$$H_n(X, A) \cong \tilde{H}_n(|X|/|A|),$$

where the latter is the singular homology of the quotient of their geometric realizations. This is an application of the general result of Proposition 2.22 in Hatcher [35].

In fact in our case, we may exploit the geometric nature of our problem to formulate the relative homology in terms of the complements of spaces rather than the quotient of the spaces. The following lemma shows the relationship between the 0-dimensional relative homology of the domain minus an offset of the boundary and the relative homology of their complements. This is

the motivation behind the choice of the word *surrounds*, as no path can leave the interior of the domain to the exterior of the boundary.

Lemma 4.3. *If $\mathcal{B}$ surrounds $\mathcal{D}$, then for all $\varepsilon > 0$, $H_0((\mathcal{D} \setminus \mathcal{B}^\varepsilon, \emptyset) \hookrightarrow (\overline{\mathcal{B}^\varepsilon}, \overline{\mathcal{D}^\varepsilon}))$ is an isomorphism.*

Proof. First we prove the map is injective. Given some non-trivial 0-chain $[x] \in H_0(\mathcal{D} \setminus \mathcal{B}^\varepsilon)$, we can pick a representative point $x \in \mathcal{D} \setminus \mathcal{B}^\varepsilon \subseteq \overline{\mathcal{B}^\varepsilon}$. Because $\mathcal{B}$ surrounds $\mathcal{D}$, there does not exist a path from $\mathcal{D} \setminus \mathcal{B}$ to $\overline{\mathcal{D}}$ that does not intersect $\mathcal{B}$, and so there does not exist a path from $\mathcal{D} \setminus \mathcal{B}^\varepsilon$ to $\overline{\mathcal{D}^\varepsilon}$ that does not intersect $\mathcal{B}^\varepsilon$. Thus, $[x] \neq 0 \in H_0(\overline{\mathcal{B}^\varepsilon}, \overline{\mathcal{D}^\varepsilon})$.

Next we prove the map is surjective. Any $[x] \in H_0(\overline{\mathcal{B}^\varepsilon}, \overline{\mathcal{D}^\varepsilon})$, is represented by a point x in a connected component of $\overline{\mathcal{B}^\varepsilon} \setminus \overline{\mathcal{D}^\varepsilon} = \mathcal{D} \setminus \mathcal{B}^\varepsilon$, and thus a homology class $[x] \in H_0(\mathcal{D} \setminus \mathcal{B}^\varepsilon)$. $\square$

Standard homology and relative homology are connected via the following exact sequence for all homology theories. For a pair of spaces (X, A) there is the canonical exact sequence called the exact homology sequence of the pair

$$\dots \rightarrow H_n(A) \rightarrow H_n(X) \rightarrow H_n(X, A) \rightarrow H_{n-1}(A) \rightarrow \dots,$$

where the map that drops a dimension is known as the connecting homomorphism. An exact sequence is one in which the image of a map is the kernel of the next map. With this in mind, the relative homology of the pair (X, A) represents the difference between the homology of X and A , and it can be computed if one knows all the homology groups of A and X .

In the same way that vector spaces have dual vector spaces, the algebraic structure dual to homology is called cohomology. Given a topological space X , recall the formulation of its singular chain complex $C_*(X)$. Consider the vector space of linear maps from $C_n(X)$ to $\mathbb{F}_2$, written succinctly by $\text{Hom}(C_n(X), \mathbb{F}_2)$. Define the singular n -dimensional **cochain group** as the vector space (or abelian group) as

$$C^n(X) := \text{Hom}(C_n(X), \mathbb{F}_2),$$

whose elements are called n -cochains. To define the coboundary map between cochain groups, consider the following diagram of vector spaces, for cochain $f \in C_n(X)$.

$$\begin{array}{ccc} C_{n+1}(X) & \xrightarrow{f\partial_n} & \mathbb{F}_2 \\ \downarrow \partial_n & \nearrow f & \\ C_n(X) & & \end{array} \quad (4.1)$$

Define the n -dimensional coboundary map as $d_n := f\partial_n$. Note that this sends n -cochains to $(n+1)$ -cochains, and that $d_{n+1}d_n = f\partial_n\partial_{n+1} = 0$. From this relationship we define the **n -dimensional cohomology groups** by

$$H^n(X) := \ker d_n / \operatorname{im} d_{n-1}.$$

Relative cohomology is defined in the analogous fashion as relative homology.

In mathematics as a whole, cohomology theories are studied more as they richer than homology theories in the sense that there are further algebraic structures built from them. In this work, we use cohomology to utilize one of the several dualities between homology and cohomology known as Alexander Duality. This duality exists in many forms, the formulation we will use follows, where a refinement is the resulting triangulation from a subdivision of the simplices of a triangulation. The complements of the simplicial complexes in Theorem 4.4 are the simplicial complexes in K that do not share vertices with the complemented complex.

Theorem 4.4 (Alexander Duality [36]). *Let K be an abstract simplicial complex which is a combinatorial oriented d -manifold. Given a pair (L, M) , where L is a subcomplex of some refinement of K , and M is a subcomplex of L , there is a natural isomorphism*

$$H^r(L, M) \rightarrow H_{d-r}(\overline{M}, \overline{L}).$$

One last prerequisite for our results is how to manage cohomology, which is less practical computationally (and intuitively) than homology. The Universal Coefficient Theorem (see Hatcher, Theorem 3.2 [35]) implies that for a chain complex $C_*(X)$ over a group G ,

$$H^n(X; G) \cong \operatorname{Hom}(H_n(X); G) \oplus \operatorname{Ext}(H_{n-1}(X), G),$$

where Ext is a technical result from algebraic topology and homological algebra and $\oplus$ is the direct sum (for example, as used in linear algebra). What is of importance to us is that when G is a field, specifically in our work where $G = \mathbb{F}_2$, Ext is 0, so $\text{Ext}(H_{n-1}(X), \mathbb{F}_2) = 0$. In addition, $\text{Hom}(H_n(X), \mathbb{F}_2)$ is the dual vector space to $H_n(X)$, and if $H_n(X)$ is finitely generated, i.e. of finite dimension, then $H_n(X) \cong \text{Hom}(H_n(X), \mathbb{F}_2)$ as all vector spaces of the same dimension over the same base field are isomorphic. In conclusion, the Universal Coefficient Theorem tells us that for topological spaces X that have finite homology groups over a field, $H^n(X) \cong H_n(X)$.

The last technical lemma that must be introduced, due to Sheehy [44], is a generalization of the Persistent Nerve Lemma which relates the relative homology of pairs of k -NN offsets to the relative homology of the k -barycentric decompositions of the corresponding pairs of Čech complexes.

Lemma 4.5. *For any $B \subset A \subseteq M$, if the coverage regions $\{\text{cov}(a, \alpha) \mid a \in A\}$ form a good cover of A_1^α and similarly for A_1^β , B_1^α , and B_1^β , then the following diagram commutes for all k and the vertical maps are isomorphisms.*

$$\begin{array}{ccc} H_*(A_k^\alpha, B_k^\alpha) & \longrightarrow & H_*(A_k^\beta, B_k^\beta) \\ \downarrow \cong & & \downarrow \cong \\ H_*(C_k^\alpha(A, B)) & \longrightarrow & H_*(C_k^\beta(A, B)) \end{array}$$

4.3 The Geometric Topological Coverage Criterion

In this section we introduce our algorithm for checking for k -wise coverage of a domain by sensors. First we will outline the geometric assumptions we make of the domain, surrounding set, and sensors, and then prove our Geometric Topological Coverage Criterion (Geometric TCC). In the next section, we will prove an algorithm that checks a sufficiently combinatorial condition for coverage.

Fix $k \geq 1$, the number of sensors we wish to cover each point in the domain. In real world applications, this could represent error-proofing sensor network coverage, ensuring that if any sensor

fails that another one will still be able to report on that part of the domain. We consider a pair of subspaces $(\mathcal{D}, \mathcal{B})$ of a length space, where $\mathcal{B}$ is closed and surrounds $\mathcal{D}$ and $\mathcal{D}$ is compact, and is homeomorphic to a subset of $\mathbb{R}^d$. These represent the domain where the sensors are, and the “border” of the domain which the sensors can detect if they are within some distance from it.

The sensors are represented by a weighted finite point sample $P \subset \mathcal{D}$, where the weights represent each sensor’s varying sensing radii, which could correspond to varying density of a physical object, or environmental obstacles like trees obstructing radio signals from a collection of towers for example. There are two non-negative sensing parameters α and β such that $3\alpha \leq \beta$, which represent the two radii at which the sensors can detect each other’s presence. Explicitly, a sensor can discern whether another is sensor is within both α and β distance from each other. The sensors that are α -close to the surrounding set we denote by $Q = \{p \in P \mid \text{cov}(p, \alpha) \cap \mathcal{B} \neq \emptyset\}$.

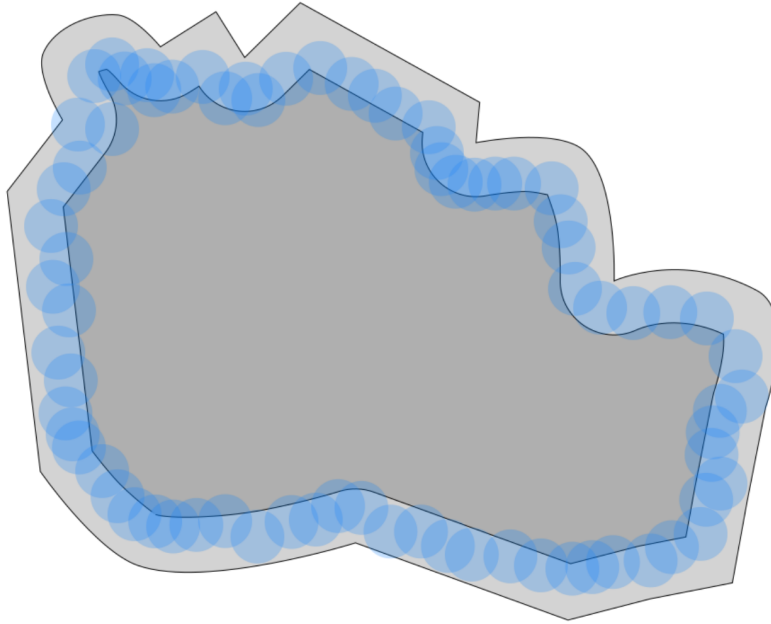


Figure 4.3: The blue metric balls are the coverage regions of the subset of sensors that are within distance α of the surrounding set.

Though each sensor can check for other sensors within distance β , the coverage radius is defined

to be α . Our ultimate goal is to check for coverage of $\mathcal{D} \setminus \mathcal{B}^{2\alpha}$ by k metric balls of P with radius α . This is called k -coverage and the coverage condition be written succinctly as $\mathcal{D} \setminus \mathcal{B}^{2\alpha} \subset P_k^\alpha$.

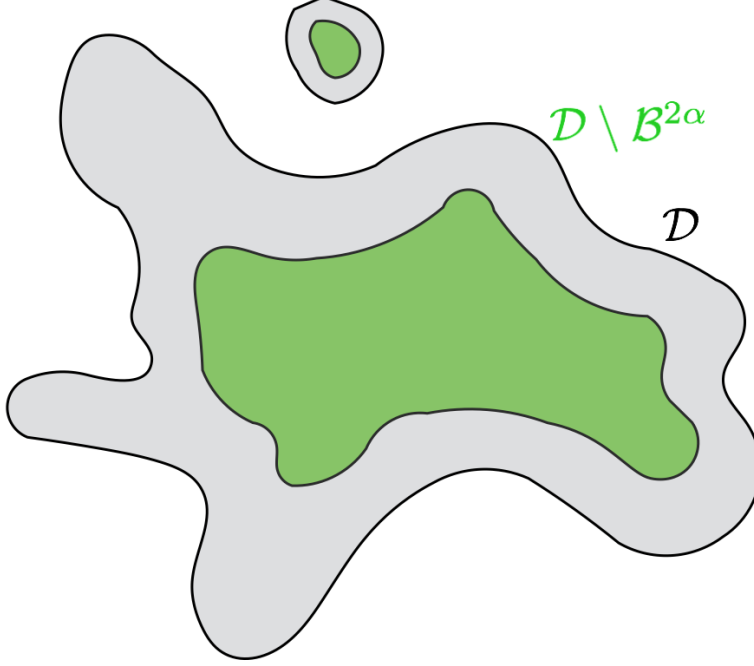


Figure 4.4: A compact domain $\mathcal{D}$ in $\mathbb{R}^2$ and the coverage region of it with respect to its boundary, in green.

We also assume that the domain and the surrounding set is that the connected components are not too small and that they are well-separated to the sensing radii, namely the map $H_0(\mathcal{D} \setminus \mathcal{B}^{\alpha+\beta} \hookrightarrow \mathcal{D} \setminus \mathcal{B}^{2\alpha})$ is surjective and the map $H_0(\mathcal{D} \setminus \mathcal{B}^{2\alpha} \hookrightarrow (\mathcal{D} \setminus \mathcal{B})^{2\alpha})$ is injective. The latter assumption will only be necessary as a hypothesis in the algorithmic TCC, but we state here for consistency of presentation. Below we summarize the assumptions made on the domain and sensors, and we number them in order to be referenced in each of the lemmas and theorems for both simplicity of presentation and minimizing the necessary hypotheses for the results to hold.

Domain and Sensor Assumptions

0. $\mathcal{D}$ is a bounded, compact length space homeomorphic to a subset of $\mathbb{R}^d$ and $\mathcal{B} \subseteq \mathcal{D}$ is closed and surrounds $\mathcal{D}$.
1. The map $H_0(\mathcal{D} \setminus \mathcal{B}^{\alpha+\beta} \hookrightarrow \mathcal{D} \setminus \mathcal{B}^{2\alpha})$ is surjective.
2. The map $H_0(\mathcal{D} \setminus \mathcal{B}^{2\alpha} \hookrightarrow (\mathcal{D} \setminus \mathcal{B})^{2\alpha})$ is injective.

Having introduced the assumptions of the sensor network and its surrounding set, we now can work to a proof of our new topological coverage criterion for k -coverage. Consider Diagram 4.2 consisting of pairs of spaces, with inclusions component-wise, whose existence is guaranteed by the fact that $\alpha < \beta$. From a high-level perspective examining at the change in the topology of the offsets at the two different sensory scales removes noise.

$$\begin{array}{ccc}
 (P_k^\alpha, Q_k^\alpha) & \hookrightarrow & (P_k^\beta, Q_k^\beta) \\
 \downarrow & & \downarrow \\
 (\mathcal{D}^{2\alpha}, \mathcal{B}^{2\alpha}) & \hookrightarrow & (\mathcal{D}^{\alpha+\beta}, \mathcal{B}^{\alpha+\beta})
 \end{array} \tag{4.2}$$

By complementing all the spaces in the aforementioned diagram we get the commutative Diagram 4.3, where j is the map between the complemented domain and surrounding set and i is the map between the complemented k -NN offsets of the sensors and the sensors close to the surrounding set. Note that here we define the complement of a subspace $X \subseteq \mathbb{R}^d$ to be $\overline{X} := (\mathbb{R}^d \cup \{\infty\}) \setminus X$, with the point at infinity being added as to allow for the use of Alexander Duality. Compactifying the space does not affect our domain and sensors' offsets however as we assume that $\mathcal{D}$ is compact and thus bounded in $\mathbb{R}^d$.

$$\begin{array}{ccc}
 (\overline{\mathcal{B}^{\alpha+\beta}}, \overline{\mathcal{D}^{\alpha+\beta}}) & \xhookrightarrow{j} & (\overline{\mathcal{B}^{2\alpha}}, \overline{\mathcal{D}^{2\alpha}}) \\
 \downarrow & & \downarrow \\
 (\overline{Q_k^\beta}, \overline{P_k^\beta}) & \xhookrightarrow{i} & (\overline{Q_k^\alpha}, \overline{P_k^\alpha})
 \end{array} \tag{4.3}$$

By applying the 0-dimensional relative homology functor, Diagram 4.3 becomes Diagram 4.4, where the homology maps i_* and j_* are the maps of interest. Namely we examine the map p_* between the offset's and domain's images,

$$p_* : \text{im } j_* \rightarrow \text{im } i_*.$$

As the 0-dimensional relative homology is isomorphic to the d -dimensional homology by the Universal Coefficient Theorem, we are indirectly looking for the lack of a hole in the d -dimensional hole in the homology of the k -NN sensors offsets relative to the surrounding set offsets, as this corresponds to a lack of coverage of $\mathcal{D} \setminus \mathcal{B}^{2\alpha}$. Looking at the 0-dimensional relative homology of the complements is easier however due to our assumptions about the connected components, Assumptions 1 & 2.

$$\begin{array}{ccc} H_0(\overline{\mathcal{B}^{\alpha+\beta}}, \overline{\mathcal{D}^{\alpha+\beta}}) & \xrightarrow{j_*} & H_0(\overline{\mathcal{B}^{2\alpha}}, \overline{\mathcal{D}^{2\alpha}}) \\ \downarrow & & \downarrow \\ H_0(\overline{Q_k^\beta}, \overline{P_k^\beta}) & \xrightarrow{i_*} & H_0(\overline{Q_k^\alpha}, \overline{P_k^\alpha}) \end{array} \quad (4.4)$$

Lemma 4.6 states that under our assumptions, p_* is surjective, which is to say that all homology elements that survive j_* have a homology class representative in $\text{im } i_*$.

Lemma 4.6. *Given Assumptions 0 and 1, the map $p_* : \text{im } j_* \rightarrow \text{im } i_*$ is surjective.*

Proof. First we have by Assumption 1 and Lemma 4.3, which states how a surrounding set interacts with relative homology, that j_* is surjective, represented by the following diagram.

$$\begin{array}{ccc} H_0(\mathcal{D} \setminus \mathcal{B}^{\alpha+\beta}) & \longrightarrow & H_0(\mathcal{D} \setminus \mathcal{B}^{2\alpha}) \\ \downarrow \cong & & \downarrow \cong \\ H_0(\overline{\mathcal{B}^{\alpha+\beta}}, \overline{\mathcal{D}^{\alpha+\beta}}) & \xrightarrow{j_*} & H_0(\overline{\mathcal{B}^{2\alpha}}, \overline{\mathcal{D}^{2\alpha}}) \end{array} \quad (4.5)$$

Choose a basis for $\text{im } i_*$ such that each basis element is represented by a point in $P_k^\alpha \setminus Q_k^\beta$. Consider $x \in P_k^\alpha \setminus Q_k^\beta$ such that $[x] \neq 0 \in \text{im } i_*$.

