

# Persistence by Parts: Multiscale Feature Detection via Distributed Persistent Homology

Hee Rhang Yoon\* and Robert W. Ghrist†

## Abstract

A method is presented for the distributed computation of persistent homology, based on an extension of the generalized Mayer-Vietoris principle to filtered spaces. Cellular cosheaves and spectral sequences are used to compute global persistent homology based on local computations indexed by a scalar field. These techniques permit computation localized not merely by geography, but by other features of data points, such as density. As an example of the latter, the construction is used in multi-scale analysis of point clouds to detect features of varying sizes that are overlooked by standard persistent homology.

## 1 Introduction

### 1.1 Cosheaves and computational persistence

Persistence has emerged as a central principle in Topological Data Analysis (TDA). As its core, persistence exploits the functoriality of homology to extract robust features from point cloud data. Given a filtered cell complex arising from such data, its homology is a persistence module – a sequence of vector spaces and linear transformations. The representation theory of such objects leads to an unambiguous decomposition into indecomposables (persistent homology classes) [4, 5, 6] which, thanks to the Stability Theorem for persistent homology [11, 10], is robust with respect to perturbations of the original data points. For details, see §2.1.

There is no inherent obstruction to computing homology from local data: Mayer-Vietoris and spectral sequences lead the way. The story is more complicated for persistent homology, due to the algebra of persistence modules. The approach of this paper is to use theory of cellular cosheaves to store and relate local homology information over the complex. This theory, like the algebraically dual theory of cellular sheaves [12], is ideal for local-to-global operations.

The main result of this paper is an application of spectral sequences and the Mayer-Vietoris principle to compute persistent homology by breaking a filtered complex into pieces based on a scalar field on the point cloud. This has utility beyond the obvious idea of breaking up a complex into parts based on geographic proximity.

In its typical interpretation, homology classes which persist over large parameter intervals are the statistically significant features – those which are not artefacts of noisy sampling. For Čech complexes of manifolds sampled with sufficiently high density, a single (practically unobtainable) homology computation yields truth [21]. For the (more realistic) case of a filtered Vietoris-Rips complex, persistent homology is a good first approximation to truth. Questions of a more epistemological bent (“*Are these really the important features?*”) are evident. Recent work of Berry and Sauer argues that persistent homology can erroneously label important small features as non-persistent, and they propose a *Continuous k-Nearest Neighbors* graph to estimate the geometry of a discretely sampled manifold [2]. They prove that this method captures multiscale connectivity data and conjecture that it works in higher homologies.

As an application of our results, we show how to use a local density as the distribution parameter in order to retrieve persistent homology classes that are weighted according to the density of sampling. The idea is

---

\*Department of Electrical & Systems Engineering, University of Pennsylvania, Philadelphia, PA (yoonhee@seas.upenn.edu)

†Department of Mathematics and Electrical & Systems Engineering, University of Pennsylvania, Philadelphia, PA (ghrist@math.upenn.edu)

that when faced with geometrically small but tightly-sampled homology classes which may be of significance, one can compute persistent homology using distance as the (usual) parameter, and sampling density as the distribution field: see §5. The oft-expressed desire of doing multiparameter persistence on both distance and density [4, 17] is full of algebraic challenges [8, 7]; in a sense, this paper gives a novel approach by separating out one parameter as a scalar field for distributed computation.

## 1.2 Related and supporting work

Algorithms for computing persistent homology now comprise a rich and intricate literature. The original algorithm for computing persistent homology [15, 14] computes persistence pairs by reducing the boundary operator. Since then, variations of the original algorithm have been developed to improve computation [22]. In particular, parallelized and distributed algorithms have been proposed in order to improve memory usage and computation time. The spectral sequence algorithm from [14] reduces blocks of matrices at each phase, which results in persistence pairs of particular lengths. In [1], the authors incorporate optimization techniques and construct a distributed algorithm of the spectral sequence algorithm in both shared and distributed memory.

The above mentioned algorithms distribute data with respect to ranges of filtration values. In [19], the authors provide a distributed computation algorithm in which data is distributed with respect to spatial decomposition of the domain. They build a Mayer-Vietoris blowup complex, and its boundary matrix is reduced by reducing submatrices in parallel. Our work shares a similar philosophy, as we distribute data according to spatial decomposition of the domain and we operate on the foundation of Mayer-Vietoris principle. Instead of using the geometric construction of Mayer-Vietoris blowup complex, we use the algebraic construction of cosheaf homology to combine local information. Our approach adapts the use of cosheaf homology to compute homology from subspaces [13], with cosheaf morphisms and spectral sequences to take the filtration into account.

This work has its origins in the Ph.D. dissertation of the first author [24]. After preparing this article, the preprint of Casas appeared [9], which, influenced by [24], builds on and extends the results. In particular, Algorithm 2 of [9] recapitulates §4.2.3 of [24]. Casas greatly extends the results of that thesis and this paper by not limiting the nerve of the distribution cover to 1-d; however, in the restricted case considered here, Casas’ diagram chase is exactly that of [24] and the present work.

## 1.3 Problem statement and contributions

We address the following question.

*Given a point cloud  $P$  (a finite subset of Euclidean  $\mathbb{R}^m$ ), compute the persistence module of  $P$  from local persistence modules subordinate to a cover of  $P$*

Let  $\mathbb{V}$  denote the persistence module obtained from  $P$  — a finite sequence of vector spaces and linear maps that encodes topological information about Vietoris-Rips filtration associated to  $P$ . We use cellular cosheaf homology to assemble the relevant information gathered from subsets of  $P$  subordinate to a cover. Morphisms of cellular cosheaves then allow us to incorporate persistence. We use spectral sequences to discover a hidden map among cosheaf homologies. Our main result, as stated in the following theorem, is a distributed construction of a persistence module  $\mathbb{V}_\Psi$  that is isomorphic to the persistence module  $\mathbb{V}$  of interest.

**Theorem 1.** *The local  $\mathbb{V}_\Psi$  and global  $\mathbb{V}$  persistence modules are isomorphic.*

We argue that this distributed computation can be used to filter and annotate persistent homology based on meta-data associated to the point-cloud, using such to build a cover. We illustrate this idea in the case where the meta-data is a sampling density estimate, yielding a method for computing multiscale persistent homology, identifying significant features that are overlooked by standard persistent homology methods.

Section 2 contains a summary of persistent homology and cellular sheaf theory. We review a distributed computation method for homology using Leray cellular cosheaves in §3. In §4, we construct a general distributed persistent homology computation mechanism subordinate to a cover with at most pairwise overlaps. Finally, in §5, we indicate the utility of this distributed computation in multiscale persistent homology, using our methods to identify persistent homology classes relative to a sampling density.

## 2 Background

Throughout this paper, we assume that every homology is computed with coefficients in a field  $\mathbb{K}$ .

### 2.1 Persistent Homology

Given a point cloud, one can interrogate its global structure via persistent homology. Let  $P$  be a finite collection of points in Euclidean  $\mathbb{R}^m$ . The **Vietoris-Rips complex**  $\mathcal{R}^\epsilon$  is the simplicial complex whose  $k$ -simplices correspond to  $(k + 1)$ -tuple of points from  $P$  that have pairwise distance  $\leq \epsilon$ . For brevity, we use the term ‘‘Rips complex’’ to refer to the Vietoris-Rips complex.

Assume that  $(\mathcal{R}^i)_{i=1}^N$  is a sequence of Rips complexes over  $P$  for increasing parameter values  $(\epsilon_i)_{i=1}^N$ , with inclusion maps between each pair of Rips complexes

$$\mathcal{R}^1 \hookrightarrow \mathcal{R}^2 \hookrightarrow \dots \hookrightarrow \mathcal{R}^N.$$

By applying the homology functor, one obtains the following sequence of vector spaces

$$\mathbb{V} : H_\bullet(\mathcal{R}^1) \rightarrow H_\bullet(\mathcal{R}^2) \rightarrow \dots \rightarrow H_\bullet(\mathcal{R}^N). \quad (1)$$

The above sequence is an instance of a **persistence module**. For the purpose of this paper, it suffices to consider a persistence module as a finite sequence of vector spaces and linear maps between them.

A **morphism of persistence modules**  $\alpha : \mathbb{V} \rightarrow \mathbb{W}$  is a collection of linear maps  $\alpha_i : V_i \rightarrow W_i$  such that the following diagram commutes.

$$\begin{array}{ccccccc} V_1 & \longrightarrow & V_2 & \longrightarrow & \dots & \longrightarrow & V_n \\ \downarrow \alpha_1 & & \downarrow \alpha_2 & & & & \downarrow \alpha_n \\ W_1 & \longrightarrow & W_2 & \longrightarrow & \dots & \longrightarrow & W_n \end{array}$$

When all  $\alpha_i$ ’s are isomorphisms, then  $\alpha$  is an isomorphism of persistence modules.

By the Structure Theorem [25], the persistence module from Equation (1) decomposes uniquely as

$$\mathbb{V} \cong \bigoplus_{l=1}^N \mathbb{I}(b_l, d_l),$$

where each  $\mathbb{I}(b, d)$  is a simpler persistence module of the form

$$\mathbb{I}(b, d) : 0 \rightarrow \dots \rightarrow 0 \rightarrow \mathbb{K} \xrightarrow{1} \mathbb{K} \xrightarrow{1} \dots \xrightarrow{1} \mathbb{K} \rightarrow 0 \rightarrow \dots \rightarrow 0.$$

The  $b$  and  $d$  each indicates the first and last index of  $\mathbb{K}$ . Each  $\mathbb{I}(b, d)$  represents a homological feature with birth time  $b$  and death time  $d$ . One can visualize such birth and death times of  $\mathbb{I}(b, d)$  using a **barcode**. Given a persistence module  $\mathbb{V}$ , a barcode diagram,  $barcode(\mathbb{V})$ , is a collection of bars that correspond to the intervals  $(b, d)$  obtained from the decomposition of  $\mathbb{V}$ . In simple settings, long bars of  $barcode(V)$  capture significant homological features; shorter bars may be due to noise.

### 2.2 Cellular Cosheaves

A cellular cosheaf is a certain assignment of algebraic structure to a cell complex [12]. Given a cell complex  $X$ , there is a face poset category whose objects are the cells of  $X$  and whose morphisms  $\tau \rightarrow \sigma$  correspond to the face relation  $\tau \trianglelefteq \sigma$  (that is,  $\tau \subset \bar{\sigma}$ ).

Following [23, 12], a **cellular cosheaf**  $\mathcal{F}$  on a cell complex  $X$  with values in category  $D$  is a contravariant functor from the associated face poset category to  $D$ . In other words,  $\mathcal{F}$  assigns to each cell  $\sigma$  of  $X$  an object  $\mathcal{F}(\sigma)$  in  $D$ , and to each face relation  $\tau \trianglelefteq \sigma$  a morphism  $\mathcal{F}(\tau \trianglelefteq \sigma) : \mathcal{F}(\sigma) \rightarrow \mathcal{F}(\tau)$  such that

- $\mathcal{F}(\tau \trianglelefteq \tau) : \mathcal{F}(\tau) \rightarrow \mathcal{F}(\tau)$  is the identity, and
- if  $\rho \trianglelefteq \tau \trianglelefteq \sigma$ , then  $\mathcal{F}(\rho \trianglelefteq \sigma) = \mathcal{F}(\rho \trianglelefteq \tau) \circ \mathcal{F}(\tau \trianglelefteq \sigma)$ .

A cellular cosheaf  $\mathcal{F}$  over a compact cell complex  $X$  has a well-defined homology associated to its chain complex

$$C_\bullet \mathcal{F} : \quad \cdots \xrightarrow{\partial_3} \bigoplus_{\dim \sigma=2} \mathcal{F}(\sigma) \xrightarrow{\partial_2} \bigoplus_{\dim \sigma=1} \mathcal{F}(\sigma) \xrightarrow{\partial_1} \bigoplus_{\dim \sigma=0} \mathcal{F}(\sigma) \xrightarrow{\partial_0} 0, \quad (2)$$

where  $C_n(X, \mathcal{F})$  is the direct sum of the data over the  $n$ -cells of  $X$ . The boundary map  $\partial_n : C_n(X, \mathcal{F}) \rightarrow C_{n-1}(X, \mathcal{F})$  is defined in the familiar manner as

$$\partial_n = \sum_{\tau \trianglelefteq \sigma} [\tau : \sigma] \mathcal{F}(\tau \trianglelefteq \sigma),$$

where  $[\tau : \sigma]$  is the incidence number. The  $n^{\text{th}}$  **cosheaf homology** of  $\mathcal{F}$  is the homology of this chain complex  $H_n(C_\bullet \mathcal{F}) = \ker \partial_n / \text{im } \partial_{n+1}$ .

Cosheaf homologies reflect global structures from locally encoded data. When comparing local data, cosheaf morphisms allow one to extract global changes from the local changes. Following [3], a **cosheaf morphism**  $\phi : \mathcal{F} \rightarrow \mathcal{F}'$  between cosheaves  $\mathcal{F}$  and  $\mathcal{F}'$  on  $X$  is a collection of morphisms  $\phi|_\sigma : \mathcal{F}(\sigma) \rightarrow \mathcal{F}'(\sigma)$  indexed by cells  $\sigma \in X$  such that the following diagram commutes

$$\begin{array}{ccc} \mathcal{F}(\sigma) & \xrightarrow{\phi|_\sigma} & \mathcal{F}'(\sigma) \\ \mathcal{F}(\tau \trianglelefteq \sigma) \downarrow & & \downarrow \mathcal{F}'(\tau \trianglelefteq \sigma) \\ \mathcal{F}(\tau) & \xrightarrow{\phi|_\tau} & \mathcal{F}'(\tau) \end{array}$$

for every face relation  $\tau \trianglelefteq \sigma$ .

Thus, a cosheaf morphism  $\phi : \mathcal{F} \rightarrow \mathcal{F}'$  is a natural transformation from the functor  $\mathcal{F}$  to  $\mathcal{F}'$ . As such, any cosheaf morphism  $\phi : \mathcal{F} \rightarrow \mathcal{F}'$  induces morphisms on homology:

$$H_n(\phi) : H_n(C_\bullet \mathcal{F}) \rightarrow H_n(C_\bullet \mathcal{F}').$$

In the special case of a cell complex  $X$  that is homeomorphic to a closed interval, a cellular cosheaf  $\mathcal{F}$  on  $X$  can be interpreted as a generalized persistence module, or **zigzag module** [6], of the form

$$V_1 \leftarrow V_2 \rightarrow V_3 \leftarrow \cdots \rightarrow V_m.$$

By Gabriel's Theorem [16], such a cosheaf  $\mathcal{F}$  can be decomposed into a direct sum of **indecomposable cosheaves**, of the form

$$\mathcal{J} : 0 \leftarrow \cdots \leftarrow 0 \leftrightarrow \mathbb{K} \leftrightarrow \cdots \leftrightarrow \mathbb{K} \leftrightarrow 0 \cdots \rightarrow 0,$$

as in Figure 1.

