PyTDA

Kirk P. Gardner[*] Donald R. Sheehy[†]

1 Sensor Networks and Simplicial Complexes

We consider the problem of determining coverage in a coordinate free sensor network. That is, we would like to determine if an unknown domain is covered by a collection of sensors without their precise coordinates. Let $\mathcal{D}$ denote our unknown domain and $P \subset \mathcal{D}$ be a collection of points, each representing a sensor in our network. At the very least each sensor is capable of detecting nodes which are sufficiently "close" in the sense that, if we endow our domain with a metric $\mathrm{d} : \mathcal{D} \times \mathcal{D} \to \mathbb{R}$ there is some radius of communication $\alpha > 0$ such that two nodes $p, q \in P$ such that $\mathrm{d}(p, q) \leq \alpha$ are capable of communication. Note that, although sensors can communicate within this distance they are not able to measure the distance itself.

With this limited capability we can construct an undirected graph $G = (V, E)$ with vertices $V = P$ and edges $E = \{\{p, q\} \subset P \mid \mathrm{d}(p, q) \leq \alpha\}$. In order to determine coverage we must at least assert that the coverage domain spanned by the points in P does not contain any holes. Assuming the coverage radius of our sensors is equal to their communication radius α we may define a hole in coverage as a cycle that cannot be "filled" with triangles. This leads us to a more natural construction known as a simplicial complex.

Definition 1. *A **simplicial complex** K is a collection of subsets, called **simplices**, of a vertex set V such that for all $\sigma \in K$ and $\tau \subset \sigma$ it must follow that $\tau \in K$.*

The **dimension** of a simplex $\sigma \in K$ is defined as $\dim(\sigma) := |\sigma| - 1$ where $|\cdot|$ denotes set cardinality. The dimension of a simplicial complex K is the maximum dimension of any simplex in K. That is, a graph is a 1-dimensional simplicial complex in which vertices and edges are 0 and 1-dimensional simplices, respectively.

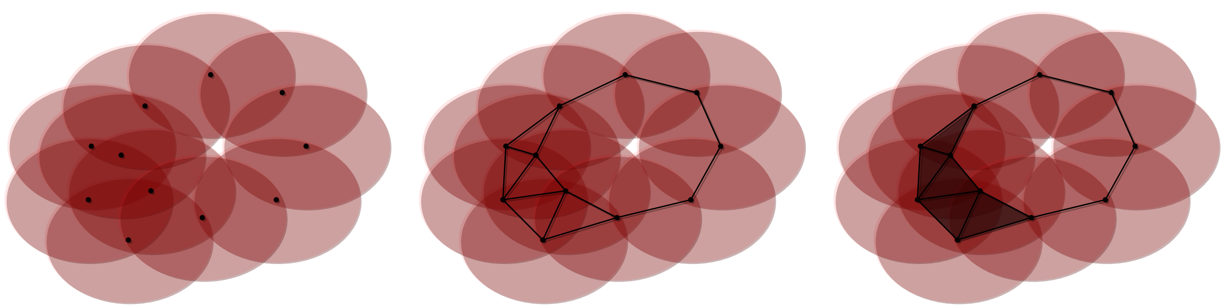

Figure 1:

It is natural to think of a k-dimensional simplicial complex as the generalization of an undirected graph consisting of vertices and edges, collections of at most 2 vertices, to collections of sets of at most $k-1$ vertices. Just as we have defined a hole in our graph G as a cycle that cannot be filled with triangles, we define a k-dimensinal hole in a simplicial complex as a k-cycle that cannot be filled with $(k+1)$-simplices. In the next section we will formally define k-cycles and introduce simplicial homology as a tool for identifying when and which cycles cannot be filled.

[*]University of Connecticut.
[†]University of Connecticut.

Let K be a simplicial complex with 0-simplices $\{v\}$ for all $p \in P$, 1-simpices $\{u, v\} \subset P$ for each edge in E, and 2-simplices $\{u, v, w\} \subset P$ whenever $\{\{\{u, v\}, \{v, w\}, \{u, w\}\} \subset E$. This particular simplicial complex is known as the Vietoris-Rips complex.

Definition 2. *The **(Vietoris-)Rips complex** is defined for a set P at scale $\varepsilon > 0$ as*

$$\mathcal{R}^{\varepsilon}(P) = \{\sigma \subseteq P \mid \forall p, q \in \sigma,\ \mathrm{d}(p, q) \leq \varepsilon\}.$$

1.1 Coverage

A subset D of an domain $\mathcal{D}$ is covered by $P \subset$ at scale $\varepsilon > 0$ if every point $x \in D$ is within distance α at least one point in P.

Definition 3. *The **coverage region** of a point $p \in P$ at scale $\varepsilon > 0$ is defined*

$$\mathrm{ball}_{\varepsilon}(p) = \{x \in \mathcal{D} \mid \mathrm{d}(x, p)\}.$$

Let P^{ε} to denote set of points in $\mathcal{D}$ within distance ε of at least one point in P:

$$P^{\varepsilon} = \bigcup_{p \in P} \mathrm{ball}_{\varepsilon}(p).$$

Definition 4. *Let $D, P \subset \mathcal{D}$. D is **covered** by P at scale ε if $D \subseteq P^{\varepsilon}$.*

For this section we will assume that the coverage radius of our sensor network is equal to the radius of communication. With this assumption we know that the converse problem is true: if a sensor network P covers a domain at scale α the topology of the domain is reflected in $\mathcal{R}^{\alpha}(P)$ however, as we will see this is in no way a tight bound on the minimal radius for coverage.