Suppose $x \in \mathcal{B}^{2\alpha}$. By definition of the offsets, there is a point $y \in \mathcal{B}$ such that $d(x, y) \leq 2\alpha$. Because $x \in \overline{Q_k^\beta}$ by hypothesis, $d_k(x, Q) > \beta$. We will show that if a point z is in the shortest path

$\overline{xy}$, then $z \in \overline{Q_k^\alpha}$. For any $z \in \overline{xy}$, we have $d(x, z) \leq d(x, y) \leq 2\alpha$, thus the following inequality holds.

$$\begin{aligned} d_k(z, Q) &\geq d_k(x, Q) - d(x, z) && [d_k \text{ is Lipschitz}] \\ &> \beta - 2\alpha && [d_k(x, Q) > \beta, d(x, z) \leq 2\alpha] \\ &\geq \alpha. && [\beta \geq 3\alpha] \end{aligned}$$

From this inequality we conclude that $z \in \overline{Q_k^\alpha}$ for all $z \in \overline{xy}$, and thus $\overline{xy} \subseteq \overline{Q_k^\alpha}$ and $y \in \mathcal{B} \cap \overline{Q_k^\alpha}$. Now suppose $y \in P_k^\alpha$ as well. By the definition of P_k^α , for some $A \in \binom{P}{k}$, $y \in \text{cov}(p, \alpha)$ for all $p \in A$, which implies that $A \subseteq Q$, so $y \in Q_k^\alpha$, a contradiction. Thus we may conclude that $y \notin P_k^\alpha$, which is equivalent to $y \in \overline{P_k^\alpha}$.

Any path $\gamma : [0, 1] \rightarrow \overline{Q_k^\alpha}$ such that $\gamma(0) = x$ and $\gamma(1) = y$, generates a class $[\gamma]$ in the chain group $C_1(\overline{Q_k^\alpha})$ containing γ . For $[\gamma] \in C_1(\overline{Q_k^\alpha}, \overline{P_k^\alpha})$ it follows $\partial([\gamma]) = [x + y] = [x]$ as $y \in \overline{P_k^\alpha}$, and therefore that there must exist $z \in \overline{xy} \cap Q_k^\alpha$. This is a contradiction as we have shown that $\overline{xy} \cap Q_k^\alpha = \emptyset$, and thus x cannot be in $\mathcal{B}^{2\alpha}$.

Now, we may assume $x \in \overline{\mathcal{B}^{2\alpha}}$. Then $x \in \mathcal{D} \setminus \mathcal{B}^{2\alpha}$ so $[x] \neq 0 \in H_0(\overline{\mathcal{B}^{2\alpha}}, \overline{\mathcal{D}^{2\alpha}})$. Because j_* is surjective, $H_0(\overline{\mathcal{B}^{2\alpha}}, \overline{\mathcal{D}^{2\alpha}}) = \text{im } j_*$ and thus $p_*([x]) = [x]$ and so $[x] \in \text{im } p_*$. It follows that p_* is surjective. $\square$

Lemma 4.7 gives a coverage condition: if p_* injective, then our formulation of coverage exists. Importantly, if p_* is injective, by Lemma 4.6, then p_* is an isomorphism between the images of i_* and j_* .

Lemma 4.7. *Given Assumptions 0 and 1, if p_* is injective then $\mathcal{D} \setminus \mathcal{B}^{2\alpha} \subseteq P_k^\alpha$.*

Proof. The proof is essentially the same as that presented by de Silva & Ghrist [22]. We include it here in our own notation for completeness.

Suppose for contradiction p_* is injective and there exists a point $x \in (\mathcal{D} \setminus \mathcal{B}^{2\alpha}) \setminus P_k^\alpha$. Thus $[x] \neq 0 \in H_0(\overline{\mathcal{B}^{2\alpha}}, \overline{\mathcal{D}^{2\alpha}})$, because the point x is in some connected component of $\mathcal{D} \setminus \mathcal{B}^{2\alpha}$. Moreover,

$[x] \in \text{im } j_*$, because j_* is surjective by Assumption 1. Consider the following sequence of maps induced by inclusions.

$$H_0(\overline{\mathcal{B}^{2\alpha}}, \overline{\mathcal{D}^{2\alpha}}) \xrightarrow{f_*} H_0(\overline{\mathcal{B}^{2\alpha}}, \overline{\mathcal{D}^{2\alpha}} \cup \{x\}) \xrightarrow{g_*} H_0(\overline{Q_k^\alpha}, \overline{P_k^\alpha})$$

As $f_*([x]) = 0$ in $H_0(\overline{\mathcal{B}^{2\alpha}}, \overline{\mathcal{D}^{2\alpha}} \cup \{x\})$, then $p_*([x]) = (g_* \circ f_*)([x]) = 0$, contradicting the assumption that p_* is injective. $\square$

The prior two lemmas lead to Theorem 4.8, which we call the Geometric Topological Coverage Criterion. It gives a topological sufficient condition guaranteeing weighted k -coverage of $\mathcal{D} \setminus \mathcal{B}^{2\alpha}$ by comparing the ranks of the maps i_* and j_* , which segues into the next section where we give an algorithmic method for checking coverage, as the ranks of linear maps are computable. This argument is simpler than that of the original TCC which focuses on analyzing the map between two embedded Rips complexes.

Theorem 4.8 (The Geometric TCC). *Let $(\mathcal{D}, \mathcal{B})$ be a pair of spaces satisfying Assumptions 0 and 1. Let α and β be constants such that $0 < 3\alpha \leq \beta$. Let $P \subset \mathcal{D}$ be a finite set with $Q = \{p \in P \mid \text{cov}(p, \alpha) \cap \mathcal{B} \neq \emptyset\}$. Let i_* and j_* be the maps in Diagram (4.4). If $\text{rk } i_* \geq \text{rk } j_*$ then $\mathcal{D} \setminus \mathcal{B}^{2\alpha} \subseteq P_k^\alpha$.*

Proof. Lemma 4.6 implies that $p_* : \text{im } j_* \rightarrow \text{im } i_*$ is surjective, so $\text{rk } i_* \leq \text{rk } j_*$. So with the assumption that $\text{rk } i_* \geq \text{rk } j_*$, $\text{rk } i_* = \text{rk } j_*$. Since P is a finite point set, $\text{im } i_*$ is finite-dimensional and by equality, $\text{im } j_*$ is as well, so p_* is an isomorphism and in turn injective. Lemma 4.7 then implies $\mathcal{D} \setminus \mathcal{B}^{2\alpha} \subseteq P_k^\alpha$. $\square$

4.4 The Algorithmic Topological Coverage Criterion

Though the Geometric TCC gives a topological verification method for coverage, it is not computable as we know little about the domain besides what local metric information the finite collection of sensors tell us, and the connected components assumptions, which only tell about what

is preserved between scales. In this section we introduce an computational method to inferring coverage that makes use of the geometric result from Section 4.3. To introduce the algorithm and its proof of correctness, we first introduce the assumptions on our input, adding to the previously stated the geometric assumptions.

We begin with a weighted finite point set P from which our input to the TCC algorithm is built. The input is two graphs (G_1, G_2) over vertex set P , and their respective restrictions to the vertex set Q , $(G_1[Q], G_2[Q])$, where we remind the reader that Q is the subset of the sensors whose coverage regions intersect the surrounding set $\mathcal{B}$. For technical purposes we assume that the vertex set P , has a point on each connected component of $\mathcal{D} \setminus \mathcal{B}^{2\alpha}$ to avoid degenerate inputs where coverage impossible.

Given the two input graphs, we also assume that the k -barycentric decomposition Čech complexes of the coverage regions at scales α and β , $\mathcal{C}_k^\alpha(P)$ and $\mathcal{C}_k^\beta(P)$ are sandwiched between the k -clique complexes of the two graphs (definition is found in Subsection 4.2.2). Though this seems like a specific condition, it in fact generalizes the input of De Silva and Ghrist, who used Rips complexes built from sensors at two different scales as their coverage algorithm's input. Specifically, if the domain has Euclidean metric and the two graphs are the intersection graphs of the k -NN metric balls of P at scale α/ϑ_d and β , then the clique complexes of these graphs are $\mathcal{R}^{\alpha/\vartheta_d}(P)$ and $\mathcal{R}^\beta(P)$, so the k -clique complexes are the k -Rips complexes and the sandwich assumption is a direct result of the Jung's Theorem corollary, Equation 2.2. Furthermore, we assume that the coverage regions of radius α and β of the sensors form a good cover, allowing for the application of the relative Persistent Nerve Lemma.

In order to apply Alexander Duality, Theorem 4.4, we assume the existence of nested triangulations of the k -NN ε -offsets P_k^ε and Q_k^ε for $\varepsilon \in \{\alpha, \beta\}$, the strong and weak sensing radii, which are consistent with some triangulation of the compactification of $\mathbb{R}^d$.

Algorithmic Assumptions

3. Each connected component of $\mathcal{D} \setminus \mathcal{B}^{2\alpha}$ contains a point in P .
4. The graphs G_1, G_2 are defined over vertex set $P \subset \mathcal{D}$, and subgraphs $G_1[Q], G_2[Q]$ induced by restriction to the vertex set $Q = \{p \in P \mid \text{cov}(p, \alpha) \cap \mathcal{B} \neq \emptyset\}$.
5. $\text{Clq}_k(G_1) \subseteq \mathcal{C}_k^\alpha(P) \subseteq \mathcal{C}_k^\beta(P) \subseteq \text{Clq}_k(G_2)$.
6. $\mathcal{U} = \{\text{cov}(p, \varepsilon) \mid p \in P\}$ is a good cover for $\varepsilon \in \{\alpha, \beta\}$.
7. There is a triangulation $\mathcal{K}$ of $\mathbb{R}^d \cup \{\infty\}$ and triangulations L_ε and M_ε of P_k^ε and Q_k^ε respectively, such that $M_\varepsilon \subset L_\varepsilon$ in $\mathcal{K}$, for $\varepsilon \in \{\alpha, \beta\}$.

First we will introduce the result which states an Alexander Duality theorem with respect to the sensors' offsets. By assumption 7, there exists a sufficiently refined triangulation $\mathcal{K}$ of $\mathbb{R}^d \cup \{\infty\}$ and respective triangulations of $P_k^\alpha, Q_k^\alpha, P_k^\beta$, and Q_k^β . Due to the naturality of the Alexander Duality theorem 4.4 and the fact that relative homology is a invariant under homeomorphisms of pairs allows us to state the following reinterpretation with our spaces.

Corollary 4.9. *Given α and $\beta > 0$ and P and Q satisfying Assumption 7, the following diagram commutes where the horizontal maps are isomorphisms.*

$$\begin{array}{ccc}
 H_0(\overline{Q_k^\alpha}, \overline{P_k^\alpha}) & \xrightarrow{\cong} & H^d(P_k^\alpha, Q_k^\alpha) \\
 \downarrow & & \downarrow \\
 H_0(\overline{Q_k^\beta}, \overline{P_k^\beta}) & \xrightarrow{\cong} & H^d(P_k^\beta, Q_k^\beta).
 \end{array} \tag{4.6}$$

The following lemma gives an upper-bound on the rank of the d -dimensional homology map between the pairs of graphs, relating the clique complexes of the graphs to the homology maps between the spaces in Diagram (4.4).

Lemma 4.10. *The rank of the map $H_d(\text{Clq}_k(G_1, G_1[Q]) \hookrightarrow \text{Clq}_k(G_2, G_2[Q]))$ induced by inclusion is at most $\text{rk } i_*$.*

Proof. Corollary 4.9 with respect to Diagram (4.4) implies that $\text{rk}(\text{H}^d(P_k^\beta, Q_k^\beta) \rightarrow \text{H}^d(P_k^\alpha, Q_k^\alpha)) = \text{rk } i_*$. By the Universal Coefficient Theorem, we then have $\text{rk}(\text{H}_d(P_k^\alpha, Q_k^\alpha) \rightarrow \text{H}_d(P_k^\beta, Q_k^\beta)) = \text{rk } i_*$, and by the relative homology iteration of the Persistent Nerve Lemma for k -barycentric decompositions, Lemma 4.5, for all $\varepsilon \geq 0$, $\text{H}_*(P_k^\varepsilon, Q_k^\varepsilon) \cong \text{H}_*(\mathcal{C}_k^\varepsilon(P, Q))$, so $\text{rk}(\text{H}_d(\mathcal{C}_k^\alpha(P, Q) \hookrightarrow \mathcal{C}_k^\beta(P, Q))) = \text{rk } i_*$.

By Assumption 5, the inclusion $\text{Clq}_k(G_1, G_1[Q]) \hookrightarrow \text{Clq}_k(G_2, G_2[Q])$ can be factored as

$$\text{Clq}_k(G_1, G_1[Q]) \hookrightarrow \mathcal{C}_k^\alpha(P, Q) \hookrightarrow \mathcal{C}_k^\beta(P, Q) \hookrightarrow \text{Clq}_k(G_2, G_2[Q]).$$

It follows that

$$\begin{aligned} & \text{rk}(\text{H}_d(\text{Clq}_k(G_1, G_1[Q]) \hookrightarrow \text{Clq}_k(G_2, G_2[Q]))) \\ & \leq \text{rk}(\text{H}_d(\mathcal{C}_k^\alpha(P, Q) \hookrightarrow \mathcal{C}_k^\beta(P, Q))) \\ & = \text{rk } i_*. \end{aligned}$$

□

The next lemma states that there is an upper-bound on the number of connected components of the coverage region of interest, $\mathcal{D} \setminus \mathcal{B}^{2\alpha}$, if we assume that its connected components are not too close together.

Lemma 4.11. *Let $(\mathcal{D}, \mathcal{B})$ be a pair of spaces satisfying Assumptions 0 and 2 for positive constants α and $\beta \geq 3\alpha$. If $P \subset \mathcal{D}$, $Q = \{p \in P \mid \text{cov}(p, \alpha) \cap \mathcal{B} \neq \emptyset\}$ and the graph G_1 satisfied Assumptions 4 – 6 for some constant $k \geq 1$, then $\dim \text{H}_0(G_1[P \setminus Q]) \geq \dim \text{H}_0(\mathcal{D} \setminus \mathcal{B}^{2\alpha})$.*

Proof. Assume there exists $p, q \in P \setminus Q$ such that p and q are connected in $\text{Clq}(G_1[P \setminus Q])$, but not in $\mathcal{D} \setminus \mathcal{B}^{2\alpha}$. By Assumption 2, we have that $d(p, q) \leq 2\alpha$ and $[p] \neq [q]$ in $\text{H}_0(\mathcal{D} \setminus \mathcal{B}^{2\alpha})$. However, the shortest path $\overline{pq} \in (\mathcal{D} \setminus \mathcal{B})^{2\alpha}$, as the distance between p and q is less than 2α , so $[p] = [q]$ in $\text{H}_0(\mathcal{D}^{2\alpha})$, which implies that $\text{H}_0(\mathcal{D} \setminus \mathcal{B}^{2\alpha} \hookrightarrow (\mathcal{D} \setminus \mathcal{B})^{2\alpha})$ is not injective, a contradiction to Assumption 2. □

This leads to the Topological Coverage Criterion algorithm, Algorithm 1. Theorem 4.12 proves the validity of this simple algorithm that checks for k -coverage, or tells the user if more sensors are

Algorithm 1 Check if $\mathcal{D} \setminus \mathcal{B}^{2\alpha} \subseteq P_k^\alpha$

1: **procedure** K-COVERAGE(G_1, G_2, P, Q, k)

2: let $c := \dim H_0(G_1[P \setminus Q])$

3: let

$$r := \text{rk } H_d(\text{Clq}_k(G_1, G_1[Q]) \hookrightarrow \text{Clq}_k(G_2, G_2[Q]))$$

4: **if** $c = r$ **then return** True

5: **else return** False

necessary, i.e. the domain needs to be resampled in the language of geometric inference. The proof of this theorem uses the fact that we can factor the maps between the clique complexes through the Čech complexes associated to the pairs (P_k^α, Q_k^α) and (P_k^β, Q_k^α) by the input assumptions. Each collection of point sets forms a good cover by assumption, so the by the Nerve Theorem there are natural isomorphisms between the homology groups of the offset pairs and the simplicial complex pairs. The proof is completed by reasoning about the ranks of the homology maps in the corresponding sequence.

Theorem 4.12 (Algorithmic TCC). *Consider a pair $(\mathcal{D}, \mathcal{B})$, a finite point sample $P \subset \mathcal{D}$, and constants k, α, β , where $0 < 3\alpha \leq \beta$, satisfying Assumptions 0–7. If*

$$\text{rk } H_d(\text{Clq}_k(G_1, G_1[Q]) \hookrightarrow \text{Clq}_k(G_2, G_2[Q])) = \dim H_0(G_1[P \setminus Q])$$

then $\mathcal{D} \setminus \mathcal{B}^{2\alpha} \subseteq P_k^\alpha$.

Proof. For simplicity, define

$$a_* := H_d(\text{Clq}_k(G_1, G_1[Q]) \hookrightarrow \text{Clq}_k(G_2, G_2[Q]))$$

and set $c = |\text{Components}(G_1[P \setminus Q])|$, $m = H_0(D \setminus \mathcal{B}^{2\alpha})$. By our hypotheses and Lemma 4.10, $\text{rk } i_* \geq \text{rk } a_* = c$. By Lemma 4.11, $c \geq m$, and Assumption 1 implies that j_* is surjective by Lemma 4.3 so by definition of $\mathcal{B}$ surrounding $\mathcal{D}$, $m = \text{rk } j_*$. Thus $\text{rk } i_* \geq \text{rk } a_* = c \geq m = \text{rk } j_*$, namely $\text{rk } i_* \geq \text{rk } j_*$, so by Theorem 4.8 we can conclude $\mathcal{D} \setminus \mathcal{B}^{2\alpha} \subseteq P_k^\alpha$. $\square$

4.5 Certified Homology Inference

In this section we will show how the Algorithmic Topological Coverage Criterion leads to a *certified* homology inference method. This procedure removes the doubt in standard homology inference methods of verifying whether a sample in practice satisfies the assumptions in the theorems. To do so, we assume the sample provides coverage as verified by the Algorithmic/Geometric Topological Coverage Criterion. If the sample does not provide coverage, one can re-sample or add more sample points until it provides coverage.

Another motivation here is solving the problem of scale selection that arises in topological data analysis. If we have a sample of points, we can attempt to infer the homology of the underlying space by replacing the sample with a union of balls, but we don't know prior what should be the radii of the balls. If we have some prior idea of the number of connected components, we could look at the first radius for which the union of balls has the right number of components, but there is no reason to believe that this is scale will give the most accurate homological reconstruction in other dimensions without more additional knowledge of the domain. In the case of an existent underlying compact space whose homology we want to infer, there really ought to be right scale, or interval of scales, for a given sampling to correctly reconstruct the space's shape, assuming the sample is well taken. Our result is the construction of a a scale that meets this task, or a provision of some evidence that the sampling is insufficient to perform homology inference on the domain.

Here, in contrast to the previous sections in this chapter, we consider a pair of spaces $(\mathcal{D}, \mathcal{B})$ where $\mathcal{D}$ is a compact Euclidean domain and $\mathcal{B}$ is the topological boundary of $\mathcal{D}$. We also will assume the boundary's weak feature size is sufficiently large with respect to the sensing parameters for the main results. The main assumption of the sample P is that it satisfies the single-coverage condition $\mathcal{D} \setminus \mathcal{B}^{2\alpha} \subset P$ which can be verified by introduced algorithm.