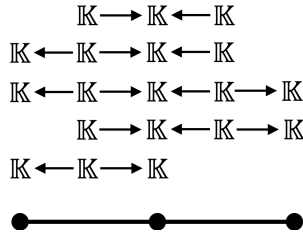


Figure 1: Direct sum of indecomposable cosheaves

Note that there are four types of indecomposable cosheaves possible: cosheaves  $\mathcal{J}_{[-]}$  whose left- and rightmost supports are 0-cells, cosheaves  $\mathcal{J}_{[-]}$  whose left and right most supports are 1-cells, cosheaves  $\mathcal{J}_{[-]}$  with leftmost support a 0-cell and rightmost support a 1-cell, and the reversed  $\mathcal{J}_{[-]}$ .

**Lemma 2.1** ([12]). *The indecomposable cosheaves  $\mathcal{J}$  satisfy*

$$H_0(C_\bullet \mathcal{J}_{[-]}) = \mathbb{K}, \quad H_1(C_\bullet \mathcal{J}_{[-]}) = \mathbb{K}, \quad H_i(C_\bullet \mathcal{J}_{[-]}) = H_i(C_\bullet \mathcal{J}_{[-]}) = 0.$$

Thus, if cosheaf  $\mathcal{F}$  can be decomposed as  $\mathcal{F} \cong \oplus \mathcal{J}$ , then  $H_i(C_\bullet \mathcal{F}) \cong \oplus H_i(C_\bullet \mathcal{J})$ . Thus,  $\dim H_0(C_\bullet \mathcal{F})$  counts the number of indecomposable cosheaves  $\mathcal{J}_{[-]}$  in the decomposition, whereas  $\dim H_1(C_\bullet \mathcal{F})$  counts the number of indecomposable cosheaves  $\mathcal{J}_{[-]}$  in the decomposition. Lemma 2.1 is the key to enriching the persistent homology barcodes in §5.

### 3 Distributed computation of homology

Our goal is to compute the persistence module

$$\mathbb{V} : H_\bullet(\mathcal{R}^1) \rightarrow H_\bullet(\mathcal{R}^2) \rightarrow \dots \rightarrow H_\bullet(\mathcal{R}^N)$$

in a distributed manner. To commence, we recall the local nature of homology computations, drawing on the classic results of Mayer-Vietoris, Leray [3], and Serre [20], in language of sheaf cohomology [13]. The following adapts the classic constructions from [13] to analyzing point cloud data.

For  $P$  a point cloud, denote by  $\mathcal{R}^\epsilon$  the Rips complex built on  $P$  for parameter  $\epsilon$ . Let  $\mathcal{V}$  be a finite cover of  $P$  and  $N_{\mathcal{V}}$  the resulting nerve complex.<sup>1</sup> To each simplex  $\sigma \in N_{\mathcal{V}}$  is then associated  $\mathcal{R}_\sigma^\epsilon$ , the Rips complex built on the subset of points of  $P$  indexed by  $\sigma$  at proximity parameter  $\epsilon$ .

We will refer to the collection  $\coprod_{\dim \sigma = n} \mathcal{R}_\sigma^\epsilon$  as the “Rips complexes over the  $n$ -simplices of  $N_{\mathcal{V}}$ ”, referring to the entire collection  $\coprod_{\sigma \in N_{\mathcal{V}}} \mathcal{R}_\sigma^\epsilon$  as the **Rips system** of the nerve.

Define a cosheaf  $\mathcal{F}_n^\epsilon$  on  $N_{\mathcal{V}}$  as the following. For each  $\sigma \in N_{\mathcal{V}}$ , let  $\mathcal{F}_n^\epsilon(\sigma) = H_n(\mathcal{R}_\sigma^\epsilon)$ . For  $\tau \trianglelefteq \sigma$ , let  $\mathcal{F}_n^\epsilon(\tau \trianglelefteq \sigma)$  be the map induced by inclusion  $\mathcal{R}_\tau^\epsilon \hookrightarrow \mathcal{R}_\sigma^\epsilon$ .

**Lemma 3.1.** *Let  $P$  be a point cloud with finite cover  $\mathcal{V}$  having one-dimensional nerve  $N_{\mathcal{V}}$ . There exists a constant  $\epsilon_*$  such that*

$$H_n(\mathcal{R}^\epsilon) \cong H_0(C_\bullet \mathcal{F}_n^\epsilon) \oplus H_1(C_\bullet \mathcal{F}_{n-1}^\epsilon) \quad (3)$$

for every  $0 < \epsilon < \epsilon_*$ .

*Proof.* Theorem 5.7 in [13] can be used to obtain the above isomorphism when the Rips system covers the full Rips complex: the technical work is in showing that this happens for sufficiently small  $\epsilon$ : see Appendix A for details.  $\square$

The following two examples illustrate the difference between  $H_0(C_\bullet \mathcal{F}_1^\epsilon)$  and  $H_1(C_\bullet \mathcal{F}_0^\epsilon)$ .

**Example 1.** Let  $P \subset \mathbb{R}^2$  be a point cloud covered by three sets  $V_1, V_2, V_3$  with nerve an interval as illustrated in Figure 2.

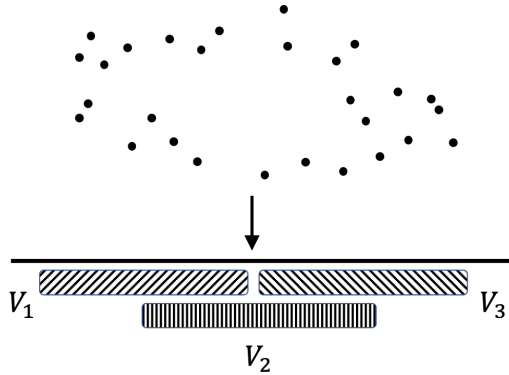


Figure 2: Point cloud  $P$  and covering by inverse image of projected intervals.

Consider  $H_1(\mathcal{R}^\epsilon)$  for some parameter  $\epsilon$ . The Rips complex  $\mathcal{R}^\epsilon$  and the Rips system over the nerve  $N_{\mathcal{V}}$  are illustrated in Figures 3a and 3b. Let  $v_1, v_2, v_3$  denote the vertices of  $N_{\mathcal{V}}$  that correspond to the cover sets  $V_1, V_2, V_3$ . Let  $e_{12}$  and  $e_{23}$  denote the edges of  $N_{\mathcal{V}}$  that correspond to  $V_1 \cap V_2$  and  $V_2 \cap V_3$ .

<sup>1</sup>The simplicial complex that indexes intersections of cover elements.

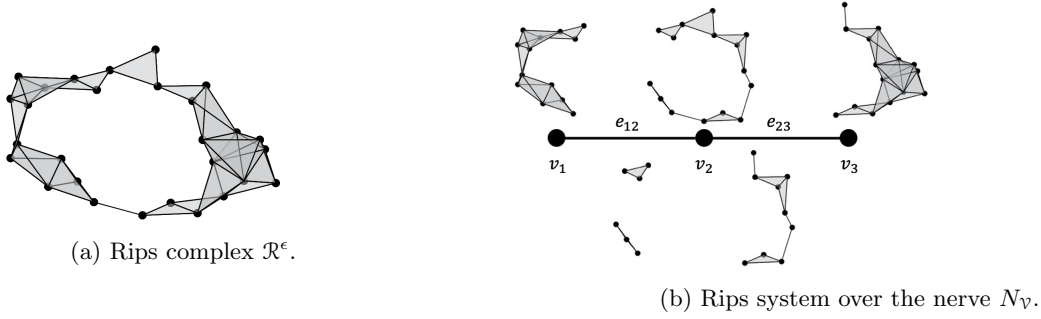


Figure 3: Rips complex and the associated Rips system.

The two relevant cosheaves,  $\mathcal{F}_0^\epsilon$  and  $\mathcal{F}_1^\epsilon$ , are illustrated in Figures 4a and 4b. The maps  $\mathcal{F}_0^\epsilon(v_1 \trianglelefteq e_{12})$  and  $\mathcal{F}_0^\epsilon(v_2 \trianglelefteq e_{12})$  are represented by the matrix  $\begin{bmatrix} 1 & 1 \end{bmatrix}$ . All other maps are identity maps.

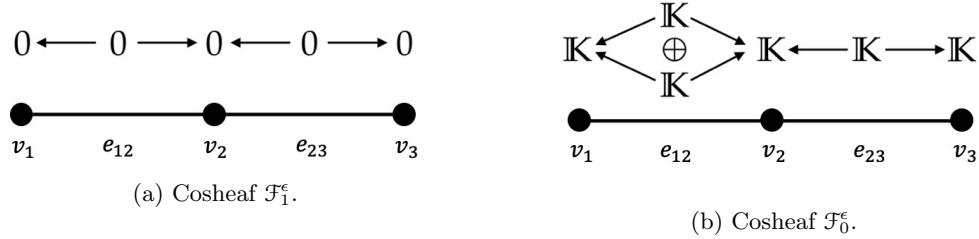


Figure 4: The two relevant cosheaves for computing  $H_1(\mathcal{R}^\epsilon)$ .

One can verify that Equation (3) holds by computing

$$H_0(C_\bullet \mathcal{F}_1^\epsilon) = 0, \quad H_1(C_\bullet \mathcal{F}_0^\epsilon) = \mathbb{K}. \quad (4)$$

**Example 2.**

Consider now a parameter  $\epsilon'$  that is larger than the parameter from Example 1. The Rips complex  $\mathcal{R}^{\epsilon'}$  and the Rips system are illustrated in Figure 5a and 5b.

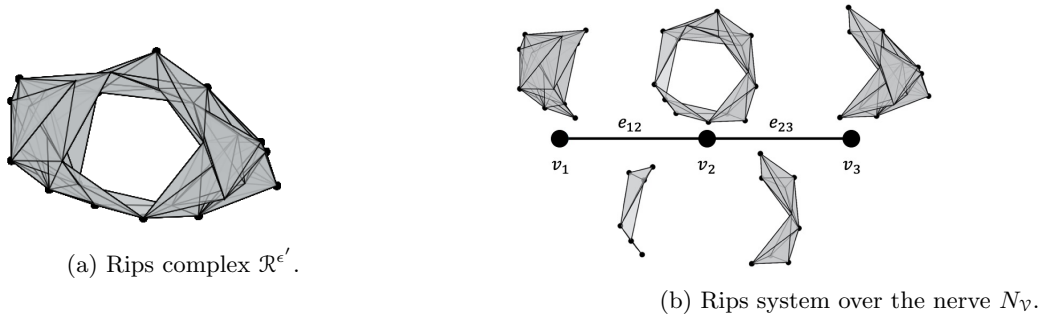


Figure 5: Rips complex and the associated Rips system.

One can compute cosheaves  $\mathcal{F}_1^{\epsilon'}$ ,  $\mathcal{F}_0^{\epsilon'}$ , and compute the following cosheaf homologies

$$H_0(C_\bullet \mathcal{F}_1^{\epsilon'}) = \mathbb{K}, \quad H_1(C_\bullet \mathcal{F}_0^{\epsilon'}) = 0. \quad (5)$$

To compare the cosheaf homologies from Equations (4) and (5) for the two parameters  $\epsilon < \epsilon'$ , note that both  $H_1(\mathcal{R}^\epsilon) = \mathbb{K}$  and  $H_1(\mathcal{R}^{\epsilon'}) = \mathbb{K}$ . In Example 1, the homology class of  $H_1(\mathcal{R}^\epsilon)$  is detected by  $H_1(C_\bullet \mathcal{F}_0^\epsilon)$ ,

while in Example 2, the homology class of  $H_1(\mathcal{R}^{\epsilon'})$  is detected by  $H_0(C_\bullet \mathcal{F}_1^{\epsilon'})$ . The difference can be explained by comparing the Rips systems from Figures 3b and 5b. In Figure 5b, the Rips complex  $\mathcal{R}_{v_2}^{\epsilon'}$  contains a non-trivial 1-cycle, while in Figure 3b, there is no such 1-cycle contained in any of the complexes  $\mathcal{R}_\sigma^\epsilon$  for  $\sigma \in N_\mathcal{V}$ .

In general,  $H_0(C_\bullet \mathcal{F}_n^\epsilon)$  reads  $n$ -cycles that exist in  $\mathcal{R}_\sigma^\epsilon$  for some  $\sigma \in N_\mathcal{V}$ . On the other hand,  $H_1(C_\bullet \mathcal{F}_{n-1}^\epsilon)$  reads  $n$ -cycles of  $\mathcal{R}^\epsilon$  that are not cycles of  $H_n(\mathcal{R}_\sigma^\epsilon)$  for any  $\sigma \in N_\mathcal{V}$ .

## 4 Distributed computation of persistent homology

We restate the main question using the terminology introduced so far.

Given a point cloud  $P$ , one can build Rips complexes for increasing parameter values  $(\epsilon_i)_{i=1}^N$ , resulting in the following sequence of Rips complexes and inclusion maps

$$\mathcal{R}^1 \xrightarrow{\iota^1} \mathcal{R}^2 \xrightarrow{\iota^2} \dots \xrightarrow{\iota^{N-1}} \mathcal{R}^N.$$

By applying the homology functor of dimension  $n$ , one obtains the persistence module

$$\mathbb{V} : H_n(\mathcal{R}^1) \xrightarrow{\iota_*^1} H_n(\mathcal{R}^2) \xrightarrow{\iota_*^2} \dots \xrightarrow{\iota_*^{N-1}} H_n(\mathcal{R}^N). \quad (6)$$

Assuming a covering  $\mathcal{V}$  of the point cloud  $P$  that has 1-d nerve (cf. Lemma 3.1), we have the following isomorphisms

$$H_n(\mathcal{R}^i) \cong H_0(C_\bullet \mathcal{F}_n^i) \oplus H_1(C_\bullet \mathcal{F}_{n-1}^i) \quad (7)$$

for every  $i$  and  $n$ . Our main question is stated as the following.