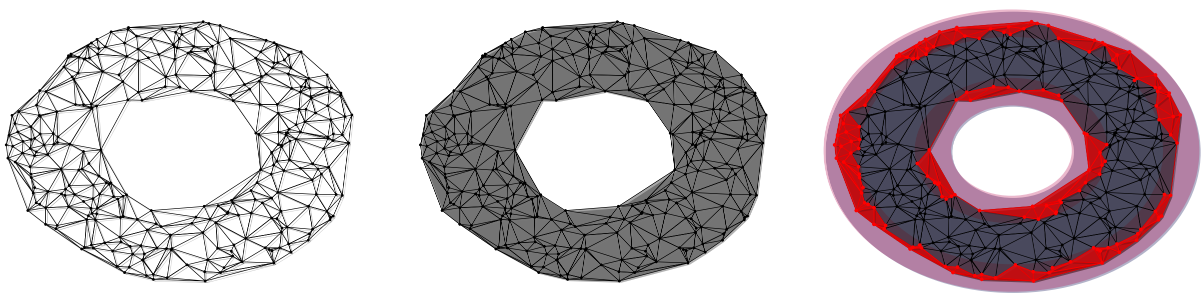

Figure 2:

In order to *verify* coverage by we need our network to sufficiently sample the extent of our domain. Moreover, if there are gaps in coverage, are they due to insufficient sampling or a gap in the domain itself? Let $\mathcal{B} \subset \mathcal{D}$ be the **boundary** of our domain and let our sensors detect whe the are within communication radius of $\mathcal{B}$. Let $Q = \{p \in P \mid \min_{x \in \mathcal{B}} \mathrm{d}(x, p)\}$ be the set of **boundary nodes** in P. The set Q induces a **subcomplex** $\mathcal{R}^{\alpha}(Q)$ of $\mathcal{R}^{\alpha}(K)$ restricted to nodes in Q.

TODO Section overview. We will show how this construction is used to verify coverage of a specific subset of a bounded domain $\mathcal{D}$ with a tight bound on the coverage radius.

2 Functions on Complexes

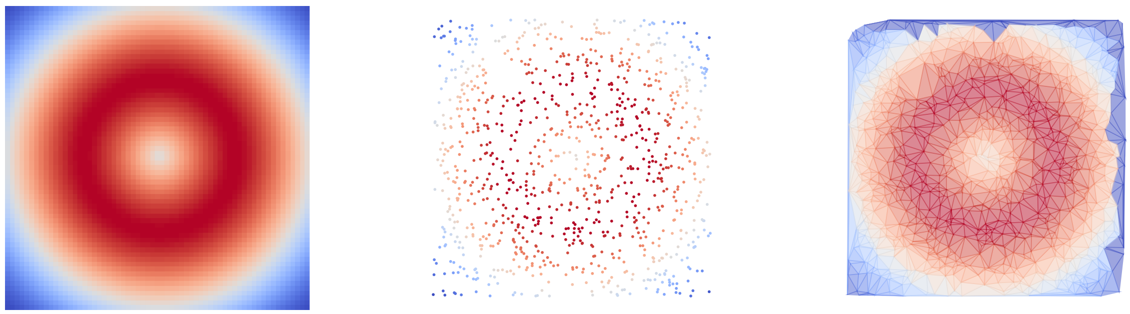

Figure 3: (Left) Function on the plane. (Middle) Function values on random sample. (Right) Function values on the 2-simplices of a simplicial complex on the random sample.

3 Homology and Cohomology

Simplicial complexes not only provide a comprehensive discretization of continuous domains, but are the primary tool for concrete calculations in algebraic topology. In particular, the study of simplicial homology groups and its dual cohomology groups rely on simplicial complexes in order to study important topological invariants of a discretized space.

3.1 Homology

The following vector spaces may be defined over any field $\mathbb{F}$, however we will assume the field $\mathbb{F}_2$ in order to avoid orienting the simplices in K. Let $C_k(K)$ denote the vector space over some field $\mathbb{F}$ consisting of linear combinations of k-simplices in K, which form a basis for $C_k(K)$, known as ***k*-chains**. These vector spaces are connected by **boundary maps** $\partial_k : C_k(K) \to C_{k+1}(K)$ which are linear transformations taking basis elements of $C_k(K)$ to the abstract sum of basis $(k+1)$-simplex faces. The collection of chains and boundary maps forms a sequence of vector spaces known as the **chain complex** $\mathcal{C} = (C_*, \partial_*)$

$$\ldots \xrightarrow{\partial_{k+1}} C_k(K) \xrightarrow{\partial_k} C_{k-1}(K) \xrightarrow{\partial_{k-1}} \ldots \xrightarrow{\partial_2} C_1(K) \xrightarrow{\partial_1} C_0(K) \xrightarrow{\partial_0} 0.$$