Now assume the geometric assumptions 0, 1 & 2 in 4.3. For a constant $\beta \geq 0$, we define the

subset $U_\beta \subseteq P$ to be

$$U_\beta = P \setminus \mathcal{B}^\beta = \{p \in P \mid d(p, \mathcal{B}) > \beta\}.$$

These definitions lead to the following lemma relating a shrunken domain to a subsampling, given that we have coverage.

Lemma 4.13. *Given a domain and boundary $\mathcal{D}$ and $\mathcal{B}$ with point sampling P that satisfies the TCC, for all γ, β such that $\gamma, \beta \geq \alpha$, we have the following.*

$$U_{\beta+\gamma}^\gamma \subseteq \mathcal{D} \setminus \mathcal{B}^\beta \text{ and } \mathcal{D} \setminus \mathcal{B}^{\beta+\gamma} \subseteq U_\beta^\gamma.$$

Proof. For the first statement, if $x \in U_{\beta+\gamma}^\gamma$, then there exists $p \in P \setminus \mathcal{B}^{\beta+\gamma}$ such that $d(p, x) \leq \gamma$. Also, $d(p, \mathcal{B}) \geq \beta + \gamma$. Together this implies that $d(x, \mathcal{B}) \geq d(p, \mathcal{B}) - d(p, x) \geq \beta$, by the distance function being 1-Lipschitz. It then follows that $x \in \mathcal{D} \setminus \mathcal{B}^\beta$.

For the second statement, if $x \in \mathcal{D} \setminus \mathcal{B}^{\beta+\gamma}$, then $x \in \mathcal{D} \setminus \mathcal{B}^{2\alpha}$, as $\beta + \gamma \geq 2\alpha$. By our coverage assumption, $x \in P^\alpha$, so there exists $p \in P$ such that $d(x, p) \leq \alpha$. We then have that $d(p, \mathcal{B}) \geq d(x, \mathcal{B}) - d(p, x) \geq \beta + \gamma - \alpha \geq \beta$. Thus $p \in U_\beta$ and $x \in U_\beta^\alpha \subseteq U_\beta^\gamma$. $\square$

We can then rewrite these two statements in terms of the complement spaces as

$$\overline{\mathcal{D}}^\beta \subseteq \overline{U_{\beta+\gamma}^\gamma} \text{ and } \overline{U_\beta^\gamma} \subseteq \overline{\mathcal{D}}^{\beta+\gamma}.$$

These containments of complements result in a topological interleaving, to which we apply Alexander duality in order to prove the following lemma connecting the TCC with homology inference, assuming the criterion is satisfied.

Lemma 4.14. *Suppose the point sample $P \subset \mathcal{D}$ has the property $\mathcal{D} \setminus \mathcal{B}^{2\alpha} \subseteq P^\alpha$ as asserted by the TCC. Let $\beta, \gamma, \varepsilon, \delta$ be constants such that $\delta \geq \varepsilon \geq \gamma \geq \alpha$ and $\beta \geq \varepsilon + \delta + \gamma$. If $\text{wfs}(\mathcal{B}) > \beta + \gamma$, then*

$$\text{rk}(H_k(U_\beta^\gamma) \rightarrow H_k(U_\delta^\varepsilon)) = \dim(H_k(\mathcal{D})), \text{ for all integers } k.$$

Proof. Consider the interleaving of inclusions as follows.

$$\overline{\mathcal{D}} \hookrightarrow \overline{U_\delta^\varepsilon} \hookrightarrow \overline{\mathcal{D}^{\varepsilon+\delta}} \hookrightarrow \overline{\mathcal{D}^{\beta-\gamma}} \hookrightarrow \overline{U_\beta^\gamma} \hookrightarrow \overline{\mathcal{D}^{\beta+\gamma}}.$$

Given the weak feature size bound, the critical point homotopy result Theorem 2.7 implies that the inclusions between offsets of $\overline{\mathcal{D}}$ are all homotopy equivalences up to $\overline{\mathcal{D}^{\beta+\gamma}}$, as weak feature size is the minimum critical value for all critical points of the distance-to-set function. The inclusions above factor through the resulting isomorphisms between the offsets of the complements of the domain at the level of homology. Using Lemma 3.2 from Chazal and Oudot [17], we have that

$$\mathrm{rk}(\tilde{H}^{d-k-1}(\overline{U_\delta^\varepsilon}) \rightarrow \tilde{H}^{d-k-1}(\overline{U_\beta^\gamma})) = \dim(\tilde{H}^{d-k-1}(\overline{\mathcal{D}})).$$

The claimed result now follows from Alexander duality. $\square$

By switching the aforementioned result to the analogous one at the level of simplicial complexes, the Čech-Rips interleaving provides a method to infer the homology computationally by looking at the Rips complexes of the subsampled offsets as viewed as a cover of their union. Note that if there is a deformation retraction $\mathcal{D}^\lambda \rightarrow \mathcal{D}$ for some $\lambda \in (0, \mathrm{wfs}(\mathcal{D}))$, then the result holds for $\lambda = 0$.

Theorem 4.15 (Certified Homology Inference). *Let $\mathcal{D} \subset \mathbb{R}^d$ be a compact space with boundary $\mathcal{B}$ such that $\mathrm{wfs}(\mathcal{B}) > (2\vartheta_d^2 + 4\vartheta_d + 2)\alpha$. Let $P \subset \mathcal{D}$ be such that $\mathcal{D} \setminus \mathcal{B}^{2\alpha} \subseteq P^\alpha$ as asserted by the TCC. Then we have the following*

$$\mathrm{im}(H_k(\mathcal{R}^\alpha(U_{(2\vartheta_d^2+4\vartheta_d+1)\alpha}) \rightarrow \mathcal{R}^{\vartheta_d\alpha}(U_{(2\vartheta_d^2+\vartheta_d)\alpha})) \cong H_k(\mathcal{D}), \text{ for all } k.$$

Proof. We have the following sequence of inclusions of Čech complexes, where the scales are chosen to correspond with those in the inclusions of the form $U_\beta^\gamma \hookrightarrow U_\delta^\varepsilon$ in Lemma 4.14, as to satisfy its hypotheses.

$$\mathcal{C}^\alpha(U_{(2\vartheta_d^2+4\vartheta_d+1)\alpha}) \hookrightarrow \mathcal{C}^{\vartheta_d\alpha}(U_{(2\vartheta_d^2+3\vartheta_d)\alpha}) \hookrightarrow \mathcal{C}^{\vartheta_d\alpha}(U_{(2\vartheta_d^2+\vartheta_d)\alpha}) \hookrightarrow \mathcal{C}^{\vartheta_d^2\alpha}(U_{\vartheta_d^2\alpha}).$$

By the Nerve Theorem, for all constants $\gamma, \delta \geq 0$, $H_*(U_\delta^\gamma) \cong H_*(\mathcal{C}^\gamma(U_\delta))$, and by the Persistent Nerve Lemma, each of the inclusions induces natural isomorphisms at the homology level. By

factoring through the Rips complexes using the Jung's Theorem interleaving we have the following interleaving.

$$\begin{aligned} \mathcal{C}^\alpha(U_{(2\vartheta_d^2+4\vartheta_d+1)\alpha}) &\hookrightarrow \mathcal{R}^\alpha(U_{(2\vartheta_d^2+4\vartheta_d+1)\alpha}) \hookrightarrow \mathcal{C}^{\vartheta_d\alpha}(U_{(2\vartheta_d^2+3\vartheta_d)\alpha}) \hookrightarrow \\ &\mathcal{C}^{\vartheta_d\alpha}(U_{(2\vartheta_d^2+\vartheta_d)\alpha}) \hookrightarrow \mathcal{R}^{\vartheta_d\alpha}(U_{(2\vartheta_d^2+\vartheta_d)\alpha}) \hookrightarrow \mathcal{C}^{\vartheta_d^2\alpha}(U_{\vartheta_d^2\alpha}). \end{aligned}$$

The constants in Lemma 4.14 translate to this interleaving as $\gamma = \alpha$, $\varepsilon = \vartheta_d^2\alpha$, $\beta = (2\vartheta_d^2 + 4\vartheta_d + 1)\alpha$, and $\delta = \vartheta_d^2\alpha$. Now by applying the homology functor and again Lemma 3.2 from Chazal and Oudot [17], we prove the desired result. $\square$

Although the indices in the above interleaving look complicated, we note that Theorem 4.15 implies that as long as $\text{wfs}(\mathcal{B}) > (6 + 4\sqrt{2})\alpha$, we can infer the Betti numbers of $\mathcal{D}$ from the rank of the maps on homology induced by the inclusion $\mathcal{R}^\alpha(U_{(5+4\sqrt{2})\alpha}) \hookrightarrow \mathcal{R}^{\sqrt{2}\alpha}(U_{(4+\sqrt{2})\alpha})$. This is due to the asymptotic bound of ϑ_d as $\sqrt{\frac{2d}{d+1}} < \sqrt{2}$, so we need the weak feature size to be roughly 12 times greater than the scale α and we need to detect $\mathcal{B}$ within a distance of 5.5 to 11 times α .

Theorem 4.15 in conjunction with the Algorithmic TCC gives a repeatable homology inference algorithm by varying the input based on the output. If the input for a given domain, boundary, and point sample does not provide coverage, then one can resample and check the TCC again. If it does provide coverage that is sufficient enough for the subsampling scheme above to correctly infer the homology of the domain.

Chapter 5

The Generalized Persistent Nerve Theorem

5.1 Introduction

In this section we extend our focus beyond homology inference of single topological spaces to that of filtrations of simplicial complexes — the goal being to approximate the so-called persistent homology of a growing sequence of triangulations, or simplicial complexes. We will provide an upper bound on the difference between its persistent homology and that of a collection of nerves of the covers, in the process generalizing a foundational result in topological data analysis.

Nerves are central to topological data analysis as they assign discrete structures, simplicial complexes, to topological covers. For example, the Delauney triangulation is the nerve of the Voronoi cell decomposition, and both Čech complexes and Rips complexes are instances of nerves. The Nerve Theorem is a result in algebraic topology [6] that equates the homotopy type of a space covered by sets to a simplicial complex derived from the cover called the nerve derived which has k -simplices for each k -wise intersections between cover elements. Nerves have a role in many algorithms in surface reconstruction, homology inference, clustering, and homological sensor networks, among

others, due to this theorem.

Persistent homology, first formulated by Edelsbrunner et al. [27], is the study of the changes in the homology of the components of a filtration under the inclusion maps. Examining geometric data at a variety of scales can shed light on geometric and topological features that are not apparent upon first glance. Since its inception, a rich theory has developed around persistent homology [14, 26, 43] utilizing tools and structures from algebraic topology [35] and category theory [39]. Collectively the homology of a filtration is represented geometrically by a persistence diagram, a multiset of the plane with each feature’s coordinates being its birth and death time, or algebraically as the direct sum of barcodes, intervals on the real line [8]. The Persistent Equivalence Theorem [30] states that if two filtrations have isomorphic persistence modules, then the two filtrations have the same persistence diagrams. Analogously, the Persistent Nerve Lemma [17] states that if one has a so-called good cover of a filtration of paracompact topological spaces, then the corresponding filtration of the nerves has identical persistence diagrams as the original filtration.

Our result extends upon the Persistent Nerve Lemma to more general parametrized good covers, which we call ε -good covers. These are covers in which the image of the homology of the intersection of the cover elements is trivial ε time later. These covers can appear naturally in cases where there are discretization errors in constructing a triangulation of a space. The main result, which we call the Generalized Persistent Nerve Theorem, is that the bottleneck distance of the K -dimensional persistence diagrams of an ε -good covered simplicial filtration and the nerve filtration is bounded above by $(K + 1)\varepsilon$, and this bound is tight. The simplicial Persistent Nerve Lemma is a direct corollary making it a strict generalization for these spaces. The Generalized Persistent Nerve Theorem is an approximation theorem in contrast to the equivalence theorem of the Persistent Nerve Lemma, which can naturally be compared to the relationship between the Algebraic Stability Theorem and the Persistence Equivalence Theorem.

5.1.1 Related Work

Approximating persistent homology has been studied before. Two filtrations' persistence diagrams are compared by the bottleneck distance metric, a maximum matching distance, which is equivalent to the ∞ -Wasserstein distance and is related to the Earth-mover's distance on graphs. This metric is stable with respect to perturbations of the generating function of the filtration [20], possibly resulting from imprecision of measurements. The original (geometric) stability result has been generalized by considering interleavings between filtrations, allowing one to discuss topological approximations of filtrations. Explicitly, the Algebraic Stability Theorem [12] states that if two persistence homology are interleaved in some distance, then the bottleneck distance between their diagrams is bounded by that same quantity. Botnan and Spremann [7] proved that the bound on the bottleneck distance between the persistence diagrams of two nerve filtrations is bounded, with the assumption that there is a triple interleaving between three filtrations of covers.

Govc and Skraba worked on the problem of generalizing the Persistent Nerve Theorem in parallel to us, culminating in their paper *An Approximate Nerve Theorem* [31]. They assumed that the reduced homology of the k -wise intersections of a simplicial cover filtrations' cover elements are ε -interleaved with the 0 module, a condition they call an ε -acyclic cover. This condition is the same as ours although expressed in the language of persistence theory and they proved an identical bound. Though the result is the same, the approaches are very different. Govc and Skraba utilized a construction from homological algebra called a spectral sequence, as well as their novel right and left persistence module interleavings to prove the theorem, which is actually a corollary of the module interleavings they compute between the pages of the spectral sequences and the persistence modules of the nerve and space filtrations.

In contrast, our proof technique focuses on constructing maps between the chain complexes of the nerve filtration, space filtration, and the barycentric subdivision of the so-called blow-up complex. We go from the homological ε -goodness condition to corresponding chain maps and chain homotopies between them among the chain complexes of the spaces of interest. This is

more aligned with the approach traditionally used to proof the Nerve Theorem and the homotopy iteration of the Persistent Nerve Lemma. Due to the fact that the chain maps are defined on chain complexes of simplicial complexes and regular CW-complexes, they are computable in practice. Our module interleaving and bottleneck distance bound results are a consequence of a chain-theoretic generalization of an interleaving at the level of chain homotopy, rather than purely acting on persistence modules. There are also the novel contributions of the chain map between the barycentric subdivision of the nerve of a cover and the space filtration at a further time scale, and the use a chain complex technique we call a *lifting* to form a chain map into the barycentric subdivision of the blowup complex. The lack of existence of such a map is the reason the Nerve Theorem fails when the cover is not good, which gives credence to the notion that the maps we construct are “natural” choices.

5.2 Definitions

5.2.1 CW-Complexes

Although the filtrations considered in this work will consist of simplicial complexes, CW-complexes and cellular homology will be necessary as the product of simplicial complexes is not a simplicial complex except in degenerate cases.

A **CW-complex** is a topological space that is the next natural generalization of a simplicial complex, relaxing the simplicial “gluing” procedure. To start, an n -cell is a space that homeomorphic to the unit ball in $\mathbb{R}^n$. From this, a CW-complex is defined inductively. Starting with a collection X_0 of 0-cells, or vertices in simplicial terms, and then defining X_k as the union of X_{k-1} and a collection of k -cells whose boundaries are glued via continuous maps, attaching maps, to the $(k-1)$ -cells of X_{k-1} . Collectively the complex is the union of all the cells with the appropriate gluings. Such a space is equipped with the quotient topology induced by the equating the k -cells’ boundaries with their images under the attaching maps.

For example, the n -sphere can be given the structure of a CW-complex by constructing it as a single 0-cell along with a single n -cell (a n -dimensional disc) whose boundary is attached to the 0-cell via the constant map. Any simplicial complex can be interpreted as a CW-complex via the same gluing procedure used to define simplicial complexes.

Given two finite CW-complexes X and Y , the product space $X \times Y$ is a finite CW-complex with cells of the form $e_X \times e_Y$, where $e_X \in X$ and $e_Y \in Y$, and $\dim(e_X \times e_Y) = \dim(e_X) + \dim(e_Y)$. As finite simplicial complexes are finite CW-complexes, the product of two finite simplicial complexes is a finite CW-complex as well.

5.2.2 Chain Homotopy and Cellular Homology

Recall the definition of a chain map — Definition 2.14 in Subsection 2.2.2. A **chain homotopy** between two chain maps $f, g : C_*(X) \rightarrow C_*(Y)$ is a collection of homomorphisms $c = (c_n)_{n \geq 0}$, where $c_n : C_n(X) \rightarrow C_{n+1}(Y)$, such that the following holds for all $n \geq 0$:

$$c_{n-1}\partial_n + \partial_{n+1}c_n = f_n - g_n,$$

or $c\partial^X + \partial^Y c = f - g$ for short, i.e. the following diagram commutes.

$$\begin{array}{ccccccc} \dots & \longrightarrow & C_{n+1}(X) & \xrightarrow{\partial_{n+1}^X} & C_n(X) & \xrightarrow{\partial_n^X} & C_{n-1}(X) \longrightarrow \dots \\ & & \downarrow f_{n+1}, g_{n+1} & \swarrow c_n & \downarrow f_n, g_n & \swarrow c_{n-1} & \downarrow f_{n-1}, g_{n-1} \\ \dots & \longrightarrow & C_{n+1}(Y) & \xrightarrow{\partial_{n+1}^Y} & C_n(Y) & \xrightarrow{\partial_n^Y} & C_{n-1}(Y) \longrightarrow \dots \end{array}$$

Given a CW-complex X , we define the **cellular chain complex**

$$C_*^{\text{CW}}(X) := (C_n^{\text{CW}}(X))_{n \geq 0},$$

where $C_n^{\text{CW}}(X)$ is naturally isomorphic to a $\mathbb{F}_2$ -vector space with basis consisting of the n -cells of X . Formally, $C_n^{\text{CW}}(X) := H_n(X_n, X_{n-1})$. For each basis n -cell e_i^n in $C_n^{\text{CW}}(X)$, the **cellular boundary map** is

$$\partial_k(e_i^n) := \sum_j d_{i,j} e_j^{n-1},$$

where $d_{i,j}$ is computed by the cellular boundary formula (see page 140 in Hatcher[35] for the exact formulation). This extends linearly to the arbitrary n -chains just like in the simplicial case. When X is a regular CW-complex, $d_{i,j}1$ for all i, j , and importantly all products of simplicial complexes are regular.

A cellular map $f : X \rightarrow Y$, a continuous function between the CW-complexes such that $f(X_n) \subseteq Y_n$ for all $n \geq 0$, induces a cellular chain map $f : C_*^{\text{CW}}(X) \rightarrow C_*^{\text{CW}}(Y)$. Simplicial maps are cellular when the simplicial complexes are interpreted as CW-complexes. Given two finite CW-complexes X and Y , there is a natural isomorphism

$$C_*^{\text{CW}}(X \times Y) \cong C_*^{\text{CW}}(X) \otimes C_*^{\text{CW}}(Y),$$

where $\otimes$ is the tensor product. It follows that the following equality holds,

$$C_n^{\text{CW}}(X \times Y) \cong \bigoplus_{p+q=n} C_p^{\text{CW}}(X) \otimes C_q^{\text{CW}}(Y).$$

The basis of $C_n^{\text{CW}}(X \times Y)$ as the collection of products $e^p \times f^q$, where e^p is a p -cell in X and f^q is a q -cell in Y for $p + q = n$. Therefore, one can naturally express the basis also as all $e^p \otimes f^q \in C_p^{\text{CW}}(X) \otimes C_q^{\text{CW}}(Y)$. The product chain complex $C_*^{\text{CW}}(X \times Y)$ has a boundary map which is defined on any basis k -chain $\sigma \otimes \tau \in C_k(X \times Y)$ as

$$\partial_n^{X \times Y}(\sigma \otimes \tau) := \partial_n^X(\sigma) \otimes \tau + \sigma \otimes \partial_n^Y(\tau) \in C_{n-1}^{\text{CW}}(X \times Y),$$

and this extends linearly to all n -chains.