*Can we construct a persistence module*

$$\mathbb{V}_\Psi : H_0(C_\bullet \mathcal{F}_n^1) \oplus H_1(C_\bullet \mathcal{F}_{n-1}^1) \xrightarrow{\Psi^1} \dots \xrightarrow{\Psi^{N-1}} H_0(C_\bullet \mathcal{F}_n^N) \oplus H_1(C_\bullet \mathcal{F}_{n-1}^N)$$

*isomorphic to the persistence module  $\mathbb{V}$  from Equation (6)?*

In §4.1, we show that the most naturally induced morphisms of cosheaf homologies are not enough to construct the desired persistence module  $\mathbb{V}_\Psi$ . In §4.2, we use spectral sequences to construct a map  $\psi^i : H_1(C_\bullet \mathcal{F}_{n-1}^i) \rightarrow H_0(C_\bullet \mathcal{F}_n^{i+1})$  that can be used to define the persistence module  $\mathbb{V}_\Psi$ . As it will be illustrated in §4.2, there are multiple choices involved in defining the map  $\psi^i$ , and the construction of  $\mathbb{V}_\Psi$  requires that the maps  $\psi^i$  be defined consistently across parameters  $(\epsilon_i)_{i=1}^N$ . In §4.3, we provide an algorithm to construct the maps  $\psi^i$  consistently across parameters  $(\epsilon_i)_{i=1}^N$ . Furthermore, we construct the persistence module

$$\mathbb{V}_\Psi : H_0(C_\bullet \mathcal{F}_n^1) \oplus H_1(C_\bullet \mathcal{F}_{n-1}^1) \xrightarrow{\Psi^1} \dots \xrightarrow{\Psi^{N-1}} H_0(C_\bullet \mathcal{F}_n^N) \oplus H_1(C_\bullet \mathcal{F}_{n-1}^N).$$

In §4.4, we show that the persistence module  $\mathbb{V}_\Psi$  is isomorphic to the persistence module  $\mathbb{V}$  from Equation (6).

### 4.1 Cosheaf morphisms

Given a pair of parameters  $\epsilon_i < \epsilon_{i+1}$  and  $\sigma \in N_\mathcal{V}$ , there exist maps  $\mathcal{F}_n^i(\sigma) \rightarrow \mathcal{F}_n^{i+1}(\sigma)$  induced by the inclusion  $\mathcal{R}_\sigma^i \hookrightarrow \mathcal{R}_\sigma^{i+1}$ . The collection of such maps is the cosheaf morphism  $\phi_n^i : \mathcal{F}_n^i \rightarrow \mathcal{F}_n^{i+1}$ . In particular, the cosheaf morphisms  $\phi_n^i$  and  $\phi_{n-1}^i$  induce the following morphisms on homology

$$\begin{aligned} H_0(\phi_n^i) &: H_0(C_\bullet \mathcal{F}_n^i) \rightarrow H_0(C_\bullet \mathcal{F}_n^{i+1}), \\ H_1(\phi_{n-1}^i) &: H_1(C_\bullet \mathcal{F}_{n-1}^i) \rightarrow H_1(C_\bullet \mathcal{F}_{n-1}^{i+1}). \end{aligned} \quad (8)$$

The maps  $H_0(\phi_n^i)$  and  $H_1(\phi_{n-1}^i)$  are insufficient to construct a persistence module isomorphic to  $\mathbb{V}$  from Equation (6), as we now demonstrate. Using the maps  $H_0(\phi_n^i)$  and  $H_1(\phi_{n-1}^i)$ , one defines

$$\omega^i : H_0(C_\bullet \mathcal{F}_n^i) \oplus H_1(C_\bullet \mathcal{F}_{n-1}^i) \rightarrow H_0(C_\bullet \mathcal{F}_n^{i+1}) \oplus H_1(C_\bullet \mathcal{F}_{n-1}^{i+1})$$

by  $\omega^i(u, v) = (H_0(\phi_n^i)(u), H_1(\phi_{n-1}^i)(v))$ . The obvious attempt to reconstruct  $\mathbb{V}$  is the persistence module

$$\mathbb{O} : H_0(C_\bullet \mathcal{F}_n^1) \oplus H_1(C_\bullet \mathcal{F}_{n-1}^1) \xrightarrow{\omega^1} \dots \xrightarrow{\omega^{N-1}} H_0(C_\bullet \mathcal{F}_n^N) \oplus H_1(C_\bullet \mathcal{F}_{n-1}^N).$$

**Claim:**  $\mathbb{O}$  cannot be isomorphic to  $\mathbb{V}$  from Equation (6).

*Proof:* The putative isomorphisms  $\Phi^i : H_0(C_\bullet \mathcal{F}_n^i) \oplus H_1(C_\bullet \mathcal{F}_{n-1}^i) \rightarrow H_n(\mathcal{R}^i)$  would yield commutative diagrams

$$\begin{array}{ccc} H_0(C_\bullet \mathcal{F}_n^i) \oplus H_1(C_\bullet \mathcal{F}_{n-1}^i) & \xrightarrow{\Phi^i} & H_n(\mathcal{R}^i) \\ \downarrow \omega^i & & \downarrow \iota_*^i \\ H_0(C_\bullet \mathcal{F}_n^{i+1}) \oplus H_1(C_\bullet \mathcal{F}_{n-1}^{i+1}) & \xrightarrow{\Phi^{i+1}} & H_n(\mathcal{R}^{i+1}) \end{array} \quad (9)$$

However, Examples 1 and 2 illustrate a situation where it is impossible to find isomorphisms  $\Phi^i$  and  $\Phi^{i+1}$  that make Diagram 9 commute. Assume that there exists an isomorphism  $\Phi^i$ , and let  $(0, s) \in H_0(C_\bullet \mathcal{F}_1^i) \oplus H_1(C_\bullet \mathcal{F}_0^i)$  be the element such that  $\Phi^i(0, s)$  represents the non-trivial 1-cycle in Figure 3a. Then,  $\iota_*^i \circ \Phi^i(0, s)$  must be the non-trivial 1-cycle in Figure 5a. On the other hand, with our current construction of  $\omega^i$ , we have  $\omega^i(0, s) = 0$ , and hence,  $\Phi^{i+1} \circ \omega^i(0, s) = 0$  for any isomorphism  $\Phi^{i+1}$ . Thus, there are no isomorphisms  $\Phi^i$  and  $\Phi^{i+1}$  that make Diagram 9 commute.

The core reason why Diagram 9 fails to commute is that as we increase the parameter from  $\epsilon_i$  to  $\epsilon_{i+1}$ , a cycle in  $H_1(C_\bullet \mathcal{F}_{n-1}^i)$  can become homologous to a cycle in  $H_0(C_\bullet \mathcal{F}_n^{i+1})$ . The current construction of  $\omega^i$  fails to take such subtlety into account. This motivates our technique: we construct a map from  $H_1(C_\bullet \mathcal{F}_{n-1}^i)$  to  $H_0(C_\bullet \mathcal{F}_n^{i+1})$ .

## 4.2 Connecting morphism via spectral sequences

We seek a map  $\psi^i : H_1(C_\bullet \mathcal{F}_{n-1}^i) \rightarrow H_0(C_\bullet \mathcal{F}_n^{i+1})$  for the construction of the persistence module  $\mathbb{V}_\Psi$ . The plan to build  $\psi^i$  is as an extension of a map  $\delta^i : \ker H_1(\phi_{n-1}^i) \rightarrow H_0(C_\bullet \mathcal{F}_n^{i+1})$  using a spectral sequence type argument.

**Theorem 2.** *Let  $P$  be a point cloud with finite cover  $\mathcal{V}$  having one-dimensional nerve  $N_\mathcal{V}$ . There exists a morphism  $\delta^i : \ker H_1(\phi_{n-1}^i) \rightarrow H_0(C_\bullet \mathcal{F}_n^{i+1})$  induced by cosheaf morphisms  $\phi_n^i$  and  $\phi_{n-1}^i$ .*

*Proof.* Consider the following commutative diagram. Let  $\iota_n^i : \bigoplus_{v \in N_\mathcal{V}} C_n(\mathcal{R}_v^i) \rightarrow \bigoplus_{v \in N_\mathcal{V}} C_n(\mathcal{R}_v^{i+1})$  denote the collection of inclusion maps of the Rips complexes over the vertices of  $N_\mathcal{V}$ , and let  $\kappa_n^i : \bigoplus_{e \in N_\mathcal{V}} C_n(\mathcal{R}_e^i) \rightarrow$

$\bigoplus_{e \in N_\mathcal{V}} C_n(\mathcal{R}_e^{i+1})$  denote the collection of inclusion maps of the Rips complexes over the edges of  $N_\mathcal{V}$ . Let  $e_n^i : \bigoplus_{e \in N_\mathcal{V}} C_n(\mathcal{R}_e^i) \rightarrow \bigoplus_{v \in N_\mathcal{V}} C_n(\mathcal{R}_v^i)$  denote the collection of inclusion maps. The front and back faces of the cube in Diagram 10 are the 0<sup>th</sup> pages of the spectral sequence in the proof of Lemma 3.1 for parameters  $\epsilon_i$  and  $\epsilon_{i+1}$  respectively.



$$\begin{array}{ccccc}
& & \downarrow & & \downarrow \\
& & \bigoplus_{v \in N_{\mathcal{V}}} C_n(\mathcal{R}_v^{i+1}) & \xleftarrow{e_n^{i+1}} & \bigoplus_{e \in N_{\mathcal{V}}} C_n(\mathcal{R}_e^{i+1}) \\
& \nearrow \iota_n^i & \downarrow \partial & & \nearrow \kappa_n^i \\
\bigoplus_{v \in N_{\mathcal{V}}} C_n(\mathcal{R}_v^i) & \xleftarrow{e_n^i} & \bigoplus_{e \in N_{\mathcal{V}}} C_n(\mathcal{R}_e^i) & & \\
\downarrow \partial & & \downarrow \partial & & \downarrow \partial \\
& \nearrow \iota_{n-1}^i & \bigoplus_{v \in N_{\mathcal{V}}} C_{n-1}(\mathcal{R}_v^{i+1}) & \xleftarrow{e_{n-1}^{i+1}} & \bigoplus_{e \in N_{\mathcal{V}}} C_{n-1}(\mathcal{R}_e^{i+1}) \\
& & \downarrow & & \downarrow \\
\bigoplus_{v \in N_{\mathcal{V}}} C_{n-1}(\mathcal{R}_v^i) & \xleftarrow{e_{n-1}^i} & \bigoplus_{e \in N_{\mathcal{V}}} C_{n-1}(\mathcal{R}_e^i) & & \\
& & \downarrow & & \downarrow
\end{array} \tag{10}$$

Computing the homology with respect to the boundary maps  $\partial$  yields Diagram 11, in which the maps  $\partial_n$  are boundary maps of the chain complexes  $C_{\bullet} \mathcal{F}_n^i$  of the respective cosheaves.

$$\begin{array}{ccccc}
& & \vdots & & \vdots \\
& & \bigoplus_{v \in N_{\mathcal{V}}} H_n(\mathcal{R}_v^{i+1}) & \xleftarrow{\partial_n^{i+1}} & \bigoplus_{e \in N_{\mathcal{V}}} H_n(\mathcal{R}_e^{i+1}) \\
& \nearrow (\phi_n^i)_v & \downarrow \vdots & & \nearrow (\phi_n^i)_e \\
\bigoplus_{v \in N_{\mathcal{V}}} H_n(\mathcal{R}_v^i) & \xleftarrow{\partial_n^i} & \bigoplus_{e \in N_{\mathcal{V}}} H_n(\mathcal{R}_e^i) & & \\
& \nearrow (\phi_{n-1}^i)_v & \bigoplus_{v \in N_{\mathcal{V}}} H_{n-1}(\mathcal{R}_v^{i+1}) & \xleftarrow{\partial_{n-1}^{i+1}} & \bigoplus_{e \in N_{\mathcal{V}}} H_{n-1}(\mathcal{R}_e^{i+1}) \\
& & \downarrow \vdots & & \downarrow \vdots \\
\bigoplus_{v \in N_{\mathcal{V}}} H_{n-1}(\mathcal{R}_v^i) & \xleftarrow{\partial_{n-1}^i} & \bigoplus_{e \in N_{\mathcal{V}}} H_{n-1}(\mathcal{R}_e^i) & & \\
& & \vdots & & \vdots
\end{array} \tag{11}$$

Computing the homology with respect to these maps  $\partial_n^i$  yields Diagram 12 of cosheaf homologies.



We will inductively fix a basis  $\mathcal{B}_{\ker}^i$  of  $\ker H_1(\phi_{n-1}^i)$  and extend it to a basis  $\mathcal{B}^i$  of  $H_1(C_\bullet \mathcal{F}_{n-1}^i)$ . We will define a set map  $\Gamma^i : \mathcal{B}^i \rightarrow \bigoplus_{v \in N_\vee} C_n(\mathcal{R}_v^i)$  that consistently chooses  $\beta^i$ 's for each element of  $\mathcal{B}^i$ . We will then define  $\delta^i$  on the basis  $\{\langle b \rangle\} \in \mathcal{B}_{\ker}^i$  by

$$\delta^i \{\langle b \rangle\} = \{ \langle -e_n^{i+1} \alpha^{i+1} + \iota_n^i \circ \Gamma^i \{\langle b \rangle\} \rangle \} \quad (17)$$

and extend the map linearly to  $\ker H_1(\phi_{n-1}^i)$ . The map  $\delta^i$  can then be extended to  $\psi^i : H_1(C_\bullet \mathcal{F}_{n-1}^i) \rightarrow H_0(C_\bullet \mathcal{F}_n^{i+1})$  as Equation (16).

Note that the construction of  $\delta^i : \ker H_1(\phi_{n-1}^i) \rightarrow H_0(C_\bullet \mathcal{F}_n^{i+1})$  requires a choice of  $\beta^i$  for only the basis elements  $\{\langle b \rangle\} \in \mathcal{B}_{\ker}^i$ . However, we choose such  $\beta^i$  for every basis element  $\{\langle b \rangle\} \in \mathcal{B}^i$  because such choice can affect the construction of  $\delta^j$  for  $j > i$ .

### Base case

Recalling Diagram 10, note that

$$H_1(C_\bullet \mathcal{F}_{n-1}^1) = A^1 \oplus \ker H_1(\phi_{n-1}^1)$$

for some subspace  $A^1$ . Let  $\mathcal{B}_A^1$  be a basis of  $A^1$ , and let  $\mathcal{B}_{\ker}^1$  be a basis of  $\ker H_1(\phi_{n-1}^1)$ . Then,

$$\mathcal{B}^1 = \mathcal{B}_A^1 \cup \mathcal{B}_{\ker}^1 \quad (18)$$

is a basis of  $H_1(C_\bullet \mathcal{F}_{n-1}^1)$ . For each basis element  $\{\langle b \rangle\} \in \mathcal{B}^1$ , fix a coset representative  $b^*$  of  $\langle b \rangle$ .

Define a set map  $\Gamma^1 : \mathcal{B}^1 \rightarrow \bigoplus_{v \in N_\vee} C_n(\mathcal{R}_v^1)$  as following. For each  $\{\langle b \rangle\} \in \mathcal{B}^1$ , let

$$\Gamma^1 \{\langle b \rangle\} = \beta^1, \quad (19)$$

where  $\beta^1 \in \bigoplus_{v \in N_\vee} C_n(\mathcal{R}_v^1)$  is any element satisfying  $\partial \beta^1 = e_{n-1}^1 b^*$ .