An important property of the boundary maps ∂_k is that the composition of subsequent boundary maps is null. That is, for all k

$$\partial_k \circ \partial_{k-1} = 0.$$

As a result the image of ∂_{k+1}, denoted im $\partial_{k+1} = \{\partial_{k+1} c \mid c \in C_{k+1}(K)$ is a subspace of the kernel, $\ker \partial_k = \{c \in C_k(K) \mid \partial_k c = 0\}$, of ∂_k. A ***k*-cycle** of $\mathcal{C}$ is a k-chain with empty boundary - an element of $\ker \partial_k$. Two cycles in $\ker \partial_k$ are said to be **homologous** if they differ by an element of im ∂_{k+1}. This leads us to the definition of the **homology groups** of K as quotient vector spaces $H_k(K)$ over $\mathbb{F}$, defined for $k \in \mathbb{N}$ as

$$H_k(K) := \ker \partial_k / \text{im } \partial_{k+1}.$$

The rank of each homology group is of particular importance and are known as the **Betti numbers** $\beta_k = \text{rk} H_k(K)$. These topological invariants can be thought of as counting the number of k-dimensional "holes" in a topological space, where 0-dimensional holes are connected components, 1-dimensional holes are loops, 2-dimensional holes are voids, and so on. Note that this is the same notion which motivated our use of simplicial complexes for determining coverage - a 1-dimensional hole exists if a gap in a neighborhood graph cannot be filled by triangles.

3.2 Cohomology

A **cochain complex** is a sequence $\mathcal{C} = (C^*, \delta_*)$ of $\mathbb{R}$-modules C^k consisting of **cochains** and module homomorphisms known as **coboundary maps** $\delta_k : C^k \to C^{k+1}$. As in homology we have the property that $\delta_{k+1} \circ \delta_k = 0$ for all k, leading to a familiar definition of the **cohomology** of $\mathcal{C}$

$$H^k(\mathcal{C}) = \ker \delta_k / \text{im } \delta_{k-1}.$$

The equivalence classes of $H^k(\mathcal{C})$ consist of ***k*-cocycles**: elements of $\ker \delta_k$ that differ by a ***k*-coboundary** in im ∂_{k-1}. Such cocycles are said to be **cohomologous** if they belong to the same equivalence class in $H^k(\mathcal{C})$.

The simplest construction of a cochain complex is to dualize a chain complex. For a simplicial complex K with chain complex (C_*, ∂_*) define $C^k(K)$ to be the module of homomorphisms $\psi : C_k \to \mathbb{R}$. The coboundary maps δ_k are defined for cochains $\psi : C_k \to \mathbb{R}$ and k-simplices $\sigma \in K$ as

$$\delta_k \psi(\sigma) = \psi(\partial_k \sigma).$$

3.3 Relative Homology

For a simplicial complex K let $L \subset K$ be a subcomplex of K. Let $\mathcal{C}(K, L)$ denote the quotient chain complex of pairs $(C_k(K, L), \overline{\partial_k})$ where $C_k(K, L) = C_k(K)/C_k(L)$ consists of the chains on K modulo chains on L, with the induced boundary maps $\overline{\partial_k}$ on the quotients. Each relative chain is an equivalence class of chains in K which are identical without the elements of L. The **relative homology groups** $H_k(K, L)$ consists of homology classes of relative cycles - chains in K whose boundaries vanish or lie in L. That is, a relative cycle can either be a cycle in K or a chain in K with a boundary in L.

As we will see, relative homology provides a powerful representation of a bounded domain that is particularily suited to verifying coverage of a sensor network. Let $\mathcal{D}$ be a bounded domain with boundary $\mathcal{B} \subset \mathcal{D}$. For illustrative purposes assume $\mathcal{D}$ is a connected, compact subset of the euclidean plane $\mathbb{R}^2$ so that certain properties of the relative homology of the pair $(\mathcal{D}, \mathcal{B})$ are known. Namely, there is exactly one equivalence class in $H_2(\mathcal{D}, \mathcal{B})$, as illustrated in figure 4. We can think of the quotient as an identification of points in the boundary, illustrated by wrapping the planar domain around a single point in $\mathbb{R}^3$. As the domain is compact this creates a single void corresponding to the one generator in $H_2(\mathcal{D}, \mathcal{B})$.

Figure 4:

Now suppose our network P covers the domain at some scale $\alpha > 0$ such that there are no gaps (1-cycles) in $\mathcal{R}^\alpha(P)$. The subset $Q = \{p \in P \mid \mathrm{ball}_\alpha(p) \cap \mathcal{B} \neq \emptyset\}$ of points within distance α of $\mathcal{B}$ induces a subcomplex $\mathcal{R}^\alpha(Q)$ of $\mathcal{R}^\alpha(K)$. Under our assumptions the relative homology $H_2(\mathcal{R}^\alpha(K), \mathcal{R}^\alpha(Q))$ should reflect that of the domain. A gap in coverage can be thought of as "popping the balloon" in the sense that, if we wrap the simplices of $\mathcal{R}^\alpha(K)$ around those in $\mathcal{R}^\alpha(Q)$ we would have no void - the gap provides a hole through which the "air" can escape, as illustrated in figure 5.

Figure 5:

3.4 Representative Cycles and Cocycles

We can find a basis for each homology group $H_p(K)$ and cohomology group $H^p(K)$ consisting of p-cycles and p-cocycles, respectively. In general, p-cycles represent p-dimensional holes in the simplicial complex K, where p-cocycles can be understood as "blocking chains."