The cellular homology groups $H_*^{\text{CW}}(X)$ are defined in the same manner as the simplicial homology groups except instead with cellular chain complex groups and boundaries. In fact, for a finite simplicial complex X , $H_*^{\text{CW}}(X) = H_*(X)$. This is clear by the fact that the cellular and simplicial chain complexes are canonically isomorphic — the n -simplices are the n -cells when X is viewed as a CW-complex in the obvious way.

5.2.3 Persistent Homology

A **filtration** is a sequence of topological spaces $\mathcal{F} = (F^\alpha)_{\alpha \geq 0}$ such that $F^\alpha \subseteq F^\beta$ if and only if $\alpha \leq \beta$. We call a filtration consisting of simplicial complexes a simplicial filtration, and if each simplicial complex is defined over a finite vertex set it is called a finite simplicial filtration. A sequence of subdivisions of a triangulation of a space is a simplicial filtration for example.

Persistent homology is the changes in the homology of a filtration as it ranges over the interval $[0, \infty)$. To be precise, it is the computation of the “birth” and “death” scales of homological features under the homology maps induced by inclusion, $i_{\mathcal{F}}^{\alpha, \beta} : H_*(F^\alpha \hookrightarrow F^\beta)$, for all α, β such that $\alpha \leq \beta$. The persistent homology data of a filtration is contained in its **persistence module**, denoted by $H_*(\mathcal{F})$, which consists of the spaces $H_*(F^\alpha)$ over all scales, and the aforementioned maps $i_{\mathcal{F}}^{\alpha, \beta}$ for $\alpha \leq \beta$. The birth and deaths scales of k -dimensional homological features in a filtration $\mathcal{F}$ are represented in a filtration’s k -dimensional **persistence diagram**, denoted by $\text{Dgm}_k(\mathcal{F})$. This is a multiset with elements being points in the plane (x, y) , where x and y are the birth and death scales respectively of features, and (x, x) for all $x \in \mathbb{R}$ with infinite multiplicity. When discussing the persistence diagrams collectively for all dimensions we denote $\text{Dgm}(\mathcal{F}) := (\text{Dgm}_k(\mathcal{F}))_{k \geq 0}$. The following theorem provides a condition under which we can say that two filtrations have identical persistence diagrams, often called the Persistence Equivalence Theorem — see chapter 26 of [30].

Theorem 5.1 (Persistence Equivalence Theorem). *Consider two filtrations, $\mathcal{F} = (F^\alpha)_{\alpha \geq 0}$ and $\mathcal{G} = (G^\alpha)_{\alpha \geq 0}$, with point-wise finite dimensional persistence modules. If their persistence modules $H_*(\mathcal{F})$ and $H_*(\mathcal{G})$ are isomorphic then the filtrations have identical persistent homology and $\text{Dgm}(\mathcal{F}) = \text{Dgm}(\mathcal{G})$.*

Observe that $H_*(\mathcal{F})$ is isomorphic to $H_*(\mathcal{G})$ if and only if there are natural isomorphisms component-wise between the modules at all scales, i.e. $H_*(F^\alpha) \cong H_*(G^\alpha)$.

A persistence diagram is finite if it has finitely many off-diagonal points. The standard metric on the space of persistence diagrams is the bottleneck distance d_B , which is efficiently computable

for finite diagrams. For two finite diagrams D and D' , the **bottleneck distance** between them is defined as

$$d_B(D, D') := \min_{\phi \in \Phi} \max_{p \in D} \|p - \phi(p)\|_\infty,$$

where Φ is the set of all bijections $\phi : D \rightarrow D'$. Two finite persistent diagrams are equivalent if and

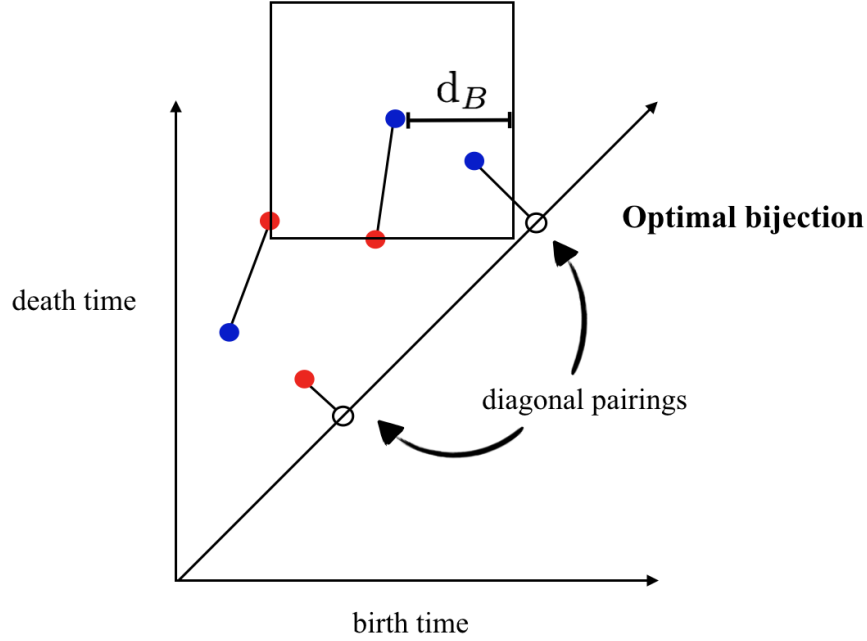


Figure 5.1: The optimal bijection between two finite persistence diagrams. The bottleneck distance is the minimum radius of sup-norm balls required to cover the points in the other persistence diagram.

only if the bottleneck distance between them is 0. A major result in topological data analysis is that the bottleneck distance is stable with respect to perturbations of the function generating the diagram, which is known as the Stability Theorem [20]. This perturbation can be generalized algebraically as follows.

Definition 5.2. Two filtrations $\mathcal{F} = (F^\alpha)_{\alpha \geq 0}$ and $\mathcal{G} = (G^\alpha)_{\alpha \geq 0}$, with persistence modules $H_*(\mathcal{F})$ and $H_*(\mathcal{G})$ respectively, are **δ -interleaved** if there exists collections of homomorphisms $f = (f^\alpha)_{\alpha \geq 0}$

and $g : (g^\alpha)_{\alpha \geq 0}$, where $f^\alpha : H_*(F^\alpha) \rightarrow H_*(G^{\alpha+\delta})$ and $g^\alpha : H_*(G^\alpha) \rightarrow H_*(F^{\alpha+\delta})$, such that

$$g^{\alpha+\delta} f^\alpha = i_{\mathcal{F}}^{\alpha, \alpha+2\delta}$$

and

$$f^{\alpha+\delta} g^\alpha = i_{\mathcal{G}}^{\alpha, \alpha+2\delta},$$

and these maps commute with all $i_{\mathcal{F}}^{\alpha, \beta}$ and $i_{\mathcal{G}}^{\alpha, \beta}$ for all $\alpha \leq \beta$.

In the literature what we call a δ -interleaving is also known as an additive interleaving, in contrast to a multiplicative interleaving where the indices of the interleaving are multiplied by a constant factor. In the language of Definition 3.11 in Chapter 3, this definition coincides with a $(+\delta, +\delta)$ -interleaving between $\mathcal{F}$ and $\mathcal{G}$ on $[0, \infty)$. Persistence module interleavings and their persistence diagrams' bottleneck distances are related by the Algebraic Stability Theorem (see [12, Thm 4.4]),

Theorem 5.3 (Algebraic Stability Theorem). *Given two filtrations $\mathcal{F} = (F^\alpha)_{\alpha \geq 0}$ and $\mathcal{G} = (G^\alpha)_{\alpha \geq 0}$ such that for all $\alpha \geq 0$, $\dim H_*(F^\alpha), \dim H_*(G^\alpha) < \infty$, if $H_*(\mathcal{F})$ and $H_*(\mathcal{G})$ are δ -interleaved then $d_B(Dgm_k(\mathcal{F}), Dgm_k(\mathcal{G})) \leq \delta$ for all k .*

5.2.4 Nerve and Cover Filtrations

Let $\mathcal{U} := \{U_0, \dots, U_n\}$ be a collection of simplicial filtrations, where $U_i := (U_i^\alpha)_{\alpha \geq 0}$ and for all i and α , U_i^α is a finite subcomplex of the same simplicial complex. Define $\mathcal{U}^\alpha := \{U_0^\alpha, \dots, U_n^\alpha\}$ — the collection of simplicial complexes at scale α from each filtration U_i . For each non-empty $v \subseteq [n]$, let $U_v^\alpha := \bigcap_{i \in v} U_i^\alpha$, which results in a simplicial filtration $U_v := (U_v^\alpha)_{\alpha \geq 0}$. Note that in this notation $U_{\{i\}}^\alpha = U_i^\alpha$ and $U_{\{i\}} = U_i$. The **nerve filtration** for a collection of filtrations $\mathcal{U}$ is defined as

$$\text{Nrv } \mathcal{U} := (\text{Nrv } (\mathcal{U}^\alpha))_{\alpha \geq 0}.$$

For each scale the union of over the elements of $\mathcal{U}^\alpha$ is the simplicial complex defined as $W^\alpha := \bigcup_{i=0}^n U_i^\alpha$ and the **union filtration** with respect to $\mathcal{U}$ is denoted by $\mathcal{W} := (W^\alpha)_{\alpha \geq 0}$. For each W^α , $\mathcal{U}^\alpha$ is a

cover and thus we say that $\mathcal{U}$ is a cover filtration of $\mathcal{W}$, or cover for short. We call $\mathcal{U}$ a **good cover filtration**, or good cover for short, of $\mathcal{W}$ if $\mathcal{U}^\alpha$ is a good cover of W^α for all $\alpha \geq 0$.

The previous definitions lead to the statement of the Persistent Nerve Lemma, which the main theorem of this paper generalizes. The following lemma was originally formulated by Frédéric Chazal and Steve Oudot in [17] as a generalization of the Nerve Theorem to filtrations.

Lemma 5.4. *Let $X \subseteq X'$ be two finite simplicial complexes with good covers, $\mathcal{V} = \{V_\alpha\}_{\alpha \in A}$ and $\mathcal{V}' = \{V'_\alpha\}_{\alpha \in A}$ respectively, such that $V_\alpha \subseteq V'_\alpha$ for all $\alpha \in A$. There exists homotopy equivalences $|Nrv \mathcal{V}| \rightarrow X$ and $|Nrv \mathcal{V}'| \rightarrow X'$ that commute with the topological inclusions $X \hookrightarrow X'$ and $|Nrv \mathcal{V}| \hookrightarrow |Nrv \mathcal{V}'|$.*

Viewing each $\mathcal{U}^\alpha$ from a good cover filtration $\mathcal{U}$ as a good cover of W^α , Theorem 5.1 can be applied to the construction of the homotopy equivalences in the proof of Lemma 5.4 to achieve the following fundamental persistent homology result.

Theorem 5.5 (Persistent Nerve Lemma). *Given a collection of finite simplicial filtrations $\mathcal{U}$ where $\mathcal{U}$ is a good cover filtration of $\mathcal{W}$, then $Dgm(Nrv \mathcal{U}) = Dgm(\mathcal{W})$.*

5.3 ε -Good Covers

In the Persistent Nerve Lemma a primary assumption is that $\mathcal{U}^\alpha := \{U_0^\alpha, \dots, U_n^\alpha\}$ is a good cover of $W^\alpha := \bigcup_{i=0}^n U_i^\alpha$ for all α , i.e. $\mathcal{U}$ is a good cover filtration of $\mathcal{W}$. However, there are common situations where a simplicial cover may not be good as shown in the figures below.

In these cases the good cover violations are relatively small and intuitively one should be able to ignore them without greatly affecting the accuracy of the homology of the shape. Persistent homology is the ideal theory for quantifying what is meant by a small violation of the good cover condition. The following is our generalization of a good cover filtration.

Definition 5.6. *A cover filtration $\mathcal{U}$ is an ε -good cover of a filtration $\mathcal{W} = (\bigcup_{i=0}^n U_i^\alpha)_{\alpha \geq 0}$ if for*

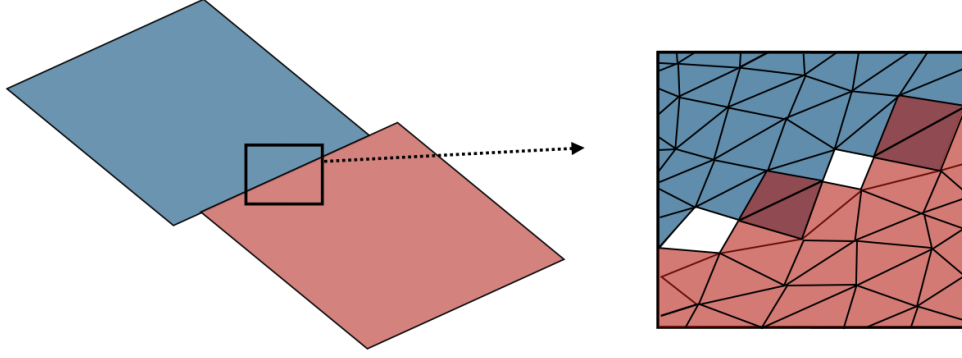


Figure 5.2: On the left is a space covered by two triangulated planes. By zooming in we see their intersection is not contractible by a small margin potentially due to an approximation error, which can be rectified by further sampling and triangulating.

all non-empty $v \subseteq [n]$, and all $\alpha \geq 0$,

$$\tilde{H}_*(U_v^\alpha \hookrightarrow U_v^{\alpha+\varepsilon}) = 0.$$

Although the definition of an ε -good covers is stated in terms of homology of the inclusions of cover intersections, it is in fact weaker than the assumption that $\mathcal{U}_v^\alpha \hookrightarrow U_v^{\alpha+\varepsilon}$ is null-homotopic, which would tie in neater with the traditional notion of a good cover. We choose to still use the term “good” despite this choice. There is the nice relation between the definitions of an ε -good cover and a good cover. If $\mathcal{U}$ is a good finite simplicial cover then for each non-empty intersection of cover elements U_v^α of $\mathcal{U}^\alpha$, U_v^α is homotopy equivalent to a point. This implies that $\text{im } \tilde{H}_*(U_v^\alpha \hookrightarrow U_v^{\alpha+\varepsilon}) \subseteq \tilde{H}_*(U_v^{\alpha+\varepsilon}) = 0$, so $\mathcal{U}$ is a 0-good cover. However, the converse does not hold — the 2-skeleton of the Poincaré 3-sphere has trivial reduced homology groups but is not contractible.

Recall the definition of the nerve filtration $\text{Nrv } \mathcal{U} := (\text{Nrv } \mathcal{U}^\alpha)_{\alpha \geq 0}$. The theorem we will prove

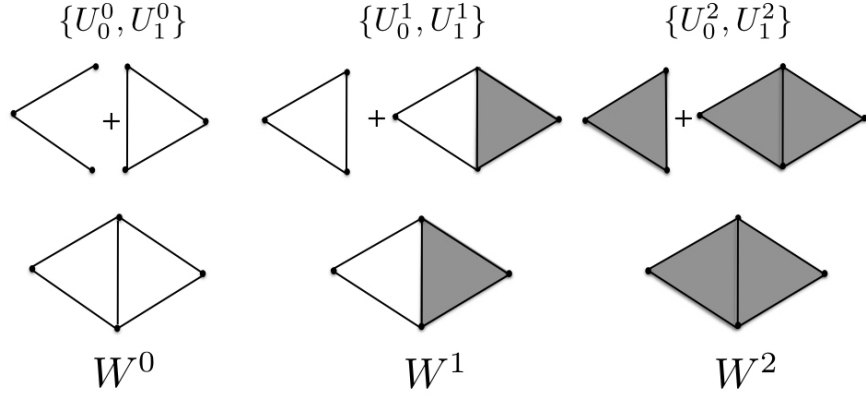


Figure 5.3: The simplicial cover filtration is not a good cover at scales 0 and 1, but is 1-good. The images of the inclusions from each intersection of covers has trivial homology at the next scale.

in this next section, our Generalized Persistent Nerve Theorem, provides a tight bound of $(K + 1)\varepsilon$ on the bottleneck distance between the K -dimensional persistence diagrams of the nerve filtration and a simplicial cover filtration, given that $\mathcal{U}$ is an ε -good cover filtration of $\mathcal{W}$.

Theorem 5.7 (Generalized Persistent Nerve Theorem). *Given a finite collection of finite simplicial filtrations $\mathcal{U} = \{U_0, \dots, U_n\}$, where $U_i := (U_i^\alpha)_{\alpha \geq 0}$ and all U_i^α are subcomplexes of a sufficiently large simplicial complex, if $\mathcal{U}$ is an ε -good cover filtration of $\mathcal{W} := (\bigcup_{i=0}^n U_i^\alpha)_{\alpha \geq 0}$, then*

$$d_B(Dgm_K(\mathcal{W}), Dgm_K(Nrv \mathcal{U})) \leq (K + 1)\varepsilon.$$

As Theorem 5.7 is true for all dimensions it implies that $d_B(Dgm(\mathcal{W}), Dgm(Nrv \mathcal{U})) \leq (D + 1)\varepsilon$ for $D := \dim Nrv \mathcal{U}$, as well as implying Theorem 5.5 (the Persistent Nerve Lemma) for the case of finite simplicial filtrations. See Subsection 5.5.1 for a construction that realizes the bottleneck distance bound over all dimensions.

The Generalized Persistent Nerve Theorem can be seen as an extension of the Persistent Nerve Lemma analogous to how the Algebraic Stability Theorem for persistence modules extends the Persistence Equivalence Theorem to interleaved modules, by viewing ε -good cover filtrations as

perturbations or approximations of good cover filtrations, the ideal object. This relationship is summarized in the following table.

	<i>Equivalence</i>	<i>Approximation</i>
<i>Persistence Modules</i>	Persistence Equivalence Theorem	Stability Theorem
<i>Nerves</i>	Persistent Nerve Lemma	Gen. Persistent Nerve Theorem

5.4 Building the Interleaving

For the duration of the construction, we will consider an arbitrary cover filtration $\mathcal{U} := \{U_0, \dots, U_n\}$ consisting of finite simplicial filtrations, where $U_i := (\bigcup_{\alpha \geq 0} U_i^\alpha)_{\alpha \geq 0}$ and all the U_i^α for each $i \in [n]$ and $\alpha \geq 0$ are defined over the same vertex set. We assume that $\mathcal{U}$ is an ε -good cover filtration of the simplicial filtration $\mathcal{W} := (W^\alpha)_{\alpha \geq 0} := (\bigcup_{i=0}^n U_i^\alpha)_{\alpha \geq 0}$. Fix a dimension K . For the remainder of the proof K will be the maximal dimension considered when discussing chain complexes, i.e. $C_*(X) := (C_k(X))_{k \leq K}$ for any space X , and likewise for homology groups.