Define  $\delta^1$  on  $\{\langle b \rangle\} \in \mathcal{B}_{\ker}^1$  by

$$\delta^1 \{\langle b \rangle\} = \{ \langle -e_n^2 \alpha^2 + \iota_n^1 \circ \Gamma^1 \{\langle b \rangle\} \rangle \},$$

where  $\alpha^2 \in \bigoplus_{e \in N_\vee} C_n(\mathcal{R}_e^{i+1})$  is any element satisfying Equation (13). Extend  $\delta^1$  linearly to  $\ker H_1(\phi_{n-1}^1)$ . The map  $\delta^1$  can be used to define the map  $\psi^1 : H_1(C_\bullet \mathcal{F}_{n-1}^1) \rightarrow H_0(C_\bullet \mathcal{F}_n^2)$  via Equation (16).

### Inductive step

- **Inductive assumption.**

Note that

$$H_1(C_\bullet \mathcal{F}_{n-1}^{i-1}) = A^{i-1} \oplus \ker H_1(\phi_{n-1}^{i-1})$$

for some subspace  $A^{i-1}$ .

- Assume that there exists a basis  $\mathcal{B}^{i-1}$  of  $H_1(C_\bullet \mathcal{F}_{n-1}^{i-1})$  that has the form

$$\mathcal{B}^{i-1} = \mathcal{B}_A^{i-1} \cup \mathcal{B}_{\ker}^{i-1},$$

where  $\mathcal{B}_A^{i-1}$  is a basis of  $A^{i-1}$  and  $\mathcal{B}_{\ker}^{i-1}$  is a basis of  $\ker H_1(\phi_{n-1}^{i-1})$ .

- Assume that for each basis element  $\{\langle b \rangle\} \in \mathcal{B}^{i-1}$ , a coset representative  $b^*$  of  $\langle b \rangle$  has been fixed.
- Assume that there exists a set map  $\Gamma^{i-1} : \mathcal{B}^{i-1} \rightarrow \bigoplus_{v \in N_\vee} C_n(\mathcal{R}_v^{i-1})$  such that

$$\partial \Gamma^{i-1} \{\langle b \rangle\} = e_{n-1}^{i-1} b^* \quad (20)$$

for every  $\{\langle b \rangle\} \in \mathcal{B}^{i-1}$ .

- **Step 1. Fix a basis  $\mathcal{C}^i$  of  $H_1(C_\bullet \mathcal{F}_{n-1}^i)$  that is compatible with  $\mathcal{B}^{i-1}$ .**

By assumption, the basis  $\mathcal{B}^{i-1}$  of  $H_1(C_\bullet \mathcal{F}_{n-1}^{i-1})$  has the form  $\mathcal{B}^{i-1} = \mathcal{B}_A^{i-1} \cup \mathcal{B}_{\ker}^{i-1}$ . Without loss of generality, assume that

$$\mathcal{B}_A^{i-1} = \{ \{ \langle b_1 \rangle \}, \dots, \{ \langle b_t \rangle \} \}.$$

One can show that  $\{ \langle \kappa_{n-1}^{i-1} b_1 \rangle \}, \dots, \{ \langle \kappa_{n-1}^{i-1} b_t \rangle \}$  are linearly independent in  $H_1(C_\bullet \mathcal{F}_{n-1}^i)$  (Appendix C). Let

$$\mathcal{C}_{\text{im}}^i = \{ \{ \langle \kappa_{n-1}^{i-1} b_1 \rangle \}, \dots, \{ \langle \kappa_{n-1}^{i-1} b_t \rangle \} \}.$$

Extend  $\mathcal{C}_{\text{im}}^i$  to a basis  $\mathcal{C}^i$  of  $H_1(C_\bullet \mathcal{F}_{n-1}^i)$ . Let  $\mathcal{C}_D^i$  denote the basis vectors of  $\mathcal{C}^i$  that are not in  $\mathcal{C}_{\text{im}}^i$ , i.e.,

$$\mathcal{C}^i = \mathcal{C}_{\text{im}}^i \cup \mathcal{C}_D^i. \quad (21)$$

If  $\{ \langle c \rangle \} \in \mathcal{C}_{\text{im}}^i$  such that  $\{ \langle c \rangle \} = \{ \langle \kappa_{n-1}^{i-1} b \rangle \}$ , then let  $c^* = \kappa_{n-1}^{i-1} b^*$  be the coset representative of  $\langle c \rangle$ . If  $\{ \langle c \rangle \} \in \mathcal{C}_D^i$ , fix any coset representative  $c^*$  of  $\langle c \rangle$ .

- **Step 2. Define a set map  $\Gamma_{\mathcal{C}}^i : \mathcal{C}^i \rightarrow \bigoplus_{v \in N_V} C_n(\mathcal{R}_v^i)$ .**

We will define a set map  $\Gamma_{\mathcal{C}}^i : \mathcal{C}^i \rightarrow \bigoplus_{v \in N_V} C_n(\mathcal{R}_v^i)$  such that

$$\partial \Gamma_{\mathcal{C}}^i \{ \langle c \rangle \} = e_{n-1}^i c^* \quad (22)$$

for every  $\{ \langle c \rangle \} \in \mathcal{C}^i$ . Define

$$\Gamma_{\mathcal{C}}^i \{ \langle c \rangle \} = \begin{cases} \iota_n^{i-1} \Gamma^{i-1} \{ \langle b \rangle \} & \text{if } \{ \langle c \rangle \} = \{ \langle \kappa_{n-1}^{i-1} b \rangle \} \in \mathcal{C}_{\text{im}}^i \\ \text{any } \beta^i \text{ satisfying } \partial \beta^i = e_{n-1}^i c^* & \text{if } \{ \langle c \rangle \} \in \mathcal{C}_D^i \end{cases} \quad (23)$$

- **Step 3. Fix a new basis  $\mathcal{B}^i$  of  $H_1(C_\bullet \mathcal{F}_{n-1}^i)$**

Note that

$$H_1(C_\bullet \mathcal{F}_{n-1}^i) = A^i \oplus \ker H_1(\phi_{n-1}^i)$$

for some subspace  $A^i$ . Let  $\mathcal{B}_A^i$  be a basis of  $A^i$ , and let  $\mathcal{B}_{\ker}^i$  be a basis of  $\ker H_1(\phi_{n-1}^i)$ . Then,

$$\mathcal{B}^i = \mathcal{B}_A^i \cup \mathcal{B}_{\ker}^i \quad (24)$$

is a basis of  $H_1(C_\bullet \mathcal{F}_{n-1}^i)$ .

The coset representative  $b^*$  for each basis vector  $\{ \langle b \rangle \} \in \mathcal{B}^i$  follows naturally from the coset representatives of  $\mathcal{C}^i$ . In other words, if  $\mathcal{C}^i = \{ \{ \langle c_1 \rangle \}, \dots, \{ \langle c_l \rangle \} \}$  and  $\{ \langle b \rangle \} \in \mathcal{B}^i$  is written as

$$\{ \langle b \rangle \} = d_1 \{ \langle c_1 \rangle \} + \dots + d_l \{ \langle c_l \rangle \},$$

then let

$$b^* = d_1 c_1^* + \dots + d_l c_l^* \quad (25)$$

be the coset representative of  $\langle b \rangle$ .

- **Step 4. Define the set map  $\Gamma_i : \mathcal{B}^i \rightarrow \bigoplus_{v \in N_V} C_n(\mathcal{R}_v^i)$ .**

Given  $\{ \langle b \rangle \} \in \mathcal{B}^i$ , if  $\{ \langle b \rangle \}$  is written as

$$\{ \langle b \rangle \} = d_1 \{ \langle c_1 \rangle \} + \dots + d_l \{ \langle c_l \rangle \}$$

where  $\{ \langle c_1 \rangle \}, \dots, \{ \langle c_l \rangle \}$  are basis  $\mathcal{C}^i$ , then define  $\Gamma_i \{ \langle b \rangle \}$  as

$$\Gamma_i \{ \langle b \rangle \} = d_1 \Gamma_{\mathcal{C}}^i \{ \langle c_1 \rangle \} + \dots + d_l \Gamma_{\mathcal{C}}^i \{ \langle c_l \rangle \}.$$

One can check that

$$\partial \Gamma_i \{ \langle b \rangle \} = e_{n-1}^i b^*. \quad (26)$$

- **Step 5. Define the maps  $\delta^i$  and  $\psi^i$ .**

For each  $\{b\} \in \mathcal{B}_{\ker}^i$ , let

$$\delta^i\{b\} = \{ \langle -e_n^{i+1}\alpha^{i+1} + \kappa_n^i \circ \Gamma^i\{b\} \rangle \},$$

where  $\alpha^{i+1} \in \bigoplus_{e \in N_{\mathcal{V}}} C_n(\mathcal{R}_e^{i+1})$  is any element satisfying  $\partial\alpha^{i+1} = \kappa_{n-1}^i b^*$ .

Extend  $\delta^i$  linearly to  $\delta^i : \ker H_1(\phi_{n-1}^i) \rightarrow H_0(C_{\bullet}\mathcal{F}_n^{i+1})$ . One can then extend  $\delta^i$  to map  $\psi^i : H_1(C_{\bullet}\mathcal{F}_{n-1}^i) \rightarrow H_0(C_{\bullet}\mathcal{F}_n^{i+1})$  via Equation (16).

Going through Steps 1-5 defines  $\delta^i$  and  $\psi^i$  inductively for every  $i$ . We can then define the map

$$\Psi^i : H_0(C_{\bullet}\mathcal{F}_n^i) \oplus H_1(C_{\bullet}\mathcal{F}_{n-1}^i) \rightarrow H_0(C_{\bullet}\mathcal{F}_n^{i+1}) \oplus H_1(C_{\bullet}\mathcal{F}_{n-1}^{i+1})$$

by

$$\Psi^i(\{x\}, \{y\}) = (H_0(\phi_n^i)\{x\} + (-1)^{n+1}\psi^i\{y\}, H_1(\phi_{n-1}^i)\{y\}), \quad (27)$$

where  $H_0(\phi_n^i)$  and  $H_1(\phi_{n-1}^i)$  are the maps defined in Equation (8). We can then define the persistence module  $\mathbb{V}_{\Psi}$

$$\mathbb{V}_{\Psi} : H_0(C_{\bullet}\mathcal{F}_n^1) \oplus H_1(C_{\bullet}\mathcal{F}_{n-1}^1) \xrightarrow{\Psi^1} \dots \xrightarrow{\Psi^{N-1}} H_0(C_{\bullet}\mathcal{F}_n^N) \oplus H_1(C_{\bullet}\mathcal{F}_{n-1}^N). \quad (28)$$

#### 4.4 Isomorphism of persistence modules

We show that the persistence module  $\mathbb{V}_{\Psi}$  constructed in Equation (28) is isomorphic to the persistence module  $\mathbb{V}$  from Equation (6). In order to do so, we will show that both  $\mathbb{V}_{\Psi}$  and  $\mathbb{V}$  are isomorphic to the following persistence module

$$\mathbb{V}_{\text{Tot}} : H_n(\text{Tot}^1) \xrightarrow{\iota_{\text{Tot}}^1} H_n(\text{Tot}^2) \xrightarrow{\iota_{\text{Tot}}^2} \dots \xrightarrow{\iota_{\text{Tot}}^{N-1}} H_n(\text{Tot}^N),$$

where each  $H_n(\text{Tot}^i)$  is the homology of the double complex from Diagram 29 for parameter  $\epsilon_i$ , and  $\iota_{\text{Tot}}^i$  is the morphism induced by maps of double complexes.

$$\begin{array}{ccc} \vdots & & \vdots \\ \downarrow \partial & & \downarrow \partial \\ \bigoplus_{v \in N_{\mathcal{V}}} C_2(\mathcal{R}_v^i) & \xleftarrow{e_2^i} & \bigoplus_{e \in N_{\mathcal{V}}} C_2(\mathcal{R}_e^i) \\ \downarrow \partial & & \downarrow \partial \\ \bigoplus_{v \in N_{\mathcal{V}}} C_1(\mathcal{R}_v^i) & \xleftarrow{e_1^i} & \bigoplus_{e \in N_{\mathcal{V}}} C_1(\mathcal{R}_e^i) \\ \downarrow \partial & & \downarrow \partial \\ \bigoplus_{v \in N_{\mathcal{V}}} C_0(\mathcal{R}_v^i) & \xleftarrow{e_0^i} & \bigoplus_{e \in N_{\mathcal{V}}} C_0(\mathcal{R}_e^i) \end{array} \quad (29)$$

Let  $H_n(\text{Tot}^i)$  denote the homology of the double complex. Note that a coset of  $H_n(\text{Tot}^i)$  is represented by  $[x, y]$ , where  $x \in \bigoplus_{v \in N_{\mathcal{V}}} C_n(\mathcal{R}_v^i)$ ,  $y \in \bigoplus_{e \in N_{\mathcal{V}}} C_{n-1}(\mathcal{R}_e^i)$ ,  $\partial y = 0$  and  $\partial x = (-1)^{n-1} e_{n-1}^i y$ . A coset  $[x, y]$  is trivial in  $H_n(\text{Tot}^i)$  if there exist  $p_{n+1} \in \bigoplus_{v \in N_{\mathcal{V}}} C_{n+1}(\mathcal{R}_v^i)$  and  $q_n \in \bigoplus_{e \in N_{\mathcal{V}}} C_n(\mathcal{R}_e^i)$  such that  $\partial q_n = y$  and  $\partial p_{n+1} + (-1)^{n+1} e_n^i(q_n) = x$ .

Given increasing parameter values  $(\epsilon_i)_{i=1}^N$ , one can construct a double complex for each parameter  $\epsilon_i$ . There exists an inclusion map from double complex associated with parameter  $\epsilon_i$  to that of parameter  $\epsilon_{i+1}$ , as illustrated in Diagram 10. The vertical maps  $\iota_n^i$  and  $\kappa_n^i$  constitute the inclusion maps of double complexes.