As we will see representative cocycles are particularily useful in distributing the information contained in each Cohomology basis element throughout a simplicial complex. While this does not have an immediate application to our investigation of coverage in homological sensor network it is a powerful tool which may be used for future research in coordinate free sensor networks.

3.5 Persistent Homology and Cohomology

Topological data analysis is an emerging field in the intersection of data analysis and algebraic topology which extends the notion of homology and cohomology groups to a more analytical tool known as **topological persistence**. Where simplicial (co)homology identifies invariants of a static simplicial complex persistent (co)homology tracks the evolution of a sequence of nested simplicial complexes which provide a more detailed topological signature, in addition to relevant geometric information.

Definition 5. *A **filtration** of a simplicial complex K is a nested sequence of simplicial complexes*

$$K_0 \subset K_1 \subset \dots K_{n-1} \subset K_n = K.$$

A filtration $\{K_i\}_{i=1,\dots,n}$ may also be interpreted as a sequence of simplicial maps, each an inclusion $K_i \to K_{i+1}$. The resulting sequence induces an algebraic sequence of homomorphisms on (co)homology by the functoriality, for all k:

$$H_k(K_0) \to H_k(K_1) \to \dots \to H_k(K_{n-1}) \to H_k(K_n) = H_k(K).$$

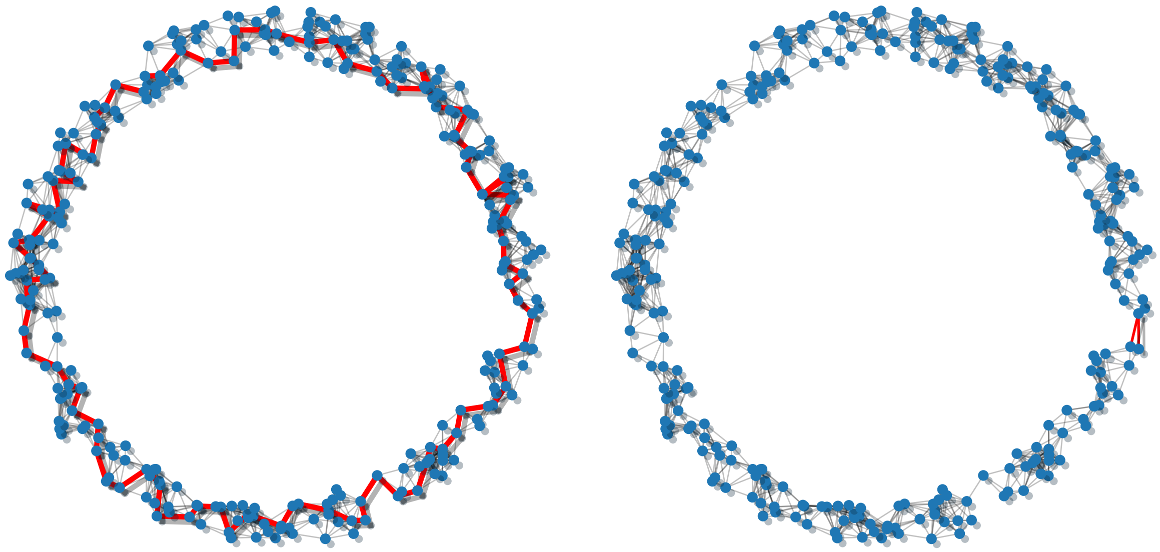

Figure 6:

This sequence encodes the local topological changes that occur at each step of the filtration. Global information is encoded in terms of the **birth** and **death** of (co)homology classes, represented as a **persistence diagram** or **barcode**.

Given a rips complex $K = \mathcal{R}^\alpha(P)$ of a finite metric space P consisting of n simplices we can order the simplices $\sigma_1, \ldots, \sigma_n$ by the minimum pairwise distance between their vertices. We first order the vertices $v_1, \ldots, v_m$ of P and let $\sigma_i = \{v_i\}$ for $i = 1, \ldots, m$ so that $K_i = \{\sigma_1, \ldots, \sigma_m\}$. We can then build a filtration $\{K_i\}_{i=1,\ldots,n}$ so that $K_i = \mathcal{R}^\varepsilon(P)$ where $\varepsilon = \max_{u,v\in\sigma_i} \mathrm{d}(u, v)$ by adding one simplex at a time, breaking ties first by dimension, then by the ordering on their constituent vertices. A k-dimensional feature is identified when a $(k + 1)$-simplex σ is added that kills a k-cycle γ. In the persistence diagram representation this feature would be represented by a point (b, d) where b is the smallest scale for which the k-cycle appears

$$\tau \in \mathcal{R}^b(P) \text{ for all } \tau \in \gamma$$

and $d = \max_{u,v\in\sigma} \mathrm{d}(u, v)$ is the scale at which σ enters the filtration. The result is a collection of points (b_i, d_i) in the half-plane for each dimension known as the persistence diagram.