The procedure to prove Theorem 5.7 is as follows. First we construct a diagram of chain complexes and chain maps that yield a $(K+1)\varepsilon$ -interleaving between the filtered chain complexes $C_k(\mathcal{W})$ and $C_k(\text{Nrv } \mathcal{U})$ for all $k \leq K$. This chain complex interleaving is analogous to one between persistence diagrams, except we only require the appropriate compositions be chain homotopic to the identity chain maps rather than equivalent. By applying homology to this chain complex diagram, the chain maps that are chain homotopic to the identity become equivalent to identity maps so there is a $(K+1)\varepsilon$ -interleaving between the persistence modules $H_k(\mathcal{W})$ and $H_k(\text{Nrv } \mathcal{U})$ for each $k \leq K$. The theorem is then proved by applying the Algebraic Stability Theorem (Theorem 5.3).

5.4.1 Homotopy Colimits

First we will introduce a mathematical construction that will be used throughout. Define Δ as a finite acyclic directed graph, or equivalently a strict poset. It can be given the structure of an abstract simplicial complex by defining k -simplices are as sequences $v_0 \rightarrow \dots \rightarrow v_k$, where $v_i \rightarrow v_{i+1}$ for all $i \in [k]$.

Consider a functor, known as a diagram, $D : \Delta \rightarrow \text{Top}$, where Top is the category of topological spaces. The **homotopy colimit** of D is defined as

$$\text{hocolim } D := \left(\bigsqcup_{\substack{\Delta \ni \sigma = v_0 \rightarrow \dots \rightarrow v_k \\ k \geq 0}} (D(v_0) \times |\sigma|) \right) / \sim,$$

where the disjoint union is taken over all simplices $\sigma \in \Delta$ and the homotopy colimit is given the quotient topology. The equivalence relation $\sim$, describing the gluing procedure along the boundaries of the k -simplices of Δ , is defined as follows. Given a k -simplex, consider the subsimplices of codimension 1, $\tau_i = v_0 \rightarrow \dots \rightarrow \hat{v}_i \rightarrow \dots \rightarrow v_k$, for each i , and their geometric realizations $|\tau_i|$. For all i , let $f_i : |\tau_i| \hookrightarrow |\sigma|$ be the i -th face map. For $i > 0$ we have the following equivalences

$$D(v_0) \times |\tau_i| \ni (x, y) \sim (x, f_i(y)) \in D(v_0) \times |\sigma|,$$

and for $i = 0$ we have the following equivalences

$$D(v_1) \times |\tau_0| \ni (D(v_0 \rightarrow v_1)(x), y) \sim (x, f_0(y)) \in D(v_0) \times |\sigma|.$$

Refer to p. 262 of Kozlov [38] for more information regarding this general construction or Welker et al.'s treatise on homotopy colimits and their applications [48].

5.4.2 Nerve Diagrams

For each $\alpha \geq 0$, define N^α as the directed graph with vertex set $\{\text{non-empty } v \subseteq [n] \mid U_v^\alpha := \bigcap_{i \in v} U_i^\alpha \neq \emptyset\}$, and directed edges (v', v) for any non-empty $v, v' \subseteq [n]$ such $v \subset v'$. Note the vertices

of this graph are in correspondence with the simplices of $\text{Nrv } \mathcal{U}^\alpha$ and thus form the 0-skeleton of the barycentric subdivision of $\text{Nrv } \mathcal{U}^\alpha$, while the (undirected) graph is its 1-skeleton. The edges correspond to inclusions between intersections of cover elements of the form $U_{v'}^\alpha \hookrightarrow U_v^\alpha$. We equip N^α with the structure of an abstract simplicial complex, where the k -simplices are sequences of vertices $v_0 \rightarrow \dots \rightarrow v_k$ such that for i and j , $0 \leq i < j \leq k$, there exists an edge (v_j, v_i) . From now on we will use the notation N^α to refer to the simplicial complex interpretation. A fact that will be important later is that its geometric realization, $|N^\alpha|$, is homeomorphic as a topological space to $|\text{Nrv } \mathcal{U}^\alpha|$.

For each $\alpha \geq 0$, define the nerve diagram as the diagram functor $D^\alpha : N^\alpha \rightarrow \text{Top}$ such that for each 0-simplex $v \in N^\alpha$, $D^\alpha(v) := \bigcap_{i \in v} U_i^\alpha$ and for each 1-simplex $v \rightarrow v'$, $D^\alpha(v \rightarrow v') := U_v^\alpha \hookrightarrow U_{v'}^\alpha$. Now we prove that the barycentric decomposition of the blow-up complex B^α is the homotopy colimit of the nerve diagram D^α .

Proposition 5.8. *For the nerve diagrams $D^\alpha : N^\alpha \rightarrow \text{Top}$, we have the following homotopy colimits for each $\alpha \geq 0$.*

$$\text{hocolim } D^\alpha = \bigcup_{\substack{N^\alpha \ni \sigma = v_0 \rightarrow \dots \rightarrow v_k \\ k \geq 0}} U_{v_0}^\alpha \times |\sigma|.$$

Proof. The equivalence relation simplifies as follows. First have $U_{v_0}^\alpha \times |\tau_i| \ni (x, y) \sim (x, y) \in U_{v_0}^\alpha \times |\sigma|$, for all $(x, y) \in U_{v_0}^\alpha \times |\tau_i|$, so for each boundary element τ_i for $i > 0$, one glues $U_{v_0}^\alpha \times |\tau_i|$ to its image under the component wise inclusion map, yielding $U_{v_0}^\alpha \times |\sigma| = ((U_{v_0}^\alpha \times |\tau_i|) \sqcup (U_{v_0}^\alpha \times |\sigma|)) / \sim$ for each $i > 0$.

The gluing procedure for τ_0 simplifies to $U_{v_0}^\alpha \times |\sigma| \ni (x, y) \sim (x, y) \in U_{v_1}^\alpha \times |\tau_0|$, for all $x \in U_{v_0} \subseteq U_{v_1}$ and $y \in |\tau_0| \subset |\sigma|$, so the gluing occurs exactly along the two sets' intersection $U_{v_0}^\alpha \times |\tau_0| = (U_{v_0}^\alpha \times |\sigma|) \cap (U_{v_1}^\alpha \times |\tau_0|)$, so we have that $((U_{v_0}^\alpha \times |\sigma|) \sqcup (U_{v_1}^\alpha \times |\tau_0|)) / (U_{v_0}^\alpha \times |\tau_0|) = (U_{v_0}^\alpha \times |\sigma|) \cup (U_{v_1}^\alpha \times |\tau_0|)$.

Since these two situations collectively determine the gluing procedures for all simplices of N^α , via induction the result follows. $\square$

From N^α and W^α we define the CW-complex,

$$B^\alpha := \operatorname{hocolim} D^\alpha,$$

which we call the **blowup complex**. It is a gluing-together of all the realizations of the simplices of N^α paired with their corresponding cover elements' intersection in W^α . Note that this is the barycentric decomposition of the so-called (Mayer–Vietoris) blow-up complex central to the proof of the Nerve Theorem, called the realization of a diagram of spaces in Hatcher [35], and other more recent persistent homology research, e.g. Zomorodian and Carlsson's work on localized homology [49].

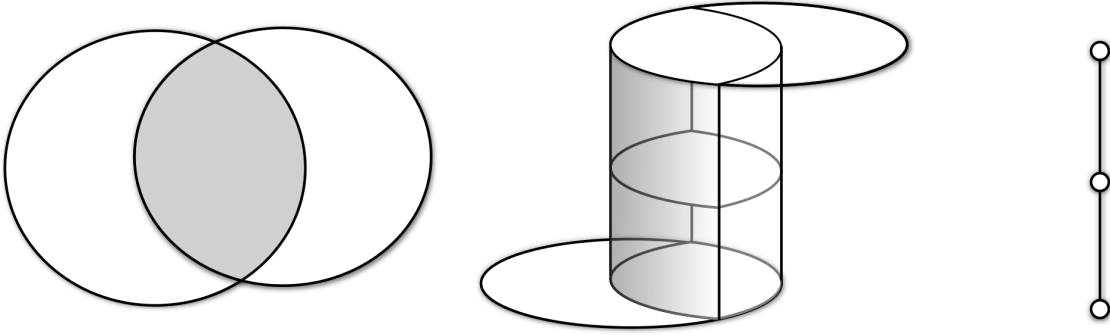


Figure 5.4: The barycentric decomposition of the blow-up complex of a two-element cover alongside the decomposition of the cover's nerve.

5.4.3 The Blowup Filtration

The blowup filtration is denoted by $\mathcal{B} := (B^\alpha)_{\alpha \geq 0}$. By definition we have that $B^\alpha \subseteq W^\alpha \times |N^\alpha|$ for all α . As the filtration $\mathcal{B}$ organizes and combines the nerves and the covered simplicial complexes, it is easier to define maps with it than the covered space filtration $\mathcal{W}$.

Using the filtration $\mathcal{B}$ we reduce the proof to constructing a $(K + 1)\varepsilon$ -interleaving between $H_K(\mathcal{B})$ and $H_K(\operatorname{Nrv} \mathcal{U})$ as follows. There are natural projection maps $b^\alpha := \pi_{W^\alpha} i$ for each $\alpha \geq 0$, where $i : B^\alpha \hookrightarrow W^\alpha \times |N^\alpha|$ and $\pi_{W^\alpha} : W^\alpha \times |N^\alpha| \rightarrow W^\alpha$. It is well-known that b^α is a homotopy

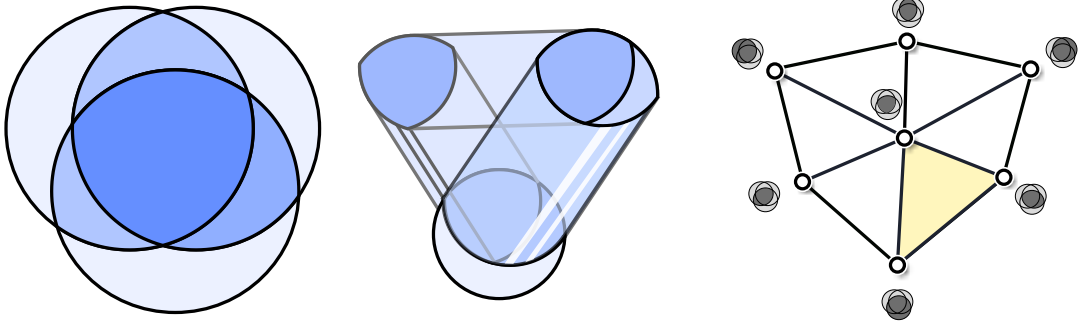


Figure 5.5: A portion of the blowup complex for a three-element cover. Highlighted is a 2-simplex in the realization of the barycentric decomposition of the nerve and its associated cell in the blowup.

equivalence (see 4G.2 in [35]) for covers of paracompact spaces, so $H_*^{\text{CW}}(B^\alpha) \cong H_*^{\text{CW}}(W^\alpha) \cong H_*(W^\alpha)$ as finite simplicial complexes are paracompact.

Moreover, two projections b^α and b^β commute with the inclusions $i_{\mathcal{B}}^{\alpha,\beta} : B^\alpha \hookrightarrow B^\beta$ and $i_{\mathcal{W}}^{\alpha,\beta} : W^\alpha \hookrightarrow W^\beta$, yielding the following commutative homological diagram for all scales α, β such that $0 \leq \alpha \leq \beta$. From Diagram 5.1 it follows that $\text{Dgm}(\mathcal{B}) = \text{Dgm}(\mathcal{W})$ by the Persistence Equivalence Theorem 5.1.

$$\begin{array}{ccc}
 H_*(W^\alpha) & \xrightarrow{i_{\mathcal{W}}^{\alpha,\beta}} & H_*(W^\beta) \\
 b^\alpha \uparrow \cong & & b^\beta \uparrow \cong \\
 H_*^{\text{CW}}(B^\alpha) & \xrightarrow{i_{\mathcal{B}}^{\alpha,\beta}} & H_*^{\text{CW}}(B^\beta)
 \end{array} \tag{5.1}$$

Next define the nerve filtration as $\mathcal{N} := (|N^\alpha|)_{\alpha \geq 0}$, where we note it is a filtration of the geometric realizations of the barycentric decomposition of the combinatorial nerves. There are also natural projection maps $p^\alpha := \pi_{|N^\alpha|} i$ for each $\alpha \geq 0$, where $i : B^\alpha \hookrightarrow W^\alpha \times |N^\alpha|$ and $\pi_{N^\alpha} : W^\alpha \times |N^\alpha| \rightarrow |N^\alpha|$. When $\mathcal{U}^\alpha$ is a good cover of W^α , and by extension $\mathcal{U}$ is a good cover filtration of $\mathcal{W}$, the projection maps p^α are homotopy equivalences that commute with the filtration inclusions. This is a central component to the proof the Nerve Theorem and Lemma 5.4. However, under our

assumptions, we do not have the homotopy equivalences resulting from the good cover condition, so instead we create a $(K+1)\varepsilon$ interleaving between $H_K(\mathcal{N})$ and $H_K^{\text{CW}}(\mathcal{B})$ to prove our theorem. As $|N^\alpha|$ is homeomorphic to $|\text{Nrv } \mathcal{U}^\alpha|$ this is sufficient, which will be expanded upon in the culminative proof of the Generalized Persistent Nerve Theorem.

5.4.4 The Fundamental Chain Maps

In this section we will construct the chain maps mapping a basis chain in $C_K^{\text{CW}}(|N^\alpha|)$ to a chain in $C_K^{\text{CW}}(B^{\alpha+(K+1)\varepsilon})$ which will result in an interleaving by symmetrizing. We will not use the geometric realization vertical bars when discussing basis cellular chains corresponding to tensor products of geometric simplices viewed as cells to avoid cumbersome presentation.

For an abstract k -simplex $\sigma \in N^\alpha$ where $\sigma = v_0 \rightarrow \dots \rightarrow v_k$, the following two shorthands will be used, $\sigma_i := |\sigma_{[v_0, \dots, v_i]}|$ and $\overline{\sigma}_i := |\sigma_{[v_i, \dots, v_k]}|$. These are the geometric realizations of the restriction of σ to the first i -vertices and the i -th through k -th vertices respectively. For notational simplicity, given some σ and some vertex $v \in \sigma$, $\sigma \setminus v := |\sigma_{[v_0, \dots, \hat{v}, \dots, v_k]}|$, where $\hat{v}$ denotes the removal of vertex v .

For each non-empty $v \subseteq [n]$, pick a vertex $x_v \in U_v^{\alpha'} \subset W^{\alpha'}$ where $\alpha' := \min\{\alpha \geq 0 \mid U_v^\alpha \neq \emptyset\}$, and note that $x_v \in U_v^\beta \subset W^\beta$ for all $\beta \geq \alpha'$. Consider the constant map $x_v : U_v^\alpha \rightarrow U_v^{\alpha+\varepsilon}$ which is the linear extension of the vertex map sending each vertex of U_v^α to $x_v \in U_v^{\alpha+\varepsilon}$. This results in the chain map $x_v : C_*^{\text{CW}}(U_v^\alpha) \rightarrow C_*^{\text{CW}}(U_v^{\alpha+\varepsilon})$ defined as follows.

$$x_v(\sigma) := \begin{cases} x_v & \text{if } \dim \sigma = 0, \\ 0 & \text{otherwise.} \end{cases}$$

The following lemma will be used to show the existence of a chain homotopy between the inclusion chain maps and the vertex chain maps.

Lemma 5.9. *Fix $\alpha \geq 0$. Given non-empty $v \subseteq [n]$, where $v \in \text{Nrv } \mathcal{U}^\alpha$, and $x_v \in U_v^\alpha$ is the chosen vertex of v , consider the chain map induced by inclusion, $i_v^{\alpha, \alpha+\varepsilon} : C_*^{\text{CW}}(U_v^\alpha) \rightarrow C_*^{\text{CW}}(U_v^{\alpha+\varepsilon})$ and*

the constant map $x_v : C_*^{\text{CW}}(U_v^\alpha) \rightarrow C_*^{\text{CW}}(U_v^{\alpha+\varepsilon})$. There exists a chain homotopy from $i_v^{\alpha, \alpha+\varepsilon}$ to x_v .

Proof. We construct the chain homotopy by induction on dimension to prove that for all n that there exists c_n, c_{n-1} such that $c_{n-1}\partial_n + \partial_{n+1}c_n = i_v^{\alpha, \alpha+\varepsilon} - x_v$. In the base case we can consider $c_{-1} = c_{-2} = 0$. Now for some $k \geq 0$, assume there exists c_{k-1}, c_{k-2} such that $c_{k-2}\partial_{k-1} + \partial_k c_{k-1} = i_v^{\alpha, \alpha+\varepsilon} - x_v$.

Let z be the k -chain such that $z = -(c_{k-1}\partial_k)(\sigma) + i_v^{\alpha, \alpha+\varepsilon}(\sigma) - x_v(\sigma)$. Observe that for $k \geq 0$, for any k -simplex $\sigma \in C_k^{\text{CW}}(U_v^\alpha)$,

$$\begin{aligned} \partial_k(z) &= -(\partial_k c_{k-1} \partial_k)(\sigma) + (\partial_k i_v^{\alpha, \alpha+\varepsilon})(\sigma) - (\partial_k x_v)(\sigma) \\ &= (c_{k-2} \partial_{k-1} \partial_k)(\sigma) - (i_v^{\alpha, \alpha+\varepsilon} \partial_k)(\sigma) + (x_v \partial_k)(\sigma) + (\partial_k i_v^{\alpha, \alpha+\varepsilon})(\sigma) - (\partial_k x_v)(\sigma) \\ &= 0 \end{aligned}$$

The second line follows by the inductive hypothesis as $\partial_k(\sigma)$ is a $(k-1)$ -chain. The third line follows as $\partial_{k-1}\partial_k = 0$ and $i_v^{\alpha, \alpha+\varepsilon}$ and x_v are chain maps so they commute with the boundary operators.