Such inclusion of double complexes induces a morphism  $\iota_{\text{Tot}}^i : H_n(\text{Tot}^i) \rightarrow H_n(\text{Tot}^{i+1})$  which can be written explicitly as

$$\iota_{\text{Tot}}^i([x, y]) = [\iota_n^i(x), \kappa_{n-1}^i(y)]. \quad (30)$$

**Lemma 4.1.** *The persistence modules  $\mathbb{V}_{\text{Tot}}$  and  $\mathbb{V}$  are isomorphic.*

*Proof.* We will define isomorphisms  $\Phi_{\text{Tot}}^i : H_n(\text{Tot}^i) \rightarrow H_n(\mathcal{R}^i)$  such that the following diagram commutes.

$$\begin{array}{ccccccc} H_n(\text{Tot}^1) & \xrightarrow{\iota_{\text{Tot}}^1} & H_n(\text{Tot}^2) & \xrightarrow{\iota_{\text{Tot}}^2} & \dots & \xrightarrow{\iota_{\text{Tot}}^{N-1}} & H_n(\text{Tot}^N) \\ \downarrow \Phi_{\text{Tot}}^1 & & \downarrow \Phi_{\text{Tot}}^2 & & & & \downarrow \Phi_{\text{Tot}}^N \\ H_n(\mathcal{R}^1) & \xrightarrow{\iota_*^1} & H_n(\mathcal{R}^2) & \xrightarrow{\iota_*^2} & \dots & \xrightarrow{\iota_*^{N-1}} & H_n(\mathcal{R}^N) \end{array} \quad (31)$$

For each parameter  $\epsilon_i$ , let  $j_n^i : \bigoplus_{v \in N_{\mathbb{V}}} C_n(\mathcal{R}_v^i) \rightarrow C_n(\mathcal{R}^i)$  be a collection of inclusion maps. Define  $\Phi_{\text{Tot}}^i$  by

$$\Phi_{\text{Tot}}^i([x, y]) = [j_n^i(x)]. \quad (32)$$

One can check that  $\Phi_{\text{Tot}}^i$  is well-defined and bijective (Appendix D).

Given  $[x, y] \in H_n(\text{Tot}^i)$ , note that

$$\iota_*^i \circ \Phi_{\text{Tot}}^i[x, y] = \iota_*^i[j_n^i(x)] = [\iota^i \circ j_n^i(x)],$$

$$\Phi_{\text{Tot}}^{i+1} \circ \iota_{\text{Tot}}^i[x, y] = \Phi_{\text{Tot}}^{i+1}[\iota_n^i(x), \kappa_{n-1}^i(y)] = [j_n^{i+1} \circ \iota_n^i(x)].$$

Then,  $\iota_*^i \circ \Phi_{\text{Tot}}^i = \Phi_{\text{Tot}}^{i+1} \circ \iota_{\text{Tot}}^i$  because all the maps involved are inclusion maps. Thus, Diagram 31 commutes.  $\square$

We now show that  $\mathbb{V}_{\Psi}$  and  $\mathbb{V}_{\text{Tot}}$  are isomorphic persistence modules. Recall that the persistence module  $\mathbb{V}_{\Psi}$  is defined as

$$\mathbb{V}_{\Psi} : H_0(C_{\bullet}\mathcal{F}_n^1) \oplus H_1(C_{\bullet}\mathcal{F}_{n-1}^1) \xrightarrow{\Psi^1} \dots \xrightarrow{\Psi^{N-1}} H_0(C_{\bullet}\mathcal{F}_n^N) \oplus H_1(C_{\bullet}\mathcal{F}_{n-1}^N), \quad (33)$$

where the maps  $\Psi^i$  are defined as

$$\Psi^i(\{\langle x \rangle\}, \{\langle y \rangle\}) = (H_0(\phi_n^i)\{\langle x \rangle\} + (-1)^{n+1}\psi^i\{\langle y \rangle\}, H_1(\phi_{n-1}^i)\{\langle y \rangle\}), \quad (34)$$

where  $H_0(\phi_n^i)$  and  $H_1(\phi_{n-1}^i)$  are the maps defined in Equation (8). Recall from Equation (15) that every  $\{\langle y \rangle\} \in H_1(C_{\bullet}\mathcal{F}_{n-1}^i)$  can be expressed explicitly as  $\{\langle y \rangle\} = \{\langle y_1 \rangle\} + \{\langle y_2 \rangle\}$ , where  $\{\langle y_1 \rangle\} \in A^i$  and  $\{\langle y_2 \rangle\} \in \ker H_1(\phi_{n-1}^i)$ . Then, we can express the map  $\Psi^i$  explicitly using maps from Diagram 10 and  $\delta^i$  as the following

$$\Psi^i(\{\langle x \rangle\}, \{\langle y \rangle\}) = (\{\langle \iota_n^i x \rangle\} + (-1)^{n+1}\delta^i\{\langle y_2 \rangle\}, \{\langle \kappa_n^i y \rangle\}) \quad (35)$$

**Lemma 4.2.** *The persistence modules  $\mathbb{V}_{\Psi}$  and  $\mathbb{V}_{\text{Tot}}$  are isomorphic.*

*Proof.* We will define isomorphisms  $\Phi^i : H_0(C_{\bullet}\mathcal{F}_n^i) \oplus H_1(C_{\bullet}\mathcal{F}_{n-1}^i) \rightarrow H_n(\text{Tot}^i)$  such that the following diagram commutes.

$$\begin{array}{ccccccc} H_0(C_{\bullet}\mathcal{F}_n^1) \oplus H_1(C_{\bullet}\mathcal{F}_{n-1}^1) & \xrightarrow{\Psi^1} & \dots & \xrightarrow{\Psi^{N-1}} & H_0(C_{\bullet}\mathcal{F}_n^N) \oplus H_1(C_{\bullet}\mathcal{F}_{n-1}^N) \\ \downarrow \Phi^1 & & & & \downarrow \Phi^N \\ H_n(\text{Tot}^1) & \xrightarrow{\iota_{\text{Tot}}^1} & \dots & \xrightarrow{\iota_{\text{Tot}}^{N-1}} & H_n(\text{Tot}^N) \end{array} \quad (36)$$

We will define  $\Phi^i : H_0(C_{\bullet}\mathcal{F}_n^i) \oplus H_1(C_{\bullet}\mathcal{F}_{n-1}^i) \rightarrow H_n(\text{Tot}^i)$  by first defining

$$\Phi_0^i : H_0(C_{\bullet}\mathcal{F}_n^i) \rightarrow H_n(\text{Tot}^i), \quad (37)$$

$$\Phi_1^i : H_1(C_{\bullet}\mathcal{F}_{n-1}^i) \rightarrow H_n(\text{Tot}^i). \quad (38)$$

Define  $\Phi_0^i : H_0(C_\bullet \mathcal{F}_n^i) \rightarrow H_n(\text{Tot}^i)$  by

$$\Phi_0^i(\{\langle x \rangle\}) = [x, 0]. \quad (39)$$

To define the map  $\Phi_1^i$ , we will define  $\Phi_1^i$  on the basis  $\mathcal{B}^i$  of  $H_1(C_\bullet \mathcal{F}_{n-1}^i)$ . Recall the fixed basis  $\mathcal{B}^i$  of  $H_1(C_\bullet \mathcal{F}_{n-1}^i)$  in Equation (24). For each basis  $\{\langle b \rangle\} \in \mathcal{B}^i$ , let

$$\Phi_1^i(\{\langle b \rangle\}) = [(-1)^{n+1} \Gamma^i \{\langle b \rangle\}, b^*], \quad (40)$$

where the coset representative  $b^*$  and the set map  $\Gamma^i$  are defined according to the construction in §4.3. Extend  $\Phi_1^i$  linearly to  $H_1(C_\bullet \mathcal{F}_{n-1}^i)$ .

We can now define  $\Phi^i : H_0(C_\bullet \mathcal{F}_n^i) \oplus H_1(C_\bullet \mathcal{F}_{n-1}^i) \rightarrow H_n(\text{Tot}^i)$  by

$$\Phi^i(\{\langle x \rangle\}, \{\langle y \rangle\}) = \Phi_0^i(\{\langle x \rangle\}) + \Phi_1^i(\{\langle y \rangle\}). \quad (41)$$

One can check that  $\Phi^i$  is well-defined and bijective (Appendix E).

To show that Diagram 36 commutes, it suffices to show that the following diagram commutes for each element of  $H_0(C_\bullet \mathcal{F}_n^i)$  and  $H_1(C_\bullet \mathcal{F}_{n-1}^i)$ .

$$\begin{array}{ccc} H_0(C_\bullet \mathcal{F}_n^i) \oplus H_1(C_\bullet \mathcal{F}_{n-1}^i) & \xrightarrow{\Psi^i} & H_0(C_\bullet \mathcal{F}_n^{i+1}) \oplus H_1(C_\bullet \mathcal{F}_{n-1}^{i+1}) \\ \downarrow \Phi^i & & \downarrow \Phi^{i+1} \\ H_n(\text{Tot}^i) & \xrightarrow{\iota_{\text{Tot}}^i} & H_n(\text{Tot}^{i+1}) \end{array} \quad (42)$$

**Case 1:** Given  $\{\langle x \rangle\} \in H_0(C_\bullet \mathcal{F}_n^i)$ , we know from Equations (30) and (39) that

$$\iota_{\text{Tot}}^i \circ \Phi^i(\{\langle x \rangle\}, 0) = \iota_{\text{Tot}}^i([x, 0]) = [\iota_n^i x, 0].$$

On the other hand, we know from Equations (27) and (39) that

$$\Phi^{i+1} \circ \Psi^i(\{\langle x \rangle\}, 0) = \Phi^{i+1}(H_0(\phi_n^i)(x), 0) = \Phi^{i+1}(\{\langle \iota_n^i x \rangle\}, 0) = [\iota_n^i x, 0].$$

Thus, the Diagram 42 commutes for every  $\{\langle x \rangle\} \in H_0(C_\bullet \mathcal{F}_n^i)$ .

**Case 2:** To show that Diagram 42 commutes for every vector in  $H_1(C_\bullet \mathcal{F}_{n-1}^i)$ , we will show that Diagram 42 commutes for every basis element  $\{\langle b \rangle\} \in \mathcal{B}^i$  of  $H_1(C_\bullet \mathcal{F}_{n-1}^i)$ . Recall that  $\mathcal{B}^i = \mathcal{B}_A^i \cup \mathcal{B}_{\ker}^i$ . We consider two cases separately: the first, if  $\{\langle b \rangle\} \in \mathcal{B}_{\ker}^i$ , and the second, if  $\{\langle b \rangle\} \in \mathcal{B}_A^i$ .

**Case 2A:** Assume  $\{\langle b \rangle\} \in \mathcal{B}_{\ker}^i$ . We know from Equations (30) and (40) that

$$\iota_{\text{Tot}}^i \circ \Phi^i(0, \{\langle b \rangle\}) = \iota_{\text{Tot}}^i([(-1)^{n+1} \Gamma^i \{\langle b \rangle\}, b^*]) = [(-1)^{n+1} \iota_n^i \circ \Gamma^i \{\langle b \rangle\}, \kappa_{n-1}^i b^*].$$

On the other hand, from Equations (35) and (40),

$$\begin{aligned} \Phi^{i+1} \circ \Psi^i(0, \{\langle b \rangle\}) &= \Phi^{i+1}((-1)^{n+1} \delta^i \{\langle b \rangle\}, 0) \\ &= \Phi^{i+1}((-1)^{n+1} \{-e_n^{i+1} \alpha^{i+1} + \iota_n^i \circ \Gamma^i \{\langle b \rangle\}\}, 0) \\ &= [(-1)^{n+1} e_n^{i+1} \alpha^{i+1} + (-1)^{n+1} \iota_n^i \circ \Gamma^i \{\langle b \rangle\}, 0]. \end{aligned}$$

Then,

$$\iota_{\text{Tot}}^i \circ \Phi^i(0, \{\langle b \rangle\}) - \Phi^{i+1} \circ \Psi^i(0, \{\langle b \rangle\}) = [(-1)^{n+1} e_n^{i+1} \alpha^{i+1}, \kappa_{n-1}^i b^*].$$

Recall from Equation (13) that  $\partial \alpha^{i+1} = \kappa_{n-1}^i b^*$ . Thus,  $\iota_{\text{Tot}}^i \circ \Phi^i(0, \{\langle b \rangle\}) - \Phi^{i+1} \circ \Psi^i(0, \{\langle b \rangle\})$  is trivial in  $H_n(\text{Tot}^{i+1})$ , and Diagram 42 commutes for  $\{\langle b \rangle\} \in \mathcal{B}_{\ker}^i$ .

**Case 2B:** If  $\{\langle b \rangle\} \in \mathcal{B}_A^i$ , then again by Equations (30) and (40),

$$\iota_{\text{Tot}}^i \circ \Phi^i(0, \{\langle b \rangle\}) = \iota_{\text{Tot}}^i([(-1)^{n+1} \Gamma^i \{\langle b \rangle\}, b^*]) = [(-1)^{n+1} \iota_n^i \circ \Gamma^i \{\langle b \rangle\}, \kappa_{n-1}^i b^*].$$

On the other hand, from Equations (35) and (40),

$$\begin{aligned}
\Phi^{i+1} \circ \Psi^i(0, \{\langle b \rangle\}) &= \Phi^{i+1}(0, \{\langle \kappa_{n-1}^i b \rangle\}) \\
&= \Phi_1^{i+1}(\{\langle \kappa_{n-1}^i b \rangle\}) \\
&= [(-1)^{n+1} \Gamma^{i+1} \{\langle \kappa_{n-1}^i b \rangle\}, \kappa_{n-1}^i b^*] \\
&= [(-1)^{n+1} \iota_n^i \circ \Gamma^i \{\langle b \rangle\}, \kappa_{n-1}^i b^*]
\end{aligned}$$

The last equality follows from the construction of  $\Gamma^{i+1}$  in Equation (23). Thus, the diagram commutes for  $\{\langle b \rangle\} \in \mathcal{B}_A^i$ .

Thus, Diagram 36 commutes.  $\square$

Lemma 4.1 and 4.2 yield the desired result.

**Theorem 3.** *The persistence modules  $\mathbb{V}_\Psi$  and  $\mathbb{V}$  are isomorphic.*

## 5 Application: Multiscale Persistence

There are a number of potential uses for distributed persistent homology computations. Perhaps scalable decentralized computation for large data sets is the most obvious: here the partition of the data set into patches is based on localization via coordinates (this is what appears in [9, 19]). However, even among small data sets, there are reasons for distributing the computation along a partition of the point cloud based on scalar fields other than coordinates. Data often comes with additional features, such as density estimates, distance to a landmark, and time dependence, that one may wish to examine. In this section, we apply the distributed computation method from §4 to point cloud data based on density as a parameter. Section 5.1 provides a general framework for multiscale persistence, and Example 5.2 illustrates the advantage of multiscale analysis when examining dataset with varying density. In such situations, multiscale persistent homology allows the user to detect significant features that are overlooked by standard persistent homology methods.