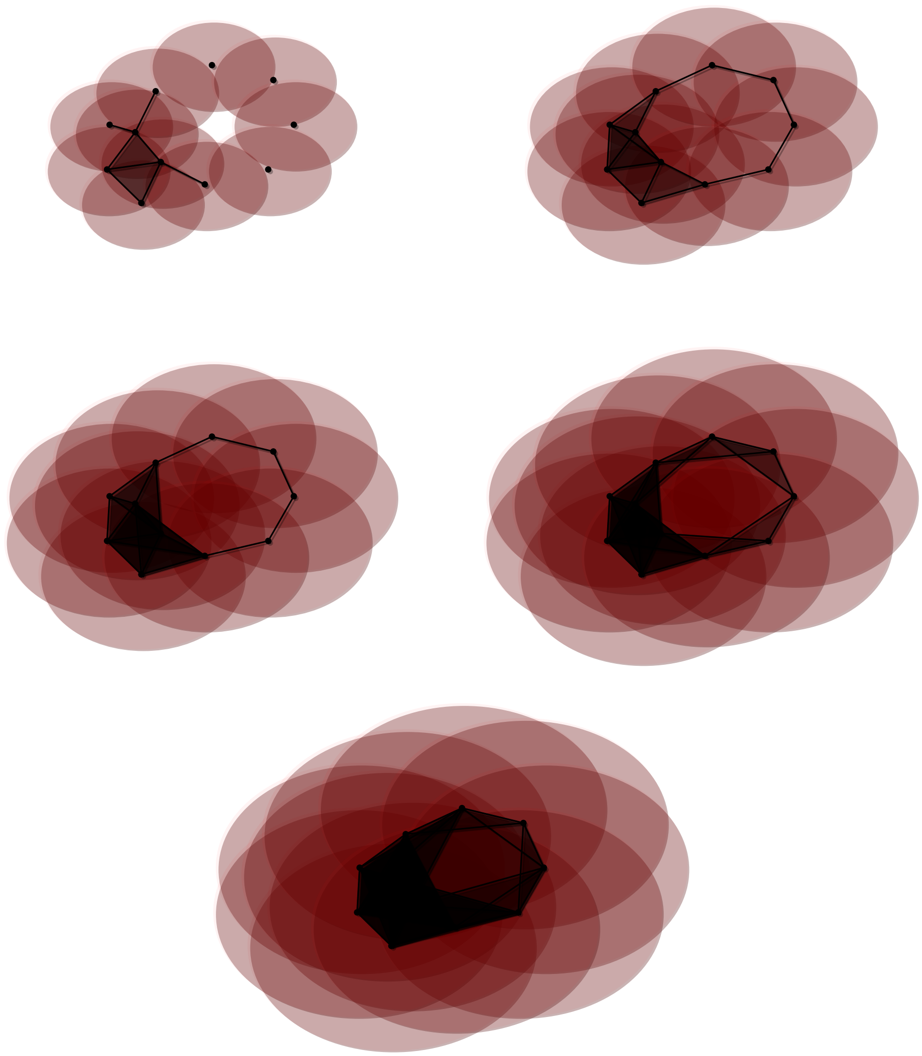

Figure 7: A filtration of rips complexes at scales 0.6, 0.8, 1.0, 1.2, and 1.4. The 1-cycle that is born at scale 0.8 persists until it dies at scale 1.4, resulting in a point $(0.8, 1.4)$ on the persistence diagram.

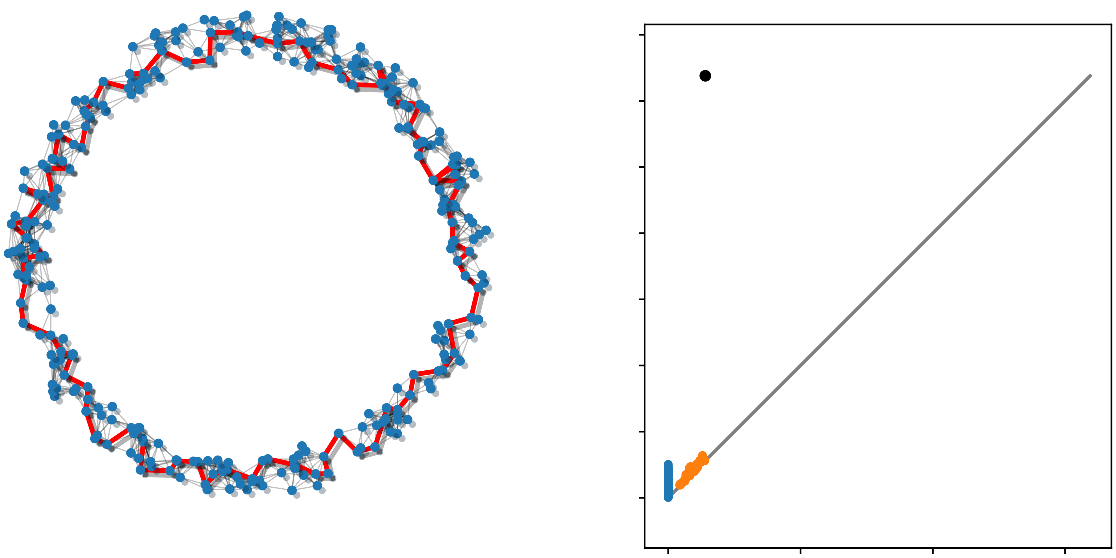

Figure 8: The representative cycle of a significant feature in the persistent homology of rips filtration of a noisy circle. The birth of the feature indicated on the persistence diagram (right) corresponds to the scale of the rips complex shown (left) when the circle, a 1-cycle, is born. The death of this feature corresponds to the scale of the rips complex at a larger scale (not shown) when a triangle first fills the interior of the circle. This scale is approximately the length of the edges in the smallest equilateral triangle with sample points as vertices that contains the centroid of the sample. This illustrates the geometric information encoded in the persistence diagram of geometric complexes as it is within a constant factor of the radius of the circle.