This proves that z is a cycle so $i_v^{\alpha, \alpha+\varepsilon}(z) = z$ is a cycle, and $\tilde{H}_*(i_v^{\alpha, \alpha+\varepsilon}(z)) = 0$ as $\mathcal{U}$ is an ε -good cover. There then must exist a boundary b such that $\partial_{k+1}(b) = z$. Define $c_k(\sigma) := b$. By the above calculations this choice of c_k satisfies the inductive hypothesis so we are done. $\square$

By Lemma 5.9, for a given $\alpha \geq 0$ and $v \in \text{Nrv } \mathcal{U}^\alpha$, there exists a chain homotopy, denoted c_v^α , between the identity chain map $i_v^{\alpha, \alpha+\varepsilon} : C_*^{\text{CW}}(U_v^\alpha) \rightarrow C_*^{\text{CW}}(U_v^{\alpha+\varepsilon})$ and the constant chain map $x_v^\alpha : C_*^{\text{CW}}(U_v^\alpha) \rightarrow C_*^{\text{CW}}(U_v^{\alpha+\varepsilon})$. By definition we have the equality

$$\partial c_v^\alpha + c_v^\alpha \partial = i_v^{\alpha, \alpha+\varepsilon} + x_v^\alpha.$$

Define $t := (K+1)\varepsilon$ for the remainder of the paper. For $k \leq K$, define the map $c^\alpha : C_k^{\text{CW}}(B^\alpha) \rightarrow C_{k+1}^{\text{CW}}(W^{\alpha+t})$ for a cellular basis element $\tau \otimes \sigma \in C_k^{\text{CW}}(B^\alpha)$, where τ is a p -cell and σ is a q -cell

corresponding to the geometric realization of the abstract q -simplex $\sigma = v_0 \rightarrow \dots \rightarrow v_k$ in N^α , as

$$c^\alpha(\tau \otimes \sigma) := c_\sigma^\alpha(\tau) := (c_{v_q}^{\alpha+q\varepsilon} \dots c_{v_0}^\alpha)(\tau).$$

Note that this is well-defined despite c_v^α 's domain each being $C_*^{\text{CW}}(U_v^\alpha)$ as for any basis cellular chain $\tau \otimes \sigma$, τ is a simplex of U_v^α for some v by the definition of B^α .

Recall the two projections from the barycentric decomposition of the blow-up complex B^α , b^α and p^α . For a cellular chain k -chain $\tau \otimes \sigma \in C_k^{\text{CW}}(B^\alpha)$, the projection-induced chain maps $b^\alpha : C_*^{\text{CW}}(B^\alpha) \rightarrow C_*^{\text{CW}}(W^\alpha)$ and $p^\alpha : C_*^{\text{CW}}(B^\alpha) \rightarrow C_*^{\text{CW}}(|N^\alpha|)$ are defined as follows,

$$b^\alpha(\tau \otimes \sigma) := \begin{cases} \tau & \text{if } \dim \sigma = 0 \\ 0 & \text{otherwise,} \end{cases}$$

and

$$p^\alpha(\tau \otimes \sigma) := \begin{cases} \sigma & \text{if } \dim \tau = 0, \\ 0 & \text{otherwise.} \end{cases}$$

Define the chain map $q^\alpha : C_*^{\text{CW}}(|N^\alpha|) \rightarrow C_*^{\text{CW}}(W^{\alpha+t})$ for a basis k -simplex $\sigma \in C_k^{\text{CW}}(|N^\alpha|)$ as the following

$$q^\alpha(\sigma) := \begin{cases} c^\alpha(x_{v_0} \otimes (\sigma \setminus v_0)) & \text{if } \dim \sigma \geq 1 \\ x_{v_0} & \text{if } \dim \sigma = 0. \end{cases}$$

The chain maps b^α and q^α are induced by topological maps so they are chain maps by construction, while q^α we prove is a chain map.

Lemma 5.10. *For $t := (K + 1)\varepsilon$ and $\alpha \geq 0$, the map $q^\alpha : C_*^{\text{CW}}(|N^\alpha|) \rightarrow C_*^{\text{CW}}(W^{\alpha+t})$ is a chain map in all dimensions less than or equal to K .*

Proof. Denote $q_k^\alpha : C_k^{\text{CW}}(|N^\alpha|) \rightarrow C_k^{\text{CW}}(W^{\alpha+t})$ for this proof to make it clear what dimension is being worked in. We will prove that q^α is a chain map by induction on the basis of $C_k^{\text{CW}}(|N^\alpha|)$ for arbitrary $\alpha \geq 0$ and $k \leq K$. These are the simplices of $|N^\alpha|$ resulting from abstract simplices in N^α of the form $\sigma = v_0 \rightarrow \dots \rightarrow v_k$.

In the base case, where $\dim \sigma = 0$, then $\sigma = |v|$ for some vertex $v \in N^\alpha$, so we have that $\partial q_0^\alpha(v) = \partial(x_v) = 0 = q_{-1}^\alpha \partial(v)$.

Now assume that for some $k \geq 1$, the following holds for any given basis k -chain $\sigma \in C_k^{\text{CW}}(|N^\alpha|)$, $q_{k-1}^\alpha \partial_k(\sigma) = \partial_k q_k^\alpha(\sigma)$. Now consider a basis $(k+1)$ -chain $\sigma' \in C_{k+1}^{\text{CW}}(|N^\alpha|)$. We have the following equalities, defining $\sigma'' := \sigma' \setminus v_{k+1}$.

$$\begin{aligned}
\partial q_{k+1}^\alpha(\sigma') &= \partial c^\alpha(x_{v_0} \otimes (\sigma' \setminus v_0)) && [\text{Def. of } q^\alpha] \\
&= \partial c_{v_{k+1}}^{\alpha+k\varepsilon} c^\alpha(x_{v_0} \otimes (\sigma' \setminus (v_0 \cup v_{k+1}))) && [\text{Def. of } c^\alpha] \\
&= (c_{v_{k+1}}^{\alpha+k\varepsilon} \partial + i_{v_{k+1}}^{\alpha+k\varepsilon, \alpha+(k+1)\varepsilon} + x_{v_{k+1}}^\alpha)(c^\alpha(x_{v_0} \otimes (\sigma' \setminus (v_0 \cup v_{k+1})))) && [\text{Chain homotopy def.}] \\
&= c_{v_{k+1}}^{\alpha+k\varepsilon} \partial c^\alpha(x_{v_0} \otimes (\sigma'' \setminus v_0)) + c^\alpha(x_{v_0} \otimes (\sigma'' \setminus v_0)) + 0 && [\text{Distributive property}] \\
&= c_{v_{k+1}}^{\alpha+k\varepsilon} \partial q_k^\alpha(\sigma'') + q_k^\alpha(\sigma'') && [\text{Def. of } q^\alpha] \\
&= c_{v_{k+1}}^{\alpha+k\varepsilon} q_{k-1}^\alpha \partial(\sigma'') + q_k^\alpha(\sigma'') && [q^\alpha \text{ is a chain map}] \\
&= \sum_{j=0}^k c_{v_{k+1}}^{\alpha+k\varepsilon} q_{k-1}^\alpha(\sigma'' \setminus v_j) + q_k^\alpha(\sigma'') && [\text{Application of } \partial] \\
&= \sum_{j=0}^k q_k^\alpha(\sigma' \setminus v_j) + q_k^\alpha(\sigma' \setminus v_{k+1}) = \sum_{i=0}^{k+1} q_k^\alpha(\sigma' \setminus v_i) = q_k^\alpha \partial(\sigma'). && [\text{Def. of } \sigma'']
\end{aligned}$$

Then, by definition, q^α is a chain map. □

Now we will define the last chain map needed, $a^\alpha : C_*^{\text{CW}}(B^\alpha) \rightarrow C_*^{\text{CW}}(W^{\alpha+t})$ — the composition of q^α and p^α . As p^α and q^α are chain maps, a^α is a chain map as it is the composition of chain maps. For any basis cellular chain $\tau \otimes \sigma \in C_k^{\text{CW}}(B^\alpha)$,

$$a^\alpha(\tau \otimes \sigma) = (q^\alpha p^\alpha)(\tau \otimes \sigma) = \begin{cases} c^\alpha(x_{v_0} \otimes (\sigma \setminus v_0)) & \text{if } \dim \sigma \geq 1 \text{ and } \dim \tau = 0 \\ x_{v_0} & \text{if } \dim \sigma = 0 \text{ and } \dim \tau = 0 \\ 0 & \text{otherwise.} \end{cases}$$

Figure 5.2 is a diagram of the k -dimensional cellular chain groups in question and the maps between

them we have just defined, where the diagonal map is q^α .

$$\begin{array}{ccc}
C_k^{\text{CW}}(W^\alpha) & \xrightarrow{i_{\mathcal{W}}^{\alpha, \alpha+t}} & C_k^{\text{CW}}(W^{\alpha+t}) \\
\uparrow b^\alpha & \nearrow i_{\mathcal{B}}^{\alpha, \alpha+t} & \uparrow b^{\alpha+t} \\
C_k^{\text{CW}}(B^\alpha) & \xrightarrow{q^\alpha} & C_k^{\text{CW}}(B^{\alpha+t}) \\
\downarrow p^\alpha & \nwarrow & \downarrow p^{\alpha+t} \\
C_k^{\text{CW}}(|N^\alpha|) & \xrightarrow{i_{\mathcal{N}}^{\alpha, \alpha+t}} & C_k^{\text{CW}}(|N^{\alpha+t}|)
\end{array} \tag{5.2}$$

In the following lemma, we prove that a^α and $i_{\mathcal{W}}^{\alpha, \alpha+t} b^\alpha$ are chain homotopic with chain homotopy c^α between them, allowing us to extend the local chain homotopies as guaranteed by Lemma 5.9 to be defined on the entire blow-up complex. This is a critical step in the process as it shows we can extract global information from a local assumption, the ε -goodness of the cover.

Lemma 5.11. *c^α is a chain homotopy between the chain maps a^α and $i_{\mathcal{W}}^{\alpha, \alpha+t} b^\alpha : C_k^{\text{CW}}(B^\alpha) \rightarrow C_k^{\text{CW}}(W^{\alpha+t})$ for all $k \leq K$, i.e. $\partial c^\alpha = c^\alpha \partial + i_{\mathcal{W}}^{\alpha, \alpha+t} b^\alpha + a^\alpha$.*

Proof. First we check its true for the initial cases and then prove it for all dimensions by induction.

Consider a basis cellular chain $\tau \otimes \sigma \in C_0^{\text{CW}}(B^\alpha)$ where $\dim \tau = 0$ and $\sigma = v_0$ some vertex $v_0 \in N^\alpha$. We have the following equalities.

$$\begin{aligned}
\partial c^\alpha(\tau \otimes v_0) &= \partial c_{v_0}^\alpha(\tau) = c_{v_0}^\alpha(\partial \tau) + \tau + x_{v_0}^\alpha(v_0) \\
&= 0 + \tau + x_{v_0} \\
&= c^\alpha \partial(\tau \otimes v) + b^\alpha(\tau \otimes v_0) + a^\alpha(\tau \otimes v_0).
\end{aligned}$$

The next case is if $\dim \tau > 0$ and $\sigma = |v_0|$ for some vertex $v_0 \in N^\alpha$.

$$\begin{aligned}
\partial c^\alpha(\tau \otimes v_0) &= \partial c_{v_0}^\alpha(\tau) = c_{v_0}^\alpha(\partial \tau) + \tau + x_{v_0}^\alpha(\tau) \\
&= c^\alpha(\partial \tau \otimes v) + \tau + 0 \\
&= c^\alpha \partial(\tau \otimes v) + b^\alpha(\tau \otimes v_0) + a^\alpha(\tau \otimes v_0)
\end{aligned}$$

Assume that $\dim \tau = 0$ and $\dim \sigma = 1$, so that $\sigma = v_0 \rightarrow v_1$.

$$\begin{aligned}
\partial c^\alpha(\tau \otimes (v_0 \rightarrow v_1)) &= \partial(c_{v_1}^{\alpha+\varepsilon}(c_{v_0}^\alpha(\tau))) = (c_{v_1}^{\alpha+\varepsilon}\partial + i_{v_1}^{\alpha+\varepsilon, \alpha+2\varepsilon} + x_{v_1}^{\alpha+\varepsilon})(c_{v_0}^\alpha(\tau)) \\
&= c_{v_1}^{\alpha+\varepsilon}\partial c_{v_0}^\alpha(\tau) + c_{v_0}^\alpha(\tau) + 0 \\
&= c_{v_1}^\alpha c_{v_0}^\alpha(\partial\tau) + \tau + x_{v_0}(\tau) + c_{v_0}^\alpha(\tau) \\
&= c^\alpha(\partial\tau \otimes \sigma) + c_{v_1}^\alpha(\tau) + c_{v_1}^\alpha(x_{v_0}) + c_{v_0}^\alpha(\tau) \\
&= c^\alpha(\partial\tau \otimes \sigma) + c^\alpha(\tau \otimes v_1) + c^\alpha(\tau \otimes v_0) + c^\alpha(x_{v_0} \otimes v_1) \\
&= c^\alpha(\partial\tau \otimes \sigma) + c^\alpha(x \otimes \partial\sigma) + a^\alpha(\tau \otimes \sigma) \\
&= c^\alpha\partial(\tau \otimes \sigma) + a^\alpha(\tau \otimes \sigma) + b^\alpha(\tau \otimes \sigma).
\end{aligned}$$

Now assume that for any basis chain $\tau \otimes \sigma$ such that $\dim \tau = 0$ and $\dim \sigma \leq k-1$,

$$\partial c^\alpha(\tau \otimes \sigma) = \partial c^\alpha + i_{\mathcal{W}}^{\alpha, \alpha+t} b^\alpha + a^\alpha.$$

For a k -simplex $\sigma \in |N^\alpha|$ we have the following equalities.

$$\begin{aligned}
\partial c^\alpha(\tau \otimes \sigma) &= \partial c_{v_k}^\alpha(c^\alpha(\tau \otimes \sigma_{k-1})) \\
&= c_{v_k}^\alpha\partial(c^\alpha(\tau \otimes \sigma_{k-1})) + c^\alpha(\tau \otimes \sigma_{k-1}) + x_{v_k}^\alpha(\tau) \\
&= c_{v_k}^\alpha(c^\alpha\partial(\tau \otimes \sigma_{k-1}) + a^\alpha(\tau \otimes \sigma_{k-1}) + b^\alpha(\tau \otimes \sigma_{k-1})) + c^\alpha(\tau \otimes \sigma_{k-1}) \\
&= c_{v_k}^\alpha(c^\alpha(\partial\tau \otimes \sigma_{k-1}) + c^\alpha(\tau \otimes \partial(\sigma_{k-1})) + c_{v_k}^\alpha(a^\alpha(\tau \otimes \sigma_{k-1}) + c_{v_k}^\alpha(b^\alpha(\tau \otimes \sigma_{k-1})))) \\
&\quad + c^\alpha(x \otimes \sigma_{k-1}) \\
&= c^\alpha(\partial\tau \otimes \sigma) + c^\alpha(\tau \otimes \partial\sigma) + 2c^\alpha(\tau \otimes \sigma_{k-1}) + c_{v_k}^\alpha(c^\alpha(x_{v_0} \otimes (\sigma_{k-1} \setminus v_0)) + 0) \\
&= c^\alpha\partial(\tau \otimes \sigma) + 0 + c^\alpha(x_{v_0} \otimes (\sigma \setminus v_0)) \\
&= c^\alpha\partial(\tau \otimes \sigma) + b^\alpha(\tau \otimes \sigma) + a^\alpha(\tau \otimes \sigma).
\end{aligned}$$

To complete the proof by induction, consider τ and σ such that $\dim \tau = s \geq 1$ and $\dim \sigma = k \geq 2$, and assume the hypothesis holds true for all $\sigma' \in$ such that $\dim \sigma' = k-1$. In particular this means it is true for σ_{k-1} so the following holds completing the proof.

$$\begin{aligned}
\partial c^\alpha(\tau \otimes \sigma) &= \partial c_{v_k}^\alpha(c^\alpha(\tau \otimes \sigma_{k-1})) \\
&= c_{v_k}^\alpha \partial(c^\alpha(\tau \otimes \sigma_{k-1})) + c^\alpha(\tau \otimes \sigma_{k-1}) + x_{v_k}^\alpha(c^\alpha(\tau \otimes \sigma_{k-1})) \\
&= c_{v_k}^\alpha(c^\alpha \partial(\tau \otimes \sigma_{k-1}) + a^\alpha(\tau \otimes \sigma_{k-1}) + b^\alpha(\tau \otimes \sigma_{k-1})) + c^\alpha(\tau \otimes \sigma_{k-1}) + 0 \\
&= c_{v_k}^\alpha(c^\alpha(\partial \tau \otimes \sigma_{k-1}) + c^\alpha(\tau \otimes \partial(\sigma_{k-1})) + c_{v_k}^\alpha(a^\alpha(\tau \otimes \sigma_{k-1}) + c_{v_k}^\alpha(b^\alpha(\tau \otimes \sigma_{k-1})))) \\
&\quad + c^\alpha(\tau \otimes \sigma_{k-1}) \\
&= c^\alpha(\partial \tau \otimes \sigma) + c^\alpha(\tau \otimes \partial \sigma) + 2c^\alpha(\tau \otimes \sigma_{k-1}) + c_{v_k}^\alpha(a^\alpha(\tau \otimes \sigma_{k-1})) + c_{v_k}^\alpha(b^\alpha(\tau \otimes \sigma_{k-1})) \\
&= c^\alpha \partial(\tau \otimes \sigma) + 0 + 0 + 0 \\
&= c^\alpha \partial(\tau \otimes \sigma) + a^\alpha(\tau \otimes \sigma) + b^\alpha(\tau \otimes \sigma).
\end{aligned}$$

□

5.4.5 Chain Map Lifting

We now introduce a general construction on chain maps which we will apply to the maps defined in the previous subsection to construct the interleaving. Consider simplicial complexes X , Y , and Z . The **Alexander-Whitney diagonal approximation chain map** (see [21]),

$$\Delta_* : C_*^{\text{CW}}(X) \rightarrow C_*^{\text{CW}}(X) \otimes C_*^{\text{CW}}(X),$$

is defined on a k -simplex σ by

$$\Delta_k(\sigma) := \sum_{i=0}^k \sigma_i \otimes \bar{\sigma}_i.$$

We remind the reader that for an n -simplex σ , $\sigma_i := \sigma|_{[v_0, \dots, v_i]}$ and $\bar{\sigma}_i := \sigma|_{[v_i, \dots, v_n]}$, the restriction of the simplex to the first $i+1$ and the last $n-i+1$ vertices respectively.

The following abstract definition will allow us to extend a chain map from $C_*(|N^{\alpha+t}|)$ into $C_*(W^{\alpha+t})$ to a chain map into $C_*(B^{\alpha+t})$.

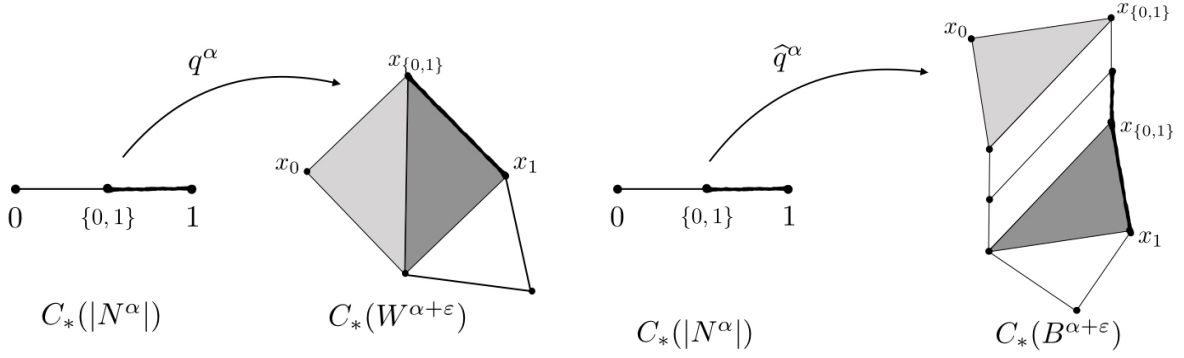


Figure 5.6: (R) q^α mapping the realization of a 1-simplex of $|N^\alpha|$. (L) The 1-simplex being mapped by $\hat{q}^\alpha$ to a 1-chain of $C_*(B^{\alpha+\epsilon})$. Viewing $B^{\alpha+\epsilon}$ vertically, $\hat{q}^\alpha$ projects onto 1-simplex q^α .