### 5.1 Multiscale Barcode Annotation

We provide a general framework for computing multiscale persistent homology. Given a point cloud  $P$ , let  $f : P \rightarrow \mathbb{R}$  be a (user-chosen) density estimate. Construct a cover  $\mathcal{V}$  of  $f(P)$  with nerve  $N_{\mathcal{V}}$  a compact interval. Let  $\mathbb{V}$  denote the persistence module

$$\mathbb{V} : H_n(\mathcal{R}^1) \rightarrow H_n(\mathcal{R}^2) \rightarrow \dots \rightarrow H_n(\mathcal{R}^N) \quad (43)$$

in the usual sense. Let  $\text{barcode}(\mathbb{V})$  denote the barcode of  $\mathbb{V}$ . If a bar of the barcode represents a feature  $\gamma$  that consists of points in  $f^{-1}(U)$  for some  $U \in \mathcal{V}$ , we say that the feature  $\gamma$  lives in  $U$ . Moreover, we can annotate the bar with its corresponding set  $U$ . The goal of multiscale persistent homology is to annotate the bars of  $\text{barcode}(\mathbb{V})$  with their corresponding sets  $U$  of  $\mathcal{V}$ .

An algorithmic summary of the annotation process is provided, followed by a detailed explanation of each step.

---

**Algorithm 1** Annotate  $\text{barcode}(\mathbb{V})$ .

---

- 1: Compute persistence module  $\mathbb{V}_*$  using distributed computation.
  - 2: Label vector spaces of  $\mathbb{V}_*$ .
  - 3: For each persistence module  $\mathbb{W}_s$  of  $\mathbb{V}_* = \bigoplus_s \mathbb{W}_s$ , annotate  $\text{barcode}(\mathbb{W}_s)$ .
  - 4: Using the annotations of  $\text{barcode}(\mathbb{V}_*)$ , annotate  $\text{barcode}(\mathbb{V})$ .
  - 5: Return annotated  $\text{barcode}(\mathbb{V})$ .
-



**Step 1. Compute persistence module  $\mathbb{V}_*$**

Let  $\mathbb{V}$  denote the persistence module of interest

$$\mathbb{V} : H_n(\mathcal{R}^1) \rightarrow H_n(\mathcal{R}^2) \rightarrow \cdots \rightarrow H_n(\mathcal{R}^N). \quad (44)$$

Let  $\epsilon_*$  be the upper bound from Lemma 3.1 such that

$$H_n(\mathcal{R}^\epsilon) \cong H_0(C_\bullet \mathcal{F}_n^\epsilon) \oplus H_1(C_\bullet \mathcal{F}_{n-1}^\epsilon)$$

for all  $\epsilon < \epsilon_*$ . Let  $\mathbb{V}|_L$  denote the sequence of vector spaces and maps of  $\mathbb{V}$  up to parameter  $\epsilon_L$  such that  $\epsilon_L < \epsilon_*$ :

$$\mathbb{V}|_L : H_n(\mathcal{R}^1) \rightarrow H_n(\mathcal{R}^2) \rightarrow \cdots \rightarrow H_n(\mathcal{R}^L).$$

We can compute the persistence module

$$\mathbb{V}_\Psi : H_0(C_\bullet \mathcal{F}_n^1) \oplus H_1(C_\bullet \mathcal{F}_{n-1}^1) \xrightarrow{\Psi^1} \cdots \xrightarrow{\Psi^{L-1}} H_0(C_\bullet \mathcal{F}_n^L) \oplus H_1(C_\bullet \mathcal{F}_{n-1}^L) \quad (45)$$

isomorphic to  $\mathbb{V}|_L$  using the distributed computation method from §4.

We will in fact compute a persistence module  $\mathbb{V}_* \cong \mathbb{V}_\Psi$  that can reveal additional information about the barcode. Recall from §2.2 that each cosheaf  $\mathcal{F}_n^i$  can be decomposed as  $\mathcal{F}_n^i \cong \oplus \mathcal{J}_n^i$ , where each  $\mathcal{J}_n^i$  is an indecomposable cosheaf over  $N_{\mathcal{V}}$ . In other words, there exists an isomorphism of cosheaves

$$D_n^i : \mathcal{F}_n^i \rightarrow \oplus \mathcal{J}_n^i. \quad (46)$$

For each parameter  $\epsilon_i$ , there exists an isomorphism

$$g^i : H_0(C_\bullet \mathcal{F}_n^i) \oplus H_1(C_\bullet \mathcal{F}_{n-1}^i) \rightarrow H_0(C_\bullet \oplus \mathcal{J}_n^i) \oplus H_1(C_\bullet \mathcal{F}_{n-1}^i)$$

defined by

$$g^i(\{\langle x \rangle\}, \{\langle y \rangle\}) = (H_0(D_n^i)\{\langle x \rangle\}, \{\langle y \rangle\}),$$

where  $H_0(D_n^i) : H_0(C_\bullet \mathcal{F}_n^i) \rightarrow H_0(C_\bullet \oplus \mathcal{J}_n^i)$  is the isomorphism induced by  $D_n^i$ .

Define the persistence module  $\mathbb{V}_*$  by

$$\mathbb{V}_* : H_0(C_\bullet \oplus \mathcal{J}_n^1) \oplus H_1(C_\bullet \mathcal{F}_{n-1}^1) \xrightarrow{\Psi_*^1} \cdots \xrightarrow{\Psi_*^{N-1}} H_0(C_\bullet \oplus \mathcal{J}_n^N) \oplus H_1(C_\bullet \mathcal{F}_{n-1}^N),$$

where the map  $\Psi_*^i$  is defined by  $\Psi_*^i = g^{i+1} \circ \Psi^i \circ (g^i)^{-1}$ .

By construction,  $\mathbb{V}_*$  is isomorphic to the persistence module  $\mathbb{V}_\Psi$  and hence isomorphic to  $\mathbb{V}|_L$ . Our mechanism of decomposing the cosheaf  $\mathcal{F}_n^i$  into indecomposable cosheaves may seem like a cumbersome step. However, such decomposition allows us to understand the cosheaf homologies in terms of the indecomposable cosheaves  $\mathcal{J}_n^i$ .

**Step 2. Label the vector spaces of  $\mathbb{V}_*$**

For any parameter  $\epsilon_i$ , recall that

$$\mathbb{V}_*^i = H_0(C_\bullet \oplus \mathcal{J}_n^i) \oplus H_1(C_\bullet \mathcal{F}_{n-1}^i).$$

By Lemma 2.1, each component of  $H_0(C_\bullet \oplus \mathcal{J}_n^i)$  corresponds to an indecomposable cosheaf of the form  $\mathcal{J}_{[-]}^i$ . We can thus annotate each component of  $H_0(C_\bullet \oplus \mathcal{J}_n^i)$  according the support of the corresponding indecomposable  $\mathcal{J}_{[-]}^i$ .

*Case 1:* If the indecomposable  $\mathcal{J}_{[-]}^i$  is supported on a unique vertex  $v \in N_{\mathcal{V}}$ , then annotate the component by  $U \in \mathcal{V}$ , where  $U$  is the open set corresponding to the vertex  $v$ .

*Case 2:* Let  $v_l, v_r \in N_{\mathcal{V}}$  each denote the left and rightmost supports of  $\mathcal{J}_{[-]}^i$ . If  $v_l, v_{l+1}, \dots, v_r$  are the vertices of  $N_{\mathcal{V}}$  between  $v_l$  and  $v_r$ , then the cosheaf  $\mathcal{J}_{[-]}^i$  represents a feature that lives in all  $U_l, U_{l+1}, \dots, U_r \in \mathcal{V}$ . The user can annotate the corresponding component by  $[U_l, U_r]$  or  $U_l$  or  $U_r$ , depending on the user's goal.

For example, assume that

$$\mathbb{V}_*^i = H_0(C_\bullet \oplus \mathcal{J}_n^i) \oplus H_1(C_\bullet \mathcal{F}_{n-1}^i) = \mathbb{K} \oplus \mathbb{K} \oplus \mathbb{K} \oplus \mathbb{K},$$

where the first three components come from  $H_0(C_\bullet \oplus \mathcal{J}_n^i)$  and the last component  $\mathbb{K}$  comes from  $H_1(C_\bullet \mathcal{F}_{n-1}^i)$ . An example of cosheaf  $\oplus \mathcal{J}_n^i$  is illustrated in Figure 6.

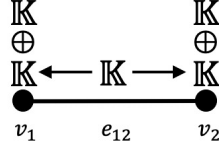


Figure 6: An example decomposition of a cosheaf  $\mathcal{F}_n^i \cong \oplus \mathcal{J}_n^i$

Then one can label the components of  $H_0(C_\bullet \oplus \mathcal{J}_n^i)$  by  $\mathbb{K}_1 \oplus \mathbb{K}_2 \oplus \mathbb{K}_{1,2}$ , where each label corresponds to the support of the indecomposable cosheaf in Figure 6. Then, the vector space  $\mathbb{V}_*^i$  can be labeled as  $\mathbb{K}_1 \oplus \mathbb{K}_2 \oplus \mathbb{K}_{1,2} \oplus \mathbb{K}$ .

**Step 3. Annotate the *barcode*( $\mathbb{W}_s$ ) for each  $\mathbb{W}_s$  of  $\mathbb{V}_* = \bigoplus_s \mathbb{W}_s$**

Note that  $\mathbb{V}_*$  can be expressed naturally as a sum of persistence modules as

$$\mathbb{V}_* = \bigoplus_s \mathbb{W}_s.$$

For example, the persistence module

$$\mathbb{V}_* : \mathbb{R}^3 \xrightarrow{\begin{bmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 1 \end{bmatrix}} \mathbb{R}^3 \xrightarrow{\begin{bmatrix} 1 & -1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \end{bmatrix}} \mathbb{R}^3$$

is the sum of persistence modules

$$\mathbb{W}_1 : \mathbb{R} \xrightarrow{\begin{bmatrix} 1 \\ 1 \end{bmatrix}} \mathbb{R}^2 \xrightarrow{\begin{bmatrix} 1 & -1 \end{bmatrix}} \mathbb{R} \quad \text{and} \quad \mathbb{W}_2 : \mathbb{R}^2 \xrightarrow{\begin{bmatrix} 1 & 1 \end{bmatrix}} \mathbb{R} \xrightarrow{\begin{bmatrix} 1 \\ 1 \end{bmatrix}} \mathbb{R}^2.$$

Moreover,  $\text{barcode}(\mathbb{V}_*)$  is the collection of barcodes  $\text{barcode}(\mathbb{W}_s)$ . For each  $\mathbb{W}_s$ , compute  $\text{barcode}(\mathbb{W}_s)$ . Let  $\bar{b}$  be a bar of  $\text{barcode}(\mathbb{W}_s)$  born at parameter  $\epsilon_i$ . Consider the annotation of the components of  $\mathbb{W}_s^i$  from Step 2. There are two cases to consider.

*Case 1:* All components of  $\mathbb{W}_s^i$  have been annotated by a unique set  $U \in \mathcal{V}$ . Bar  $\bar{b}$  then represents some linear combination of features in  $U$ , so annotate  $\bar{b}$  with  $U$ .

*Case 2:* The components of  $\mathbb{W}_s^i$  have been annotated by two or more sets in  $\mathcal{V}$ , say  $U_j$  and  $U_k$ . The user can decide to either not annotate the bar at all, to annotate the bar by  $U_j$ , or to annotate the bar by  $U_k$ , depending on the question of interest.

The result of Step 3 is an annotation of  $\text{barcode}(\mathbb{V}_*)$ .

**Step 4. Annotate *barcode*( $\mathbb{V}$ )**

We can use the annotations of  $\text{barcode}(\mathbb{V}_*)$  to annotate  $\text{barcode}(\mathbb{V})$ . Note that  $\text{barcode}(\mathbb{V}_*)$  can be obtained from  $\text{barcode}(\mathbb{V})$  by truncating  $\text{barcode}(\mathbb{V})$  at parameter  $\epsilon_L$ . Let  $[b, d]$  be a bar of  $\text{barcode}(\mathbb{V}_*)$  that has been annotated by a set  $U$  in Step 3. There are two cases to consider.

*Case 1:* If  $d < \epsilon_L$ , then annotate the bar  $[b, d]$  of  $\text{barcode}(\mathbb{V})$  with  $U$ .

*Case 2:* If  $d = \epsilon_L$  and  $[b, d]$  is the unique bar of  $\text{barcode}(\mathbb{V}_*)$  with birth time  $b$ , then there exists a unique bar  $[b, d']$  in  $\text{barcode}(\mathbb{V})$  with birth time  $b$ . Annotate the bar  $[b, d']$  of  $\text{barcode}(\mathbb{V})$  with  $U$ .

The final result of the algorithm is an annotation of  $\text{barcode}(\mathbb{V})$ . One can use this annotated barcode to perform finer data analysis.

## 5.2 Example: Multiscale Persistence

Consider a situation where the size of a feature depends on the density of the constituting points, as illustrated in Figure 7. Figure 8 illustrates the corresponding barcode, which suggests that there is one significant feature. Standard persistent homology fails to detect the small, but densely sampled features. Multiscale persistent homology can select the bars that correspond to small but densely sampled features and annotate them as being significant.

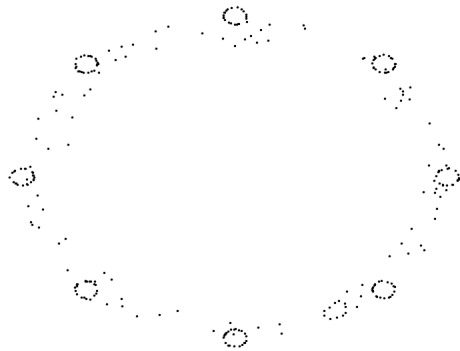


Figure 7: A point cloud with varying density



Figure 8: Barcode from standard persistent homology in dimension 1

Let  $P$  denote the point cloud in Figure 7, and let  $f : P \rightarrow \mathbb{R}$  be the function mapping each point to its estimated density value. In our example,  $f(p)$  represents the number of points in a radius  $r$ -ball centered at  $p$ .

We chose a covering  $\mathcal{V} = \{U_s, U_d\}$  of  $f(P)$  by the following. We first plotted a histogram of density values as illustrated in Figure 9, and decided to cover  $f(P)$  with two sets,  $U_s$  and  $U_d$ , where  $U_s = (0, 18)$  and  $U_d = (8, 26)$ . We will refer to points in  $f^{-1}(U_s)$  as the sparse points, and we will refer to points in  $f^{-1}(U_d)$  as the dense points. Figures 10a and 10b illustrate the sparse and dense points.