4 Discrete Exterior Calculus

4.1 Circular Coordinates

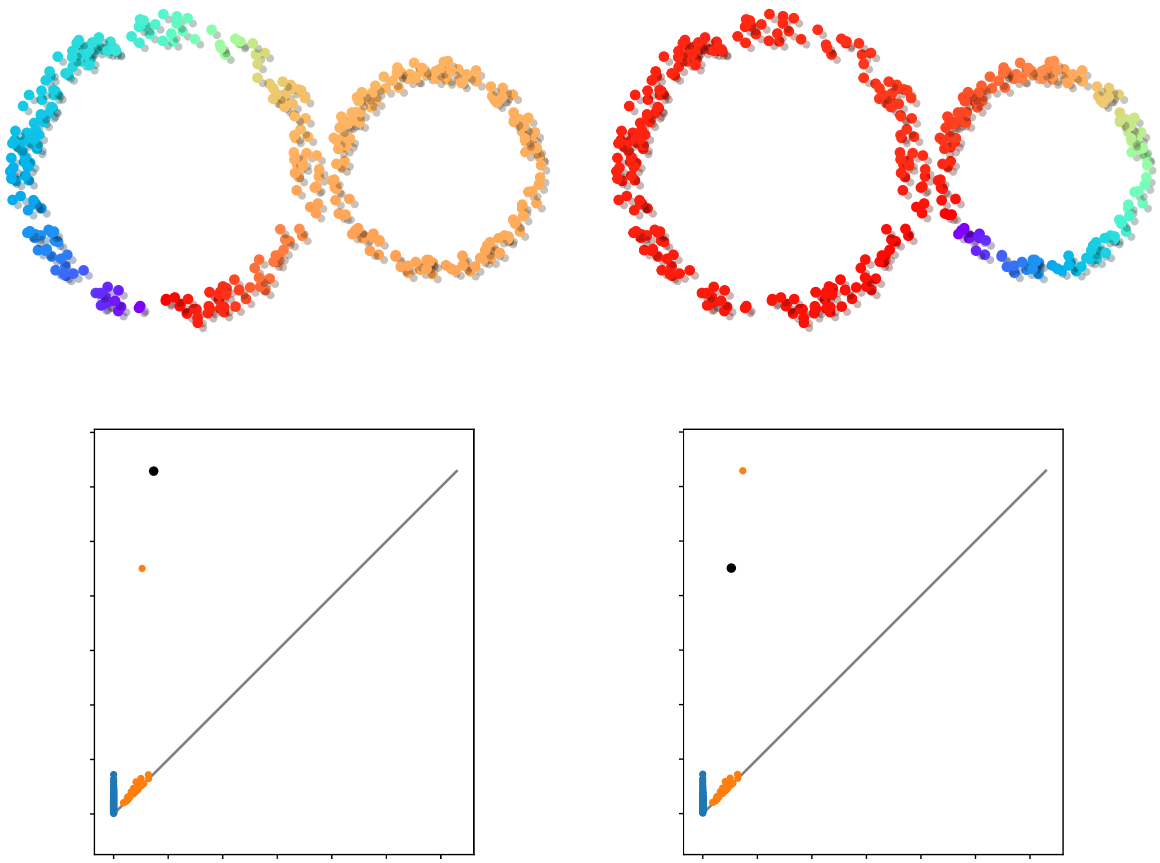

Figure 9:

5 The Topological Coverage Criterion

There is a specific simplicial complex defined for a collection of sets that precisely captures the coverage information we require. The **Čech complex** of a finite collection of points P at scale $\varepsilon > 0$ is defined

$$\check{\mathcal{C}}^{\varepsilon}(P) := \left\{ \sigma \subseteq P \mid \bigcap_{p \in \sigma} \mathrm{ball}_{\varepsilon}(p) \neq \emptyset \right\}.$$

The Čech complex is a special case of a more general construction known as a the **nerve** $\mathbb{N}(\mathcal{U})$ of a collection of sets $\mathcal{U} = \{U_i\}_{i \in I}$, where I is any indexing set. The nerve of $\mathcal{U}$ is defined as the simplicial complex with vertex set I such that $\sigma \subseteq I$ is a simplex if and only if

$$\bigcap_{i \in \sigma} U_i \neq \emptyset.$$

The collection $\mathcal{U}$ is a **good cover** if for each $\sigma \subset I$ the set $\bigcap_{i \in \sigma} U_i$ is contractible if it is nonempty. The **nerve lemma** states that if $\mathcal{U}$ is a good cover then its nerve $\mathbb{N}(\mathcal{U})$ is homotopy equivalent to $\bigcup_{i \in I} U_i$. That is, for a set of nodes $P \subset \mathcal{D}$ such that $\mathcal{U} = \{\mathrm{ball}_{\varepsilon}(p) \mid p \in P\}$ is a good cover the nerve $\mathbb{N}(\mathcal{U})$ is homotopy equivalent to $P^{\varepsilon} = \bigcup_{p \in P} \mathrm{ball}_{\varepsilon}(p)$. It follows that the Čech complex $\check{\mathcal{C}}_{\varepsilon}(P)$ of P at scale ε is a suitable representation of the coverage region P^{ε}.