Definition 5.12. Given a chain map $f : C_*^{\text{CW}}(X) \otimes C_*^{\text{CW}}(Y) \rightarrow C_*^{\text{CW}}(Z)$, the **lift** of f is the chain map

$$\hat{f} : C_*^{\text{CW}}(X) \otimes C_*^{\text{CW}}(Y) \rightarrow C_*^{\text{CW}}(Z) \otimes C_*^{\text{CW}}(Y)$$

defined by the composition $\hat{f} := (f \otimes 1_Y) \circ (1_X \otimes \Delta_*)$.

Note that the lift of a chain map is a chain map as it is the composition of chain maps. In alignment with the definition of the lift, we will interpret $C_*^{\text{CW}}(B^\alpha)$ as a subgroup of $C_*^{\text{CW}}(W^\alpha) \otimes C_*^{\text{CW}}(|N^\alpha|)$ via the chain complex isomorphism

$$C_*^{\text{CW}}(W^\alpha) \otimes C_*^{\text{CW}}(|N^\alpha|) \cong C_*^{\text{CW}}(W^\alpha \times |N^\alpha|).$$

We must still check that the maps we are lifting are well-defined, i.e. their images are properly contained in $C_*^{\text{CW}}(B^{\alpha+t})$.

The lift of q^α is

$$\hat{q}^\alpha(\sigma) := \sum_{i=0}^k q^\alpha(\sigma_i) \otimes \bar{\sigma}_i.$$

Although the definition is sound with respect to how we defined the lift, it remains to be proven that it is well-defined, i.e. the image of $\hat{q}^\alpha$ lies within $C_k(B^{\alpha+t})$ for all $k \leq K$.

Lemma 5.13. *For all $\alpha \geq 0$, the image of the map $\widehat{q}^\alpha : C_*^{\text{CW}}(|N^\alpha|) \rightarrow C_*^{\text{CW}}(W^{\alpha+t}) \otimes C_*^{\text{CW}}(|N^{\alpha+t}|)$ is contained in $C_*^{\text{CW}}(B^{\alpha+t})$.*

Proof. To see that $\widehat{q}^\alpha$ is well-defined, for a basis 0-chain $\sigma = |v_0| \in C_0(|N^\alpha|)$, note that

$$\widehat{q}^\alpha(\sigma) = x_{v_0} \otimes v_0 \in C_0(U_{v_0}^{\alpha+\varepsilon} \times |v_0|),$$

so $\widehat{q}^\alpha(\sigma) \in C_0(B^{\alpha+t})$. For $k \leq K$ and a basis k -chain $\sigma = |v_0 \rightarrow \dots \rightarrow v_k|$ we have the equality

$$\widehat{q}^\alpha(\sigma) = \sum_{i=0}^k c^\alpha(x_{v_0} \otimes (\sigma_i \setminus v_0)) \otimes \bar{\sigma}_i.$$

For each i , we have that $c^\alpha(x_{v_0} \otimes (\sigma_i \setminus v_0)) \subseteq U_{v_i}^{\alpha+i\varepsilon}$ so

$$c^\alpha(x_{v_0} \otimes (\sigma_i \setminus v_0)) \times |\bar{\sigma}_i| \subset U_{v_i}^{\alpha+i\varepsilon} \times |\bar{\sigma}_i| \subset B^{\alpha+t}.$$

We then have that for all i , each chain of the form $c^\alpha(x_{v_0} \otimes (\sigma_i \setminus v_0)) \otimes \bar{\sigma}_i$ is in $C_k^{\text{CW}}(B^{\alpha+t})$, so their sum is as well, thus $\widehat{q}^\alpha$ is well-defined. $\square$

The maps q^α and $\widehat{q}^\alpha$ define a way to map the nerve filtration into the simplicial union filtration and the subdivided blow-up complex filtration respectively, albeit at a further scale.

The rest of the lifted maps are well-defined as one can easily check that $\widehat{a}^\alpha = \widehat{q}^\alpha p^\alpha$, $p^{\alpha+t} \widehat{q}^\alpha = i_{\mathcal{N}}^{\alpha, \alpha+t}$, and $\widehat{i_{\mathcal{W}}^{\alpha, \alpha+t} b^\alpha} = i_{\mathcal{B}}^{\alpha, \alpha+t}$. For example, the latter equality holds for a basis cellular chain $\tau \otimes \sigma \in C_*^{\text{CW}}(B^\alpha)$,

$$\widehat{i_{\mathcal{W}}^{\alpha, \alpha+t} b^\alpha}(\tau \otimes \sigma) = \sum_{i=0}^k b(\tau \otimes \sigma_i) \otimes \bar{\sigma}_i = \tau \otimes \sigma = i_{\mathcal{B}}^{\alpha, \alpha+t}(\tau \otimes \sigma).$$

These lifted maps constitute Diagram 5.3, adapted from Diagram 5.2 after the lifting procedure.

$$\begin{array}{ccc} C_k^{\text{CW}}(W^\alpha) & \xrightarrow{i_{\mathcal{W}}^{\alpha, \alpha+t}} & C_k^{\text{CW}}(W^{\alpha+t}) \\ \uparrow b^\alpha & & \uparrow b^{\alpha+t} \\ C_k^{\text{CW}}(B^\alpha) & \xrightarrow{\widehat{i_{\mathcal{W}}^{\alpha, \alpha+t} b^\alpha}} & C_k^{\text{CW}}(B^{\alpha+t}) \\ \downarrow p^\alpha & \nearrow \widehat{q}^\alpha & \downarrow p^{\alpha+t} \\ C_k^{\text{CW}}(|N^\alpha|) & \xrightarrow{i_{\mathcal{N}}^{\alpha, \alpha+t}} & C_k^{\text{CW}}(|N^{\alpha+t}|) \end{array} \quad (5.3)$$

The following propositions, stated in terms of general chain complexes, cumulatively imply that chain homotopies are preserved by liftings as stated in Lemma 5.16.

Proposition 5.14. *Consider chain complexes A_* and B_* over $\mathbb{F}_2$ and homotopic chain maps $f, f' : A_* \rightarrow B_*$. The tensored chain maps $f \otimes 1_C, f' \otimes 1_C : A_* \otimes C_* \rightarrow B_* \otimes C_*$, where 1_C is the identity chain map from C_* to itself, are chain homotopic.*

Proof. By hypothesis, there exists morphisms $c : A_* \rightarrow B_{*+1}$, such that $\partial_B c + c \partial_A = f + f'$. We then have the following equalities, where $D := \partial_B \otimes 1_C + 1_B \otimes \partial_C$ is the chain complex differential of $B_* \otimes C_*$, using the fact that the tensor product of two chain maps $f \otimes g$ can be decomposed as the composition

$$\begin{aligned}
f \otimes 1_C + f' \otimes 1_C &= (f + f') \otimes 1_C = (\partial_B c + c \partial_A) \otimes 1_C \\
&= \partial_B c \otimes 1_C + c \partial_A \otimes 1_C \\
&= \partial_B c \otimes 1_C + c \otimes \partial_C + c \otimes \partial_C + c \partial_A \otimes 1_C \\
&= (\partial_B \otimes 1_C + 1_B \otimes \partial_C)(c \otimes 1_C) + (c \otimes 1_C)(\partial_B \otimes 1_C + 1_B \otimes \partial_C) \\
&= D(c \otimes 1_C) + (c \otimes 1_C)D.
\end{aligned}$$

The sequence of equalities proves that $f \otimes 1_C$ is chain homotopic to $f' \otimes 1_C$ via chain homotopy $c \otimes 1_C$. □

Proposition 5.15. *Given chain complexes A_*, B_* and C_* over $\mathbb{F}_2$ and two pairs of chain maps $f, f' : A_* \rightarrow B_*$ and $g, g' : B_* \rightarrow C_*$ where $f \simeq f'$ via chain homotopy c , and $g \simeq g'$ via chain homotopy d , $gf \simeq g'f'$ via chain homotopy $gc + df'$.*

Proof. By hypothesis, $f - f' = c \partial_A + \partial_B c$ and $g - g' = d \partial_B + \partial_C d$. We then have the following

equalities, using the defining property of chain maps and their linearity.

$$\begin{aligned}
gf - g'f' &= gf - gf' + gf' - g'f' = g(f - f') + (g - g')f' \\
&= g(c\partial_A + \partial_B c) + (d\partial_B + \partial_C d)f' \\
&= g c\partial_A + g\partial_B c + d\partial_B f' + \partial_C d f' \\
&= g c\partial_A + \partial_C g c + d f' \partial_A + \partial_C d f' \\
&= (gc + df')\partial_A + \partial_C(gc + df').
\end{aligned}$$

Thus by definition $gc + df'$ is a chain homotopy between gf and $g'f'$. $\square$

Lemma 5.16. *If chain maps $f, g : C_*^{\text{CW}}(X) \otimes C_*^{\text{CW}}(Y) \rightarrow C_*^{\text{CW}}(Z)$ are chain homotopic, then $\widehat{f}$ and $\widehat{g}$ are chain homotopic, where $\widehat{f}$ and $\widehat{g}$ are the lifts of f and g respectively.*

Proof. Denote the chain homotopy between f and g by c . By Proposition 5.14, $c \otimes \text{id}_Y$ is a chain homotopy between $f \otimes 1_Y$ and $g \otimes 1_Y$. Now consider $\widehat{f} = (f \otimes 1_Y) \circ (1_X \otimes \Delta_*)$ and $\widehat{g} = (g \otimes 1_Y) \circ (1_X \otimes \Delta_*)$. These two maps are chain homotopic by Proposition 5.15, with chain homotopy $\widehat{c} := (c \otimes 1_Y) \circ (1_X \otimes \Delta_*)$. $\square$

We are now ready to prove the existence of a $(K + 1)\varepsilon$ -interleaving between the blowup persistence module and the nerve persistence module, which leads to the Generalized Persistent Nerve Theorem.

Lemma 5.17. *There is a $(K + 1)\varepsilon$ -interleaving for all dimensions $K \geq 0$ between the blowup persistence module $H_*(\mathcal{B})$ and the nerve persistence module $H_*(\mathcal{N})$.*

Proof. Fix a dimension $K \geq 0$. Define $t := (K + 1)\varepsilon$. Consider the two collections of chain maps in dimension K :

$$\mathbf{p} := (p^{\alpha+t} i_{\mathcal{B}}^{\alpha, \alpha+t})_{\alpha \geq 0}$$

and

$$\widehat{\mathbf{q}} := (\widehat{q}^\alpha)_{\alpha \geq 0}.$$

Lemma 5.11 states that c^α is a chain homotopy between a^α , where $a^\alpha := q^\alpha p^\alpha$, and $i_{\mathcal{W}}^{\alpha, \alpha+t} b^\alpha$. Lemma 5.16 applied to c^α , a^α and $i_{\mathcal{W}}^{\alpha, \alpha+t} b^\alpha$ implies that for all $\alpha \geq 0$, $\widehat{q}^\alpha p^\alpha \simeq i_{\mathcal{B}}^{\alpha, \alpha+t}$, recalling that $\widehat{a}^\alpha = \widehat{q}^\alpha p^\alpha$ and $\widehat{i_{\mathcal{W}}^{\alpha, \alpha+t} b^\alpha} = i_{\mathcal{B}}^{\alpha, \alpha+t}$. Thus for all $\alpha \geq 0$ we have that

$$\widehat{q}^{\alpha+t} p^{\alpha+t} i_{\mathcal{B}}^{\alpha, \alpha+t} \simeq i_{\mathcal{B}}^{\alpha+t, \alpha+2t} i_{\mathcal{B}}^{\alpha, \alpha+t} = i_{\mathcal{B}}^{\alpha, \alpha+2t}.$$

We also have the equality $p^{\alpha+2t} i_{\mathcal{B}}^{\alpha+t, \alpha+2t} \widehat{q}^\alpha = i_{\mathcal{N}}^{\alpha, \alpha+2t}$, as $p^{\alpha+2t} i_{\mathcal{B}}^{\alpha+t, \alpha+2t} = i_{\mathcal{N}}^{\alpha+t, \alpha+2t} p^{\alpha+t}$ due to the facts that projection and inclusion chain maps commute and that $p^{\alpha+t} \widehat{q}^\alpha = i_{\mathcal{N}}^{\alpha, \alpha+t}$.

Consider the collections of homology maps $H_K^{\text{CW}}(\mathbf{p})$ and $H_K^{\text{CW}}(\widehat{\mathbf{q}})$ where right composition and left composition are depicted pictorially in Diagrams 5.4 and 5.5 respectively below at some arbitrary scale $\alpha \geq 0$.

$$\begin{array}{ccccc} H_K^{\text{CW}}(B^\alpha) & \xrightarrow{i_{\mathcal{B}}^{\alpha, \alpha+t}} & H_K^{\text{CW}}(B^{\alpha+t}) & \longrightarrow & H_K^{\text{CW}}(B^{\alpha+2t}) \\ \downarrow & \nearrow & \downarrow p^{\alpha+t} & \nearrow \widehat{q}^{\alpha+t} & \downarrow \\ H_K^{\text{CW}}(|N^\alpha|) & \longrightarrow & H_K^{\text{CW}}(|N^{\alpha+t}|) & \longrightarrow & H_K^{\text{CW}}(|N^{\alpha+2t}|) \end{array} \quad (5.4)$$

$$\begin{array}{ccccc} H_K^{\text{CW}}(B^\alpha) & \longrightarrow & H_K^{\text{CW}}(B^{\alpha+t}) & \xrightarrow{i_{\mathcal{B}}^{\alpha+t, \alpha+2t}} & H_K^{\text{CW}}(B^{\alpha+2t}) \\ \downarrow & \nearrow \widehat{q}^\alpha & \downarrow & \nearrow & \downarrow p^{\alpha+2t} \\ H_K^{\text{CW}}(|N^\alpha|) & \longrightarrow & H_K^{\text{CW}}(|N^{\alpha+t}|) & \longrightarrow & H_K^{\text{CW}}(|N^{\alpha+2t}|) \end{array} \quad (5.5)$$

Now we apply the K -dimensional cellular homology functor to the chain maps p and $\widehat{q}$ to obtain the follow equalities:

$$H_K^{\text{CW}}(\widehat{q}^{\alpha+t}) H_K^{\text{CW}}(p^{\alpha+t} i_{\mathcal{B}}^{\alpha, \alpha+t}) = H_K^{\text{CW}}(i_{\mathcal{B}}^{\alpha, \alpha+2t})$$

$$H_K^{\text{CW}}(p^{\alpha+2t} i_{\mathcal{B}}^{\alpha+t, \alpha+2t}) H_K^{\text{CW}}(\widehat{q}^\alpha) = H_K^{\text{CW}}(i_{\mathcal{N}}^{\alpha, \alpha+2t}).$$

The collections $H_K^{\text{CW}}(\mathbf{p})$ and $H_K^{\text{CW}}(\widehat{\mathbf{q}})$ thus form a $(K+1)\varepsilon$ -interleaving between $H_K^{\text{CW}}(\mathcal{B})$ and $H_K^{\text{CW}}(\mathcal{N})$ by Definition 5.2. $\square$

5.5 The Main Theorem

We now conclude with our main result, the Generalized Persistent Nerve Theorem.

Theorem 5.18 (Generalized Persistent Nerve Theorem). *Given a finite collection of finite simplicial filtrations $\mathcal{U} = \{U_0, \dots, U_n\}$, where $U_i := (U_i^\alpha)_{\alpha \geq 0}$ and all U_i^α are defined over the same vertex set, if $\mathcal{U}$ is an ε -good cover filtration of $\mathcal{W} := (\bigcup_{i=0}^n U_i^\alpha)_{\alpha \geq 0}$, then for all $K \geq 0$,*

$$d_B(Dgm_K(\mathcal{W}), Dgm_K(Nrv \mathcal{U})) \leq (K + 1)\varepsilon.$$

Proof. Fix a dimension $K \geq 0$. By result 4G.2 in [35], for each $\alpha \geq 0$, the map $b^\alpha : B^\alpha \rightarrow W^\alpha$ is a homotopy equivalence so $H_*^{CW}(B^\alpha)$ and $H_*^{CW}(W^\alpha)$ are naturally isomorphic at all scales. Theorem 5.1 implies that $Dgm_K(\mathcal{B}) = Dgm_K(\mathcal{W})$ for all dimensions $K \geq 0$, so by definition $d_B(Dgm_K(\mathcal{B}), Dgm_K(\mathcal{W})) = 0$.

We remind the reader of the fact that for three persistence diagrams D, D', D'' , $d_B(D, D'') \leq d_B(D, D') + d_B(D', D'')$, as the bottleneck distance is a metric on the space of finite persistence diagrams and thus obeys the triangle inequality. What then remains to be proven is that $d_B(Dgm_K(\mathcal{B}), Dgm_K(Nrv \mathcal{U})) \leq (K + 1)\varepsilon$.

By Lemma 5.17, there is a $(K + 1)\varepsilon$ -interleaving between $H_K^{CW}(\mathcal{B})$ and $H_K^{CW}(\mathcal{N})$ for all $k \leq K$. Then, by the Algebraic Stability Theorem, $d_B(Dgm_K(\mathcal{B}), Dgm_K(\mathcal{N})) \leq (K + 1)\varepsilon$.

To complete the proof, N^α is the barycentric decomposition of $Nrv \mathcal{U}^\alpha$, so their geometric realizations are homeomorphic and thus by Theorem 5.1, $Dgm(\mathcal{N}) = Dgm(Nrv \mathcal{U})$. Also we have that $H_K^{CW}(\mathcal{N}) \cong H_K(\mathcal{N})$ as cellular and simplicial homology agree for simplicial complexes. By considering the isomorphisms maps corresponding to the previous equality and the equality $Dgm(\mathcal{B}) = Dgm(\mathcal{W})$, we can conclude there is $(K + 1)\varepsilon$ -interleaving between $H_K(\mathcal{W})$ and $H_K(Nrv \mathcal{U})$, so $d_B(Dgm_K(\mathcal{W}), Dgm_K(Nrv \mathcal{U})) \leq (K + 1)\varepsilon$ by Theorem 5.3. $\square$

From this theorem we can further conclude that $d_B(Dgm(\mathcal{W}), Dgm(Nrv \mathcal{U})) \leq (D + 1)\varepsilon$, where

D is the maximum dimension of $\text{Nrv } \mathcal{U}$. TPersistent Nerve Lemma is also a corollary in the case of finite simplicial filtrations.