We now follow Algorithm 1. Let

$$\mathbb{V} : H_1(\mathcal{R}^1) \rightarrow \dots \rightarrow H_1(\mathcal{R}^N)$$

be the persistence module obtained from the point cloud  $P$ . For this example, the maximum parameter is  $\epsilon_N = 1.6$ . Let  $\epsilon_*$  be the upper bound of the parameter  $\epsilon$  from Lemma 3.1 for which the isomorphism

$$H_n(\mathcal{R}^\epsilon) \cong H_0(C_\bullet \mathcal{F}_n^\epsilon) \oplus H_1(C_\bullet \mathcal{F}_{n-1}^\epsilon)$$

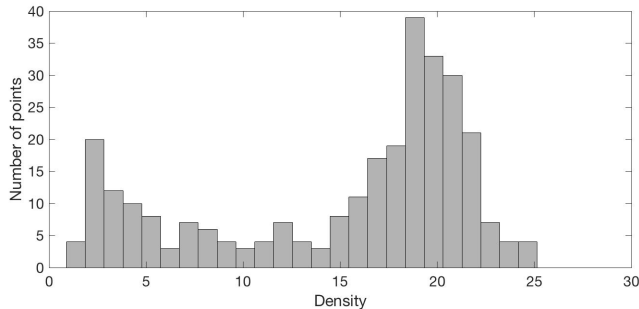


Figure 9: Histogram plot of estimated density values.



Figure 10: Sparse and dense points.

holds. For this example,  $\epsilon_*$  is 0.0719. Compute the persistence module

$$\mathbb{V}_* : H_0(C_\bullet \oplus \mathcal{J}_1^1) \oplus H_1(C_\bullet \mathcal{F}_0^1) \rightarrow \cdots \rightarrow H_0(C_\bullet \oplus \mathcal{J}_1^{\epsilon_*}) \oplus H_1(C_\bullet \mathcal{F}_0^{\epsilon_*}) \quad (47)$$

following Step 1 of Algorithm 1.

Step 2 of Algorithm 1 labels the components of vector space  $H_0(C_\bullet \oplus \mathcal{J}_1^i)$  according to the support of the indecomposable cosheaves  $\mathcal{J}_{[-]}^i$ .

Step 3 of Algorithm 1 results in an annotated version of  $\text{barcode}(\mathbb{V}_*)$ , illustrated in Figure 11. The top two gray bars have been annotated by  $U_s$ . The two bars represent features in the sparse points. The remaining black bars have been annotated by  $U_d$ , and they represent features in the dense points.



Figure 11: Annotated  $\text{barcode}(\mathbb{V}_*)$ . The top two gray bars have been annotated by  $U_s$ , and they represent features in the sparse points. The remaining black bars have been annotated by  $U_d$ , and they represent features in the dense points.

Step 4 of Algorithm 1 allows us to transfer the annotation of  $\text{barcode}(\mathbb{V}_*)$  to  $\text{barcode}(\mathbb{V})$  resulting in an annotated version of  $\text{barcode}(\mathbb{V})$  illustrated in Figure 12. The two bars enclosed by the gray box are annotated by  $U_s$ , and the bars enclosed by black box are annotated by  $U_d$ .

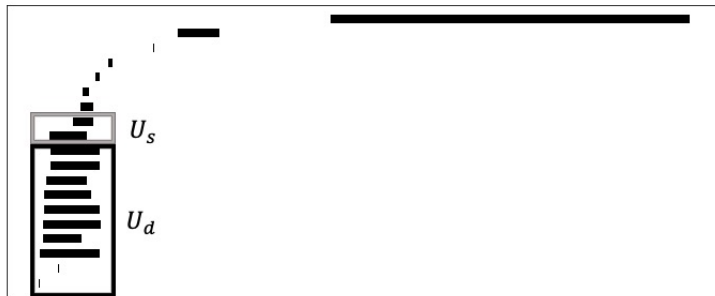


Figure 12: Annotated  $\text{barcode}(\mathbb{V})$ . The two bars enclosed by the gray box are annotated by  $U_s$ , and the bars dyenclosed by black box are annotated by  $U_d$ .

The goal is to determine the small but significant features that consist of the denser points. Thus, we focus on the bars of Figure 12 that have been annotated by  $U_d$ . By restricting our attention to only the bars that represent features in  $U_d$ , we are able to determine the significant features built among the dense points. By examining the bars annotated by  $U_d$ , one can conclude that there are eight significant bars.

Lastly, we return to  $\text{barcode}(\mathbb{V})$  and indicate the significant features. We then obtain the barcode in Figure 13, where the black bars indicate significant features and the gray bars indicate noise. Note that we have one long black bar, which is deemed significant because of its length. We have eight additional shorter significant bars which were identified via Algorithm 1.



Figure 13: Final annotation of  $\text{barcode}(\mathbb{V})$

The persistent homology computation Julia package *Eirene.jl* [18] identifies persistent homology generators in the point cloud. Using this, we identified the points of  $P$  that constitute each significant feature. The eight significant short bars identified indeed correspond to the eight small but densely sampled features in Figure 7.

## Appendix A Proof of Lemma 3.1

**Lemma 3.1.** *Let  $P$  be a point cloud with finite cover  $\mathcal{V}$  having one-dimensional nerve  $N_{\mathcal{V}}$ . There exists a constant  $\epsilon_*$  such that*

$$H_n(\mathcal{R}^\epsilon) \cong H_0(C_\bullet \mathcal{F}_n^\epsilon) \oplus H_1(C_\bullet \mathcal{F}_{n-1}^\epsilon) \quad (3)$$

for every  $0 < \epsilon < \epsilon_*$ .

*Proof.* We first specify  $\epsilon_*$ . In what follows, use the convention that the minimum over an empty set is  $\infty$ .

Each  $p \in P$  lies in either one or two elements of the cover  $\mathcal{V}$ . If unique,  $p \in U \in \mathcal{V}$ , then set

$$K_p = \min_{\{q \notin U\}} d(p, q). \quad (48)$$

If  $p$  lies in two sets of the cover,  $p \in U \cap W$ , then first let

$$K'_p = \min_{\{q \notin U \cup W\}} d(p, q), \quad (49)$$

and let

$$K''_p = \min_{\substack{\{q, q' \mid d(p, q) < K'_p, \\ d(p, q') < K'_p, \\ q \notin U, \\ q' \notin W\}} d(q, q'). \quad (50)$$

Then set

$$K_p = \min\{K'_p, K''_p\}.$$

Let  $\epsilon_* = \min_{p \in P} K_p$ . Assume  $\epsilon < \epsilon_*$ . We assert Equation (3) by showing that the Rips system covers  $\mathcal{R}^\epsilon$ . Let  $\omega = (v_0, \dots, v_l)$  be a simplex of  $\mathcal{R}^\epsilon$  with vertices as listed. The pairwise distances thus satisfy  $d(v_i, v_j) < \epsilon$ .

If there exists a vertex of  $\omega$ , say  $v_0$ , such that  $v_0$  belongs to a unique  $U \in \mathcal{V}$ , then by construction,  $\epsilon_* < K_{v_0} = \min_{\{q \notin U\}} d(v_0, q)$  from Equation (48). Thus, for any other vertex  $v$  of  $\omega$ , we have  $d(v_0, v) < \epsilon < K_{v_0}$ , and hence all of  $\omega$  is covered by  $U$ .

Otherwise, every vertex  $v$  of  $\omega$  is covered by two sets in  $\mathcal{V}$ . Without loss of generality, assume that  $v_0 \in U \cap W$ . Note that for any other vertex  $v$  of  $\omega$ , we have

$$d(v_0, v) < \epsilon < K_{v_0} \leq K'_{v_0}, \quad (51)$$

where  $K'_{v_0}$  is given by Equation (49). So  $v \in U \cup W$  for every  $v \in \omega$ . In fact, we can show that all of  $\omega$  lies in  $U$  or in  $W$ . Assuming the contrary, there exist distinct vertices, say  $v_1$  and  $v_2$ , such that  $v_1 \notin U$  and  $v_2 \notin W$ . By construction,  $d(v_0, v_1) < \epsilon < K'_{v_0}$ , and  $d(v_0, v_2) < \epsilon < K'_{v_0}$ . By definition of  $K''_{v_0}$  from Equation (50), we know that  $K''_{v_0} \geq d(v_1, v_2)$ . However, this contradicts the fact that  $d(v_1, v_2) < \epsilon < K''_{v_0}$ . Thus,  $\omega$  is covered by some subcomplex  $\mathcal{R}_\sigma^\epsilon$ . Thus, the Rips system covers  $\mathcal{R}^\epsilon$ , and Lemma 3.1 follows from the proof of the analogous result in [13].  $\square$

## Appendix B Independence of $\alpha^{i+1}$

**Lemma B.1.** *The construction of  $\delta^i$  on a basis element  $\{b\} \in \mathcal{B}_{\ker}^i$  in Equation (14) does not depend on the choice of  $\alpha^{i+1}$ .*

*Proof.* Let  $\alpha_1, \alpha_2 \in \bigoplus_{e \in N_\mathcal{V}} C_n(\mathcal{R}_e^{i+1})$  be two different choices of  $\alpha$  that satisfy Equation (13):  $\partial\alpha = \kappa_{n-1}^i b^*$ .

Note that  $\partial(\alpha_1 - \alpha_2) = \kappa_{n-1}^i b^* - \kappa_{n-1}^i b^* = 0$ . So  $\langle \alpha_1 - \alpha_2 \rangle$  represents an element in  $\bigoplus_{e \in N_\mathcal{V}} H_n(\mathcal{R}_e^{i+1})$ .

Then,  $\langle e_n^{i+1}(\alpha_1 - \alpha_2) \rangle = \partial_n^{i+1} \langle \alpha_1 - \alpha_2 \rangle \in \text{im } \partial_n^{i+1}$ . So  $\{\langle e_n^{i+1}(\alpha_1 - \alpha_2) \rangle\}$  is trivial in  $H_0(C_\bullet \mathcal{F}_n^{i+1})$ . Hence,  $\{\langle -e_n^{i+1} \alpha_1 + \iota_n^i \circ \beta^i \rangle\} = \{\langle -e_n^{i+1} \alpha_2 + \iota_n^i \circ \beta^i \rangle\}$  in  $H_0(C_\bullet \mathcal{F}_n^{i+1})$ . Thus, the map  $\delta^i$  does not depend on the choice of  $\alpha$ .  $\square$

## Appendix C Obtaining basis $\mathcal{C}_{\text{im}}^i$ from basis $\mathcal{B}_A^{i-1}$ of $A^{i-1}$

**Lemma C.1.** *Let  $\mathcal{B}^{i-1} = \mathcal{B}_A^{i-1} \cup \mathcal{B}_{\ker}^{i-1}$  be a basis of  $H_1(C_\bullet \mathcal{F}_{n-1}^{i-1}) = A^{i-1} \oplus \ker H_1(\phi_{n-1}^{i-1})$ . Let  $\mathcal{B}_A^{i-1} = \{\langle b_1 \rangle, \dots, \langle b_t \rangle\}$  be the basis of  $A^{i-1}$ . Then,*

$$\{\langle \kappa_{n-1}^{i-1} b_1 \rangle, \dots, \langle \kappa_{n-1}^{i-1} b_t \rangle\}$$

*are linearly independent in  $H_1(C_\bullet \mathcal{F}_{n-1}^i)$ .*

*Proof.* Assume the contrary, that

$$c_1\{\langle \kappa_{n-1}^{i-1} b_1 \rangle\} + \cdots + c_t\{\langle \kappa_{n-1}^{i-1} b_t \rangle\} = \{\langle 0 \rangle\}$$

for some  $c_1, \dots, c_t$  that are not all zero. By construction, this implies that

$$c_1\langle \kappa_{n-1}^{i-1} b_1 \rangle + \cdots + c_t\langle \kappa_{n-1}^{i-1} b_t \rangle = \langle 0 \rangle$$

for some  $c_1, \dots, c_t$  that are not all zero. Then,  $\langle c_1 b_1 + \cdots c_t b_t \rangle \in \ker H_1(\phi_{n-1}^{i-1})$ . Note that  $\langle c_1 b_1 + \cdots c_t b_t \rangle \in A^{i-1}$  as well since  $A^{i-1}$  is a subspace of  $H_1(C_\bullet \mathcal{F}_{n-1}^{i-1})$ . This contradicts the fact that  $H_1(C_\bullet \mathcal{F}_{n-1}^{i-1})$  is a direct sum of  $\ker H_1^{i-1}(\phi_{n-1})$  and  $A^{i-1}$ . Thus,  $\{\langle \kappa_{n-1}^{i-1} b_1 \rangle\}, \dots, \{\langle \kappa_{n-1}^{i-1} b_t \rangle\}$  are linearly independent.  $\square$

## Appendix D Details of proof of Lemma 4.1

**Lemma D.1.** *The map  $\Phi_{\text{Tot}}^i : H_n(\text{Tot}^i) \rightarrow H_n(\mathcal{R}^i)$  is well-defined and bijective.*

For clarity, we omit the superscript  $i$  indicating the parameter  $\epsilon_i$ .

*Proof.* We first show that  $\Phi_{\text{Tot}} : H_n(\text{Tot}) \rightarrow H_n(\mathcal{R})$  is a well-defined map. Assume that  $[x_1, y_1] = [x_2, y_2]$  in  $H_n(\text{Tot})$ . One can build a commutative diagram whose rows are short exact sequences

$$0 \longleftarrow C_n(\mathcal{R}) \xleftarrow{j_n} \bigoplus_{v \in N_v} C_n(\mathcal{R}_v) \xleftarrow{e_n} \bigoplus_{e \in N_e} C_n(\mathcal{R}_e) \longleftarrow 0 \quad (52)$$

and whose vertical maps are the boundary operator  $\partial_n$ . Since  $[x_1, y_1] = [x_2, y_2]$  in  $H_n(\text{Tot})$ , there exists  $p_{n+1} \in \bigoplus_{v \in N_v} C_{n+1}(\mathcal{R}_v)$  and  $q_n \in \bigoplus_{e \in N_e} C_n(\mathcal{R}_e)$  such that  $y_2 - y_1 = \partial q_n$  and  $x_2 - x_1 = \partial p_{n+1} + (-1)^{n+1} e_n q_n$ .

Then,

$$\begin{aligned} j_n(x_2 - x_1) &= j_n(\partial p_{n+1} + (-1)^{n+1} e_n q_n) \\ &= j_n(\partial p_{n+1}) \\ &= \partial(j_{n+1}(p_{n+1})). \end{aligned}$$

The second equality follows from the exactness of Diagram 52. Thus,  $[j_n(x_2)] = [j_n(x_1)]$ , and the map  $\Phi_{\text{Tot}}$  is well-defined.