While the Čech complex captures the topology of the coverage region in question exactly it can only be computed when the precise pairwise distances between nodes is known. If instead we are only provided pairwise *proximity* information indicating when nodes are within some fixed distance we may use the **(Vietoris-)Rips complex**, defined for a set P at scale $\varepsilon > 0$ as

$$\mathcal{R}^{\varepsilon}(P) := \{\sigma \subseteq P \mid \forall p, q \in \sigma,\ \mathrm{d}(p, q) \leq \varepsilon\}.$$

An important result about the relationship of Čech and Rips complexes follows from Jung's Theorem [2] relating the diameter of a point set P and the radius of the minimum enclosing ball:

$$\check{\mathcal{C}}^{\varepsilon}(P) \subseteq \mathcal{R}^{\varepsilon}(P) \subseteq \check{\mathcal{C}}^{\vartheta_d \varepsilon}(P), \tag{1}$$

where the constant $\vartheta_d = \sqrt{\frac{2d}{d+1}}$ (see [1]).

As we will see the Rips complex may be used to verify coverage of a domain satisfying some minimal assumptions when we allow the sensors in our network to communicate at two radii α and β. In short, if the the inclusion of Rips complexes at scales $\alpha < \beta$ resembles the structure of a subset of the domain then the Čech complex at scale α, and therefore P^{α}, does as well.

5.1 Implementation

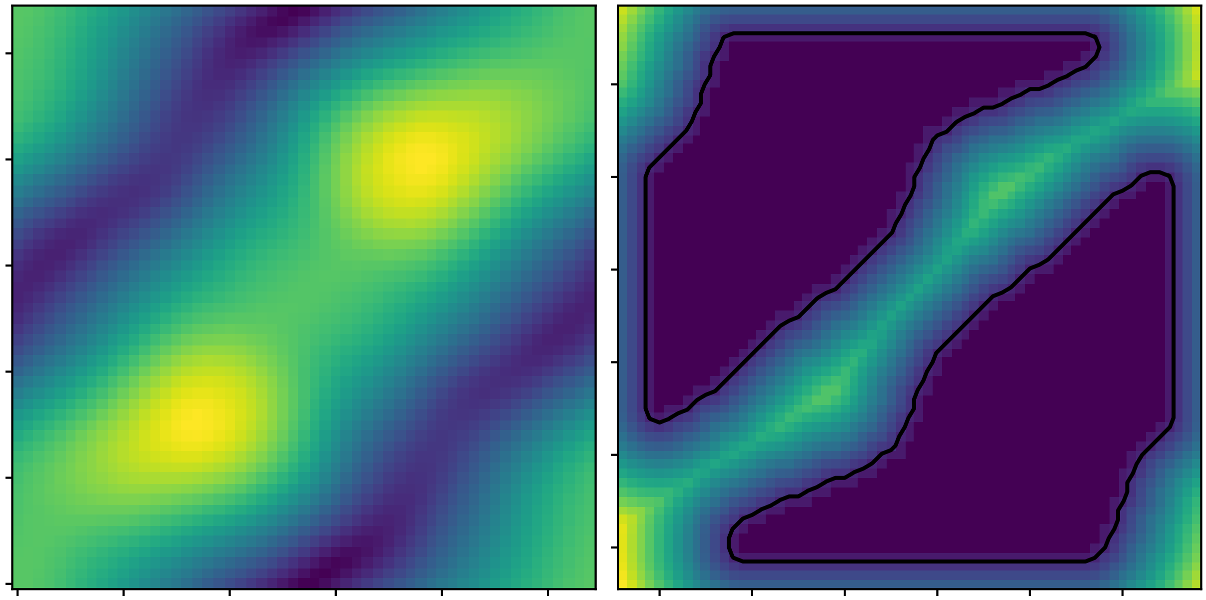

Figure 10:

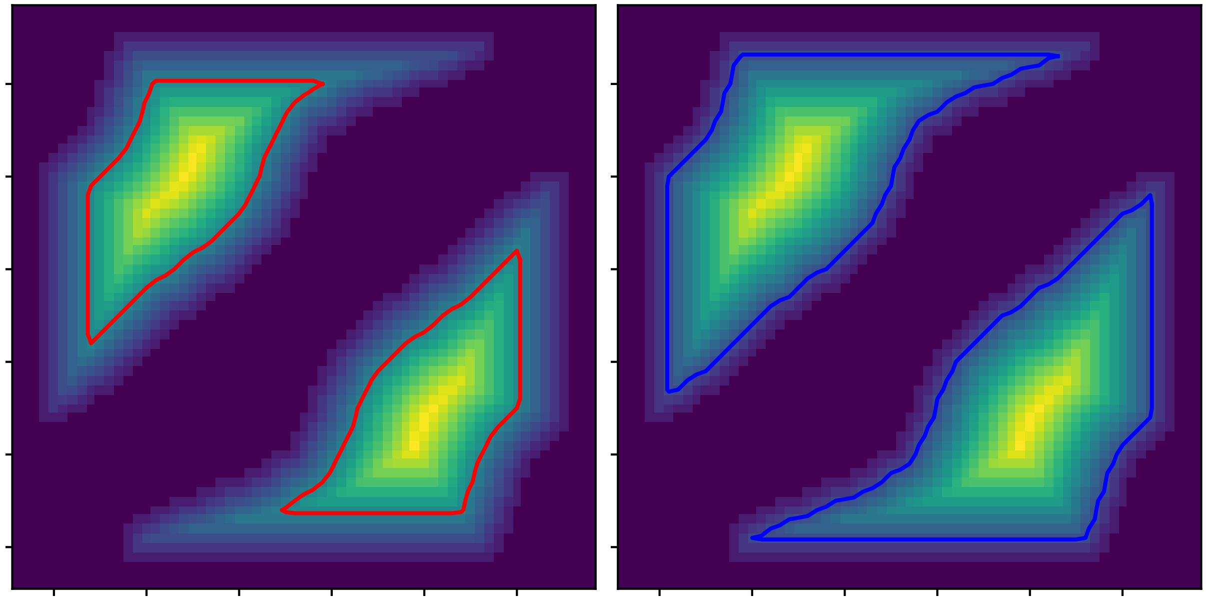

Figure 11:

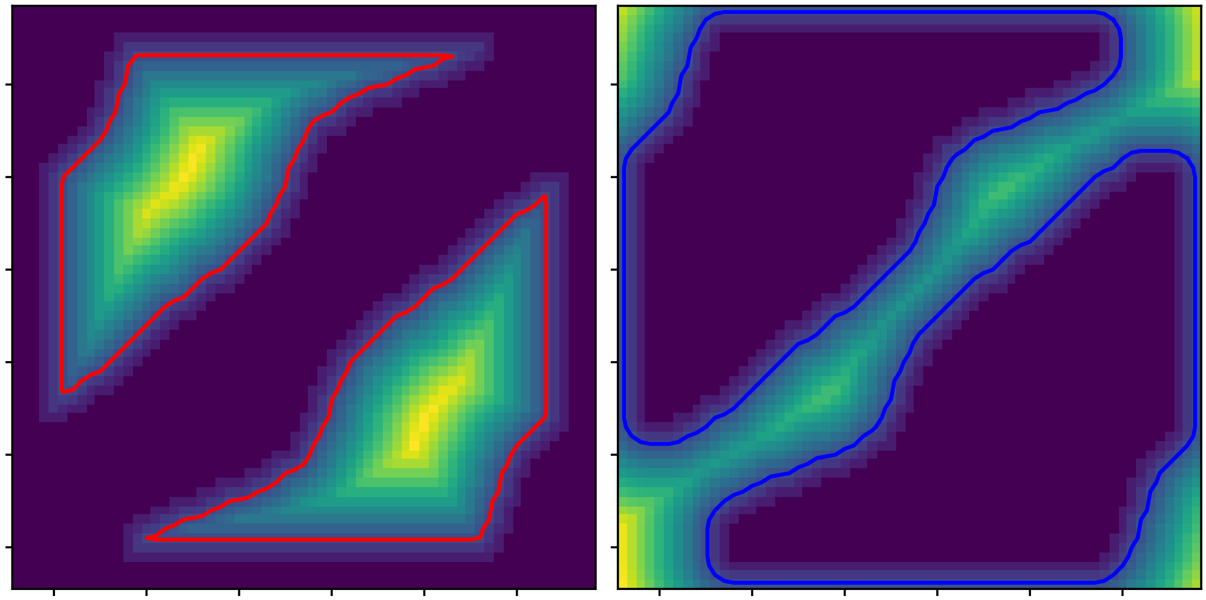

Figure 12:

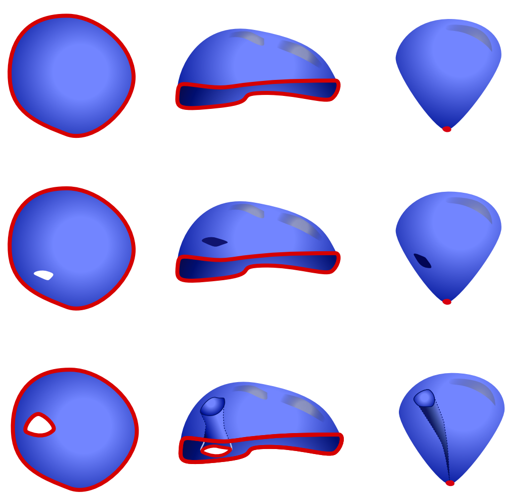

Figure 13:

References

[1] Mickaël Buchet, Frédéric Chazal, Steve Y. Oudot, and Donald R. Sheehy. Efficient and robust persistent homology for measures. In *ACM-SIAM Symposium on Discrete Algorithms*, pages 168–180, 2015.

[2] Heinrich Jung. Über die kleinste kugel, die eine räumliche figur einschließt. *J. Reine Angew. Math*, 123:214–257, 1901.