Corollary 5.19 (The Simplicial Persistent Nerve Lemma). *Given a finite collection of finite simplicial filtrations $\mathcal{U} = \{U_0, \dots, U_n\}$ where $U_i := (U_i^\alpha)_{\alpha \geq 0}$ and all U_i^α are all defined over the same vertex set, if $\mathcal{U}$ is a good cover filtration of the $\mathcal{W} = (\bigcup_{i=0}^n U_i^\alpha)_{\alpha \geq 0}$, then*

$$Dgm(\mathcal{W}) = Dgm(\text{Nrv } \mathcal{U}).$$

Proof. Since $\mathcal{U}$ is a good cover, it is also a 0-good cover. By the Generalized Persistent Nerve Theorem we know that $d_B(Dgm(\mathcal{W}), Dgm(\text{Nrv } \mathcal{U})) = 0$. As both persistence diagrams are finite, $Dgm(\text{Nrv } \mathcal{U}) = Dgm(\mathcal{W})$. $\square$

5.5.1 The Bound is Tight

In this section a cover and a filtration is presented that proves the tightness of the bottleneck distance bound of Theorem 5.18 over all dimensions. In particular, we present a 1-good cover $\mathcal{U}$ of a simplicial filtration $\mathcal{W}$ such that $d_B(Dgm_k(\text{Nrv } \mathcal{U}), Dgm_k(\mathcal{W})) = k + 1$, for all $k \leq K$. This can be transformed into an ε -good cover filtration for arbitrary $\varepsilon > 0$ by appropriately scaling the filtration indices and the definition of each U_i^α .

Consider the collection of $n+1$ $(n-1)$ -simplices where each simplex is defined as $t_i := |[n] \setminus \{i\}|$, where $[n] = \{0, 1, \dots, n\}$. These are the facets of the standard n -simplex. Define the cover filtration as the collection of simplicial filtrations $\mathcal{U} = \{U_0, \dots, U_n\}$ where

$$U_i^\alpha := \begin{cases} t_i & \text{for } \alpha \in [0, n+1) \\ t_i \cup \left(\bigcup_{k=0}^{\lfloor \alpha \rfloor - n - 1} t_k \right) & \text{for } \alpha \in [n+1, 2n+2) \\ |[n]| & \text{for } \alpha = 2n+2. \end{cases}$$

First we compute $H_*(\mathcal{W})$. For $0 \leq \alpha < 2n+2$, $W^\alpha = \bigcup_{i=0}^n t_i = \partial|[n]|$, so $\text{rk } \tilde{H}_r(W^\alpha) = 1$ if $r = n$, and $\text{rk } \tilde{H}_r(W^\alpha) = 0$ otherwise. For $\alpha = 2n+2$, $W^\alpha = |[n]|$, which is contractible, so

$\text{rk } \tilde{H}_r(W^\alpha) = 0$ for all $r \geq 0$. The ranks of these homology groups imply that there is one persistent n -dimensional homological feature from $\alpha = 0$ to $\alpha = 2n + 2$, and none in any other dimension.

Next we compute $H_*(\text{Nrv } \mathcal{U})$. Note that any k -wise intersection of the faces of the standard n -simplex is non-empty iff $k \leq n$, so for $0 \leq \alpha < n + 1$, $\text{Nrv } U^\alpha$ is the abstract simplicial complex $2^{[n]} \setminus [n]$. Recall the notation $U_\sigma^\alpha := \bigcap_{i \in \sigma} U_i^\alpha$. For $n + 1 \leq \alpha < 2n + 2$ and non-empty $\sigma \subseteq [n]$, we have that $U_\sigma^\alpha = (\bigcap_{i \in \sigma} t_i) \cup (\bigcup_{k=0}^{|\alpha| - n - 1} t_k)$ which is always non-empty, so $\text{Nrv } U^\alpha$ is the standard n -simplex. The same is true for $\text{Nrv } U^{2n+2}$ as each cover element is identical. The homology groups then, for all $\alpha \in [0, n + 1)$, are $\tilde{H}_r(\text{Nrv } U^\alpha) = 1$ if $r = n$, and 0 otherwise. For all $\alpha \geq n + 1$, for all $r \geq 0$, $\tilde{H}_r(\text{Nrv } U^\alpha) = 0$, so there is one persistent n -dimensional homological feature from $\alpha = 0$ to $\alpha = n + 1$.

By the above computations $\text{Dgm}_n(\mathcal{W})$ consists of the diagonal and the point $(0, 2n + 2) \in \mathbb{R}^2$, while $\text{Dgm}_n(\text{Nrv } \mathcal{U})$ consists of the diagonal and the point $(0, n + 1) \in \mathbb{R}^2$. Regardless of whether $(0, 2n + 2)$ is mapped to $(0, n + 1)$ or the diagonal, the ℓ_∞ distance between the matched points is $n + 1$, so $d_B(\text{Dgm}_n(\mathcal{W}), \text{Dgm}_n(\text{Nrv } \mathcal{U})) = n + 1$.

Next we prove that $\mathcal{U}$ is a 1-good cover — in other words, for all σ and all α , we need to prove that $\tilde{H}_*(U_\sigma^\alpha \hookrightarrow U_\sigma^{\alpha+1}) = 0$. Note that $\mathcal{U}^\alpha$ is a good cover of W^α for all $\alpha \in [0, n + 1)$, as the intersection between k of the faces of an n -simplex is a full $(n - k)$ -simplex, which is contractible. In order to prove that the rest of the cover filtration is 1-good, we show that there are no scales $\alpha \geq n + 1$ for which $\tilde{H}_*(U_\sigma^\alpha) \neq 0$ and $\tilde{H}_*(U_\sigma^{\alpha+1}) \neq 0$. To do so, we will check the $\sigma \subseteq [n]$ and scales result in non-trivial homology. By the definition of the cover filtration's elements, it suffices to check only U_σ^{n+1+j} for $j \in [n]$.

Lemma 5.20. *For $\sigma \subseteq [n]$ and $m > j$ such that $m \notin \sigma$, U_σ^{n+1+j} is contractible to $|m|$ and $\tilde{H}_*(U_\sigma^{n+1+j}) = 0$.*

Proof. Recall that $U_i^{n+1+j} = t_i \cup (\bigcup_{k=0}^j t_k)$ and $U_\sigma^{n+1+j} = \bigcap_{i \in \sigma} t_i \cup (\bigcup_{k=0}^j t_k)$. Given that $m \notin \sigma$, then $|m| \in t_i$ for all $i \in \sigma$, so $|m| \in \bigcap_{i \in \sigma} t_i$. In addition, $|m| \in t_k$ for all $k \neq m$, so $|m| \in t_k$ for all

$k \leq j$ as $m > j$, so $|m|$ is in every simplex in U_σ^{n+1+j} , and thus U_σ^{n+1+j} contracts to $|m|$. $\square$

Lemma 5.21. *For $p \leq j < n$ and $\sigma = ([n] \setminus [j]) \cup \{p\}$, $\tilde{H}_*(U_\sigma^{n+1+j}) = 0$.*

Proof. By assumption we know that $U_\sigma^{n+1+j} = (\bigcap_{i \in \sigma} t_i) \cup (\bigcup_{i=0}^j t_j)$, where $\bigcap_{i \in \sigma} t_i = |[0, \dots, \hat{p}, \dots, j]|$. As $|[0, \dots, \hat{p}, \dots, j]| \subseteq t_p$, $|[0, \dots, \hat{p}, \dots, j]| \subset \bigcup_{i=0}^j t_j$, so we know $U_\sigma^{n+1+j} = \bigcup_{i=0}^j t_j$. By the assumption that $j < n$, U_σ^{n+1+j} is not the entire boundary of $|[n]|$, so $\tilde{H}_*(U_\sigma^{n+1+j}) = 0$. $\square$

Lemma 5.22. *Given non-empty $\sigma \supset [n] \setminus [j]$, for some $j \in [n]$, where $\#\{\sigma \cap [j]\} \geq 2$, $\tilde{H}_*(U_\sigma^{n+1+j}) = 0$.*

Proof. First we claim that for any $p \in \sigma \cap [j]$, $U_\sigma^{n+1+j} = U_{\sigma \setminus p}^{n+1+j}$. By definition, $U_\sigma^{n+1+j} \subseteq U_{\sigma \setminus p}^{n+1+j}$. Now given some $x \in U_{\sigma \setminus p}^{n+1+j}$, $x \in \bigcap_{i \in \sigma \setminus p} t_i$ or $x \in \bigcup_{k=0}^j t_j$. Clearly in the latter case, $x \in U_\sigma^j$. In the former case, we have that $\bigcap_{i \in \sigma} t_i = |[0, \dots, \hat{p}_1, \dots, \hat{p}_l, \dots, j]|$ for some vertices $\hat{p}_1, \dots, \hat{p}_l \in (\sigma \cap [j]) \setminus \{p\}$, so $|[0, \dots, \hat{p}_1, \dots, \hat{p}_l, \dots, j]| \subseteq t_p \subseteq \bigcup_{k=0}^j t_k$, and thus $x \in U_\sigma^j$, giving us the desired equality.

By exhaustion, the only case left to check is when $\#\{\sigma \cap [j]\} = 1$. Applying Lemma 5.21 in this case proves that $\tilde{H}_*(U_\sigma^{n+1+j}) = 0$. $\square$

Lemma 5.20, 5.21 and 5.22 collectively check all non-trivial cases so U_σ^{n+1+j} has non-trivial homology if only if $\sigma = [n] \setminus [j]$ where $j \in [n]$. This fact along with the second half of the cover filtration being good proves that $\mathcal{U}$ is a 1-good cover of $\mathcal{W}$. To complete the construction simply take the disjoint union of the cover filtrations $\mathcal{U}$ over all required dimensions along with the corresponding covered space filtration.

Funding Acknowledgments

This research was supported in part by the National Science Foundation grants Grants CCF-1464379 and CFF-1525978.

Bibliography

- [1] Henry Adams and Gunnar Carlsson. Evasion paths in mobile sensor networks. *International Journal of Robotics Research* 34 (2014), pages 90–104, 2013.
- [2] Pankaj K. Agarwal, Kyle Fox, and Oren Salzman. An efficient algorithm for computing high quality paths amid polygonal obstacles. In *Proceedings of the 27th Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 1179–1192, 2016.
- [3] N. Amenta, S. Choi, T. K. Dey, and N. Leekha. A simple algorithm for homeomorphic surface reconstruction. In *Proceedings of the Sixteenth Annual Symposium on Computational Geometry*, SCG '00, pages 213–222, New York, NY, USA, 2000. ACM.
- [4] Nina Amenta and Marshall Bern. Surface reconstruction by voronoi filtering. In *Proceedings of the Fourteenth Annual Symposium on Computational Geometry*, SCG '98, pages 39–48, New York, NY, USA, 1998. ACM.
- [5] Nina Amenta, Marshall Bern, and Manolis Kamvysselis. A new Voronoi-based surface reconstruction algorithm. In *SIGGRAPH*, pages 415–421, 1998.
- [6] Karol Borsuk. On the imbedding of systems of compacta in simplicial complexes. *Fund. Math.*, 35:217–234, 1948.

- [7] Magnus Bakke Botnan and Gard Spreemann. Approximating persistent homology in Euclidean space through collapses. *Applicable Algebra in Engineering, Communication and Computing*, pages 1–29, 2015.
- [8] Gunnar Carlsson, Afra Zomorodian, Anne Collins, and Leonidas Guibas. Persistence Barcodes for Shapes. In Roberto Scopigno and Denis Zorin, editors, *Symposium on Geometry Processing*. The Eurographics Association, 2004.
- [9] Nicholas J. Cavanna, Kirk P. Gardner, and Donald R. Sheehy. When and why does the topological coverage criterion work. In *Proceedings of the ACM-SIAM Symposium on Discrete Algorithms*, 2017.
- [10] Nicholas J. Cavanna and Donald R. Sheehy. Adaptive metrics for adaptive samples. In *Proceedings of the 28th Canadian Conference on Computational Geometry, CCCG 2016, August 3-5, 2016, Simon Fraser University, Vancouver, British Columbia, Canada*, pages 128–132, 2016.
- [11] F. Chazal and A. Lieutier. Stability and homotopy of a subset of the medial axis. In *Proceedings of the Ninth ACM Symposium on Solid Modeling and Applications, SM '04*, pages 243–248, Aire-la-Ville, Switzerland, Switzerland, 2004. Eurographics Association.
- [12] Frédéric Chazal, David Cohen-Steiner, Marc Glisse, Leonidas J. Guibas, and Steve Y. Oudot. Proximity of persistence modules and their diagrams. In *Proceedings of the Twenty-fifth Annual Symposium on Computational Geometry, SCG '09*, pages 237–246, New York, NY, USA, 2009. ACM.
- [13] Frédéric Chazal, David Cohen-Steiner, and André Lieutier. A sampling theory for compact sets in Euclidean space. *Discrete & Computational Geometry*, 41:461–479, 2009.

- [14] Frédéric Chazal, Vin de Silva, Marc Glisse, and Steve Y. Oudot. *The Structure and Stability of Persistence Modules*. Springer Briefs in Mathematics. Springer, 2016.
- [15] Frédéric Chazal and André Lieutier. Weak feature size and persistent homology: Computing homology of solids in R^n from noisy data samples. In *Proceedings of the 21st ACM Symposium on Computational Geometry*, pages 255–262, 2005.
- [16] Frédéric Chazal and André Lieutier. Smooth manifold reconstruction from noisy and non-uniform approximation with guarantees. *Computational Geometry: Theory and Applications*, 40:156–170, 2008.
- [17] Frédéric Chazal and Steve Yann Oudot. Towards persistence-based reconstruction in euclidean spaces. In *Proceedings of the Twenty-fourth Annual Symposium on Computational Geometry*, SCG '08, pages 232–241, New York, NY, USA, 2008. ACM.
- [18] Jennifer Gamble Harish Chintakunta and Hamid Krim. Applied topology in static and dynamic sensor networks. *Signal Processing and Communications (SPCOM), 2012 International Conference*, pages 1–5, 2011.
- [19] Kenneth L. Clarkson. Building triangulations using ε -nets. In *Proceedings of the thirty-eighth annual ACM symposium on Theory of computing*, pages 326–335, 2006.
- [20] David Cohen-Steiner, Herbert Edelsbrunner, and John Harer. Stability of persistence diagrams. *Comput. Geom*, 37:103–120, 2007.
- [21] James F. Davis and Paul Kirk. *Lecture Notes in Algebraic Topology (Graduate Studies in Mathematics, 35)*. American Mathematical Society, 2001.
- [22] Vin de Silva and Robert Ghrist. Coverage in sensor networks via persistent homology. *Algebraic & Geometric Topology*, 7:339–358, 2007.

- [23] Vin de Silva and Robert Ghrist. Homological sensor networks. *Notices Amer. Math. Soc.*, 54(1):10–17, 2007.
- [24] T. K. Dey. *Curve and Surface Reconstruction : Algorithms with Mathematical Analysis*. Cambridge University Press, 2007.
- [25] Paweł Dłotko, Robert Ghrist, Mateusz Juda, and Marian Mrózek. Distributed computation of coverage in sensor networks by homological methods. *Applicable Algebra in Engineering, Communication and Computing*, 23:29–58, 2012.
- [26] Herbert Edelsbrunner and John Harer. *Computational Topology. An Introduction*. American Mathematical Society, 2010.
- [27] Herbert Edelsbrunner, David Letscher, and Afra Zomorodian. Topological persistence and simplification. *Discrete & Computational Geometry*, 4(28):511–533, 2002.
- [28] Jennifer Gamble, Harish Chintakunta, and Hamid Krim. Coordinate-free quantification of coverage in dynamic sensor networks. *CoRR*, abs/1411.7337, 2014.
- [29] Robert Ghrist and Abubakr Muhammad. Coverage and hole-detection in sensor networks via homology. In *Fourth International Conference on Information Processing in Sensor Networks (IPSN’05)*, UCLA, pages 254–260, 2005.
- [30] Jacob E. Goodman, Joseph O’Rourke, and Csaba D. Tóth, editors. *Handbook of Discrete and Computational Geometry, Third Edition*. CRC Press, Inc., Boca Raton, FL, USA, 2017.
- [31] Dejan Govc and Primož Skraba. An approximate nerve theorem. *Foundations of Computational Mathematics*, Sep 2017.
- [32] R.E. Greene and H. Wu. c^∞ approximations of convex, subharmonic, and plurisubharmonic functions. *Ann. Sci. École Norm. Sup.*, 12(1):47–84, 1979.

- [33] Karsten Grove. Critical point theory for distance functions. In *Proceedings of the Symposia in Pure Mathematics*, volume 54, 1993.
- [34] Shaun Harker, Miroslav Kramar, Rachel Levanger, and Konstantin Mischaikow. A comparison framework for interleaved persistence modules. arXiv:1801.06725, Jan 2018.
- [35] Allen Hatcher. *Algebraic Topology*. Cambridge University Press, 2001.
- [36] William Julian, Ray Mines, and Fred Richman. Alexander duality. *Pacific J. Math*, 106(1):115–127, 1983.
- [37] Heinrich Jung. Über die kleinste kugel, die eine räumliche figur einschließt. *J. Reine Angew. Math*, 123:214–257, 1901.
- [38] Dmitry Kozlov. *Combinatorial Algebraic Topology*. Springer-Verlag Berlin Heidelberg, 2008.
- [39] Saunders MacLane. *Categories for the Working Mathematician*. Springer-Verlag, New York, 1971. Graduate Texts in Mathematics, Vol. 5.
- [40] Elizabeth Munch, Michael Shapiro, and John Harer. Failure filtrations for fenced sensor networks. *International Journal of Robotics Research*, 2012.
- [41] James R. Munkres. *Elements of Algebraic Topology*. Addison-Wesley Publishing Company, 1984.
- [42] Partha Niyogi, Stephen Smale, and Shmuel Weinberger. Finding the homology of submanifolds with high confidence from random samples. *Discrete & Computational Geometry*, 39(1-3):419–441, 2008.
- [43] Steve Y. Oudot. *Persistence Theory - From Quiver Representations to Data Analysis*, volume 209 of *Mathematical surveys and monographs*. American Mathematical Society, 2015.

- [44] Donald R. Sheehy. A multicover nerve for geometric inference. In *The Canadian Conference in Computational Geometry*, pages 309–314, 2012.
- [45] Vin De Silva and Robert Ghrist. Coordinate-free coverage in sensor networks with controlled boundaries via homology. *International Journal of Robotics Research*, 25:1205–1222, 2006.
- [46] Edwin H. Spanier. *Algebraic Topology*. Springer New York, 1966.
- [47] Ron Wein, Jur van den Berg, and Dan Halperin. Planning high-quality paths and corridors amidst obstacles. *The International Journal of Robotics Research*, 27(11-12):1213–1231, November/December 2008.
- [48] Volkmar Welker, Günter M. Ziegler, and Rade T Zivaljevic. Homotopy colimits – comparison lemmas for combinatorial applications. *J. Reine Angew. Math.*, 509:117–149, 1999.
- [49] Afra Zomorodian and Gunnar Carlsson. Localized homology. *Computational Geometry: Theory and Applications*, 41(3):126–148, 2008.