We now show that  $\Phi_{\text{Tot}}$  is surjective. Let  $[\gamma] \in H_n(\mathcal{R})$ . Since the rows of Diagram 52 are exact, there exists  $x_n \in \bigoplus_{v \in N_v} C_n(\mathcal{R}_v)$  such that  $\gamma = j_n(x_n)$ . Then,

$$j_{n-1} \circ \partial x_n = \partial \circ j_n(x_n) = \partial \gamma = 0.$$

So  $\partial x_n \in \ker j_{n-1}$ . By exactness, there exists  $y_{n-1} \in \bigoplus_{e \in N_e} C_{n-1}(\mathcal{R}_e)$  such that  $\partial x_n = e_{n-1}(y_{n-1})$ . Moreover,

$$e_{n-2} \circ \partial y_{n-1} = \partial \circ e_{n-1} y_{n-1} = \partial \partial x_n = 0.$$

Since  $e_{n-2}$  is injective, we know that  $\partial y_{n-1} = 0$ . Then,  $[x_n, (-1)^{n-1} y_{n-1}] \in H_n(\text{Tot})$ , and  $\Psi_{\text{Tot}}[x_n, (-1)^{n-1} y_{n-1}] = [\gamma]$ . Thus,  $\Psi_{\text{Tot}}$  is surjective.

Lastly, we show that  $\Phi_{\text{Tot}}$  is injective. Assume that  $\Phi_{\text{Tot}}([x, y]) = [j_n(x)] = 0$ . Then, there exists  $p_{n+1} \in C_{n+1}(\mathcal{R})$  such that  $\partial p_{n+1} = j_n(x)$ . Since  $j_n$  and  $j_{n+1}$  are surjective, there exists  $p'_{n+1} \in \bigoplus_{v \in N_v} C_{n+1}(\mathcal{R}_v)$  such that  $p_{n+1} = j_{n+1}(p'_{n+1})$ . Then,

$$\begin{aligned} j_n(\partial p'_{n+1} - x) &= j_n \circ \partial p'_{n+1} - j_n(x) \\ &= \partial \circ j_{n+1}(p'_{n+1}) - j_n(x) \\ &= \partial p_{n+1} - j_n(x) = 0. \end{aligned}$$

Thus,  $\partial p'_{n+1} - x \in \ker j_n$ . From the exactness of rows of Diagram 52, there exists  $q_n \in \bigoplus_{e \in N_v} C_n(\mathcal{R}_e)$  such that  $\partial p'_{n+1} - x = e_n(q_n)$ . Note that while  $\partial(\partial p'_{n+1} - x) = \partial e_n(q_n)$ , this is equal to  $-\partial x = -(-1)^{n+1}e_n(y)$  by definition. Then,  $e_n \partial q_n = \partial e_n(q_n) = -\partial x = -(-1)^{n+1}e_n(y)$ . Since  $e_n$  is injective, this implies that  $\partial q_n = -(-1)^{n+1}y$ . Let  $q'_n = -(-1)^{n+1}q_n$ , so that  $\partial q'_n = y$ .

So far, we found  $p'_{n+1} \in \bigoplus_{v \in N_v} C_{n+1}(\mathcal{R}_v)$  and  $q'_n \in \bigoplus_{e \in N_v} C_n(\mathcal{R}_e)$  such that  $\partial q'_n = y$  and

$$x = \partial p'_{n+1} - e_n(q_n) = \partial p'_{n+1} - e_n(-(-1)^{n+1}q'_n) = \partial p'_{n+1} + (-1)^{n+1}e_n q'_n.$$

Thus,  $[x, y] = 0$  in  $H_n(\text{Tot})$ , and  $\Phi_{\text{Tot}}$  is injective.  $\square$

## Appendix E Details of the proof of Lemma 4.2

**Lemma E.1.** *The map  $\Phi^i : H_0(C \bullet \mathcal{F}_n^i) \oplus H_1(C \bullet \mathcal{F}_{n-1}^i) \rightarrow H_n(\text{Tot}^i)$  defined in Equation (41) is well-defined and bijective.*

*Proof.* We first show that  $\Phi^i$  is well-defined. Assume that  $(\langle x \rangle, \langle y \rangle) = (\langle x' \rangle, \langle y' \rangle)$  in  $H_0(C \bullet \mathcal{F}_n^i) \oplus H_1(C \bullet \mathcal{F}_{n-1}^i)$ , i.e.,  $\langle x \rangle = \langle x' \rangle$  in  $H_0(C \bullet \mathcal{F}_n^i)$  and  $\langle y \rangle = \langle y' \rangle$  in  $H_1(C \bullet \mathcal{F}_{n-1}^i)$ .

Note that

$$\Phi_0^i(\langle x \rangle) - \Phi_0^i(\langle x' \rangle) = [x - x', 0].$$

Since  $\langle x \rangle = \langle x' \rangle$  in  $H_0(C \bullet \mathcal{F}_n^i)$ , there exists  $p_{n+1} \in \bigoplus_{v \in N_v} C_{n+1}(\mathcal{R}_v^i)$  such that  $\partial p_{n+1} = x - x'$ . Thus,  $[x - x', 0]$  is trivial in  $H_n(\text{Tot})$ , and  $\Phi_0^i$  is a well-defined map.

By construction,  $\Phi_1^i(\langle y \rangle) = \Phi_1^i(\langle y' \rangle)$ . Thus,  $\Phi^i$  is a well-defined map.

We now show that  $\Phi^i$  is surjective. Given  $[x, y] \in H_n(\text{Tot}^i)$ , we know that  $\partial y = 0$ , so  $\langle y \rangle$  is an element of  $H_1(C \bullet \mathcal{F}_{n-1}^i)$ . If  $\mathcal{B} = \{\langle b_1 \rangle, \dots, \langle b_t \rangle\}$  is a basis of  $H_1(C \bullet \mathcal{F}_{n-1}^i)$ , and  $b_1^*, \dots, b_t^*$  are the coset representatives of the basis, then  $\langle y \rangle$  can be written as

$$\langle y \rangle = c_1 \langle b_1^* \rangle + \dots + c_t \langle b_t^* \rangle$$

for some  $c_1, \dots, c_t$ . That is, there exists  $q_n \in \bigoplus_{e \in N_v} C_n(\mathcal{R}_e^i)$  such that

$$c_1 b_1^* + \dots + c_t b_t^* - y = \partial q_n. \quad (53)$$

Recall from Equation (26) that

$$\partial \Gamma^i \langle b \rangle = e_{n-1}^i(b^*) \quad (54)$$

for every  $\langle b \rangle \in \mathcal{B}$ . Let

$$r_n = x - (-1)^{n+1}(c_1 \Gamma^i \langle b_1^* \rangle + \dots + c_t \Gamma^i \langle b_t^* \rangle) + (-1)^{n+1}e_n^i(q_n). \quad (55)$$

We know that  $\partial x = (-1)^{n-1}e_{n-1}^i(y)$ , and, by commutativity of Diagram 10, we have  $\partial e_n^i(q_n) = e_{n-1}^i(\partial q_n)$ . Thus, it follows that

$$\begin{aligned} \partial r_n &= \partial x - (-1)^{n+1} \partial(c_1 \Gamma^i \langle b_1^* \rangle + \dots + c_t \Gamma^i \langle b_t^* \rangle) + (-1)^{n+1} \partial e_n^i(q_n) \\ &= (-1)^{n-1} e_{n-1}^i(y) - (-1)^{n+1} e_{n-1}^i(c_1 b_1^* + \dots + c_t b_t^*) + (-1)^{n+1} e_{n-1}^i(\partial q_n) \\ &= 0. \end{aligned}$$

The second equality follows from Equation (54) and Diagram 10. The third equality follows from Equation (53). Thus,  $\langle r_n \rangle$  represents an element of  $H_0(C \bullet \mathcal{F}_n^i)$ . Then,

$$\begin{aligned} \Phi^i(\langle r_n \rangle, \langle y \rangle) &= \Phi_0^i(\langle r_n \rangle) + \Phi_1^i(\langle y \rangle) \\ &= [r_n + (-1)^{n+1}(c_1 \Gamma^i \langle b_1^* \rangle + \dots + c_t \Gamma^i \langle b_t^* \rangle), c_1 b_1^* + \dots + c_t b_t^*] \\ &= [x + (-1)^{n+1}e_n^i(q_n), y + \partial q_n] \\ &= [x, y] + [(-1)^{n+1}e_n^i(q_n), \partial q_n] \\ &= [x, y] \end{aligned}$$



The third equality follows from Equations (53) and (55). Thus,  $\Phi^i$  is surjective.

Lastly, we show that  $\Phi^i$  is injective. Let  $(\{x\}, \{y\}) \in H_0(C_\bullet \mathcal{F}_n^i) \oplus H_1(C_\bullet \mathcal{F}_{n-1}^i)$ . If  $\mathcal{B} = \{\{b_1\}, \dots, \{b_t\}\}$  is a basis of  $H_1(C_\bullet \mathcal{F}_{n-1}^i)$ , and  $b_1^*, \dots, b_t^*$  are the coset representatives of the basis, then  $\{y\}$  can be written as

$$\{y\} = c_1\{b_1^*\} + \dots + c_t\{b_t^*\}$$

for some  $c_1, \dots, c_t$ . Assume that

$$\Phi^i(\{x\}, \{y\}) = [x + (-1)^{n+1}(c_1\Gamma^i\{b_1^*\}) + \dots + c_t\Gamma^i\{b_t^*\}], c_1b_1^* + \dots + c_tb_t^*] = 0.$$

Then, there exists  $q_n \in \bigoplus_{e \in N_v} C_n(\mathcal{R}_e^i)$  and  $p_{n+1} \in \bigoplus_{v \in N_v} C_{n+1}(\mathcal{R}_v^i)$  such that

$$\partial q_n = c_1b_1^* + \dots + c_tb_t^*, \quad (56)$$

$$\partial p_{n+1} + (-1)^{n-1}e_n^i(q_n) = x + (-1)^{n+1}(c_1\Gamma^i\{b_1^*\} + \dots + c_t\Gamma^i\{b_t^*\}). \quad (57)$$

From Equation (56), we know  $c_1\{b_1^*\} + \dots + c_t\{b_t^*\} = \{y\}$  is trivial in  $H_1(C_\bullet \mathcal{F}_{n-1}^i)$ . Thus,  $\Phi^i(\{x\}, \{y\}) = \Phi^i(\{x\}, 0) = [x, 0]$ .

If  $[x, 0]$  is trivial in  $H_n(\text{Tot}^i)$ , then there exists  $a_n \in \bigoplus_{e \in N_v} C_n(\mathcal{R}_e^i)$  and  $b_{n+1} \in \bigoplus_{v \in N_v} C_{n+1}(\mathcal{R}_v^i)$  such that

$$\partial a_n = 0,$$

$$\partial b_{n+1} + (-1)^{n-1}e_n^i a_n = x.$$

The above two equations imply that  $\{x\}$  is trivial in  $H_0(C_\bullet \mathcal{F}_n^i)$  as well. Thus,  $\Phi^i$  is injective.  $\square$

## References

- [1] U. Bauer, M. Kerber, and J. Reininghaus. Clear and compress: computing persistent homology in chunks. In *Topological Methods in Data Analysis and Visualization III*, Mathematics and Visualization, pages 103–117. 2014.
- [2] T. Berry and T. Sauer. Consistent manifold representation for topological data analysis. *Foundations of Data Science*, 1(1):1–38, 2019.
- [3] G. Bredon. *Sheaf Theory*. Springer, 1997.
- [4] G. Carlsson. Topology and data. *Bull. Amer. Math. Soc. (N.S.)*, 46(2):255–308, 2009.
- [5] G. Carlsson. The shape of data. In *Foundations of computational mathematics, Budapest 2011*, volume 403 of *London Math. Soc. Lecture Note Ser.*, pages 16–44. Cambridge Univ. Press, Cambridge, 2013.
- [6] G. Carlsson and V. de Silva. Zigzag persistence. *Found. Comput. Math.*, 10(4):367–405, 2010.
- [7] G. Carlsson, G. Singh, and A. Zomorodian. Computing multidimensional persistence. *J. Comput. Geom.*, 1(1):72–100, 2010.
- [8] G. Carlsson and A. Zomorodian. The theory of multidimensional persistence. *Discrete Comput. Geom.*, 42(1):71–93, 2009.
- [9] A. T. Casas. Distributing persistent homology via spectral sequences. arXiv:1907:05228, July 2019.
- [10] F. Chazal, V. de Silva, M. Glisse, and S. Oudot. The structure and stability of persistence modules. *Arxiv preprint arXiv:1207.3674*, 2012.
- [11] D. Cohen-Steiner, H. Edelsbrunner, and J. Harer. Stability of persistence diagrams. *Discrete Comput. Geom.*, 37(1):103–120, 2007.

- [12] J. Curry. *Sheaves, Cosheaves and Applications*. PhD thesis, University of Pennsylvania, 2014.
- [13] J. Curry, R. Ghrist, and V. Nanda. Discrete Morse theory for computing cellular sheaf cohomology. *Found. Comput. Math*, Dec. 2013.
- [14] H. Edelsbrunner and J. Harer. *Computational Topology: an Introduction*. American Mathematical Society, Providence, RI, 2010.
- [15] H. Edelsbrunner, D. Letscher, and A. Zomorodian. Topological persistence and simplification. *Discrete and Computational Geometry*, 28:511–533, 2002.
- [16] P. Gabriel. Unzerlegbare Darstellungen I. *Manuscripta Mathematica*, 6:71–103, 1972.
- [17] R. Ghrist. Barcodes: the persistent topology of data. *Bull. Amer. Math. Soc. (N.S.)*, 45(1):61–75, 2008.
- [18] G. Henselman and R. Ghrist. Matroid Filtrations and Computational Persistent Homology *Arxiv e-print arXiv:1606.00199*, 2016.
- [19] R. Lewis and D. Morozov. Parallel computation of persistent homology using the blowup complex. In *Proceedings of the 27th ACM Symposium on Parallelism in Algorithms and Architecture*, pages 323–331, 2015.
- [20] J. McCleary. *A User’s Guide to Spectral Sequences*. Cambridge Studies in Advanced Mathematics. Cambridge University Press, 2001.
- [21] P. Niyogi, S. Smale, and S. Weinberger. Finding the homology of submanifolds with high confidence from random samples. *Discrete & Computational Geometry*, 39:419–441, 2008.
- [22] N. Otter, M. Porter, U. Tillmann, P. Grindrod, and H. Harrington. A roadmap for the computation of persistent homology. *EPJ Data Science*, 6(17), 2017.
- [23] A. Shepard. *A Cellular Description of the Derived Category of a Stratified Space*. PhD thesis, Brown University, May 1985.
- [24] H. R. Yoon. *Cellular Sheaves and Cosheaves for Distributed Topological Data Analysis*. PhD thesis, University of Pennsylvania, 2018.
- [25] A. Zomorodian and G. Carlsson. Computing persistent homology. *Discrete Comput. Geom.*, 33(2):249–274, 2005.