Persistent Homology of Zig-Zag Families of Filtrations

Emily Lauren Saunders

Submitted in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy
under the Executive Committee
of the Graduate School of Arts and Sciences

COLUMBIA UNIVERSITY

2025

Abstract

Persistent Homology of Zig-Zag Families of Filtrations

Emily Lauren Saunders

This thesis introduces a new variant of multi-parameter persistent homology where one parameter is the standard $\mathbb{R}$, and the other is the zig-zag poset $\mathcal{Z}$. This construction is motivated by the need for a general scale parameter in various applications of zig-zag persistent homology. We begin by outlining the fundamental construction and defining an analog of the interleaving distance for $\mathcal{Z} \times \mathbb{R}$ persistence modules. We then establish stability of the metric with respect to the Gromov-Hausdorff distance and discuss convergence and stability under the topological bootstrapping sampling regime. Finally, we will discuss possibilities for a useful invariant on the space of $\mathcal{Z} \times \mathbb{R}$ persistence modules and explore applications of density sensitive bifiltrations to zig-zag sequences of point clouds.

Table of Contents

List of Figures

Acknowledgements

I would like to thank my advisor Andrew Blumberg for his unwavering support and guidance, as well as Julien Dubedat, Francesco Lin, Michael Lesnick and Elizabeth Munch for serving on my dissertation committee. I would also like to thank my partner Todd Morrill, my close friend Aaron Slipper, and my beautiful dog Ziggy for their invaluable emotional support.

Introduction

Topological data analysis (TDA) is an area of study interested in techniques for finding notable features in data. One such example that has been widely studied is clustering, which examines whether data points naturally form groups. Beyond clusters, one may be interested in features of a different form, for instance tendrils or holes in the data. For finding and understanding features of these forms, in particular “holes", a method known as persistent homology has proven to be successful.

Given a dataset of finitely many points, we “fatten up" each point by a factor of a scale parameter r to obtain an object that has potentially nontrivial homology, hopefully reflecting features of the underlying sample space. However the issue arises of choosing the correct scale parameter where it might be unknown or nonexistent. For instance, it may be the case that different features emerge at different scale parameters, as in the example shown in figure 1. The key insight of persistent homology is to analyze the homology of this space for all choices of r, along with maps that carry the data of how a feature persists as r varies.

In broad terms, standard persistent homology as applied to data analysis typically works as follows: Beginning with a data set X, we construct a family of simplicial complexes or topological spaces $\{S_r(X)\}_{r\in\mathbb{R}}$ dependent on a scale parameter r such that for $r' < r$, $S_{r'}(X)$ includes into $S_r(X)$ (we call this a *filtration*). We then take i-th homology of these simplicial complexes with coefficients in a field to obtain a *persistence module*, a family of k-vector spaces $\{PH^i_r(X)\}_{r\in\mathbb{R}}$ equipped with maps $PH^i_{r'}(X) \to PH^i_r(X)$ for $r' < r$. A structure theorem allows us to uniquely decompose these modules into a direct sum of standard interval

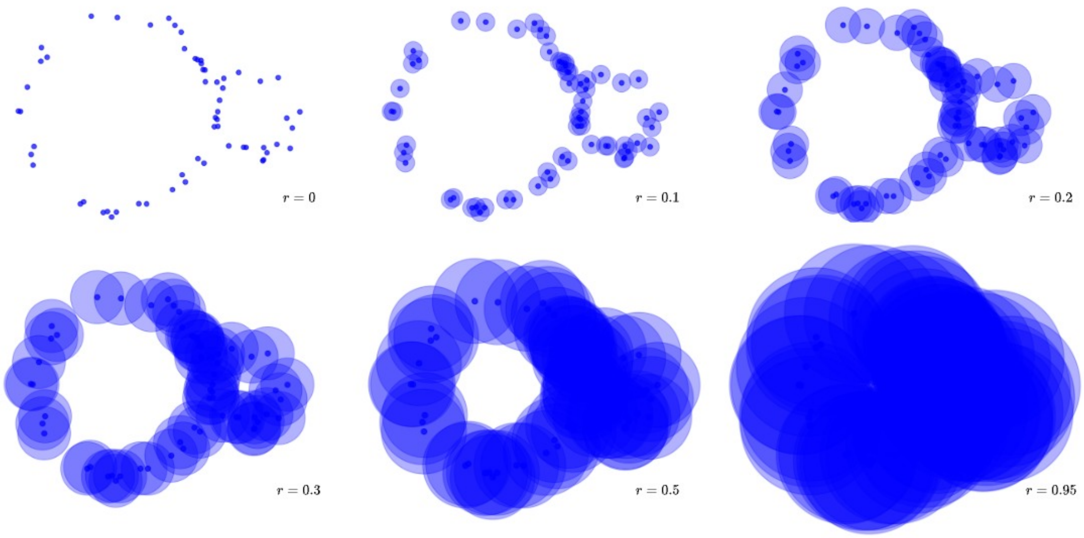


Figure 1: Union of balls filtration for various scale parameters r

persistence modules $\bigoplus_{[a,b)\in\mathcal{I}} k_{[a,b)}$[1], $a, b \in \mathbb{R} \cup \{\pm\infty\}$, where $k_{[a,b)}$ is defined to be

$$k_{[a,b)}(r) = \begin{cases} k & r \in [a,b) \\ 0 & \text{else} \end{cases} \qquad k_{[a,b)}(r \to r') = \begin{cases} id_k & r, r' \in [a,b) \\ 0 & \text{else} \end{cases}$$

This allows us to represent the data of the persistence module as the multi-set of intervals $\mathcal{I}$ that indexes this decomposition. We refer to this as a *barcode*, and can represent it visually as in figure 2.

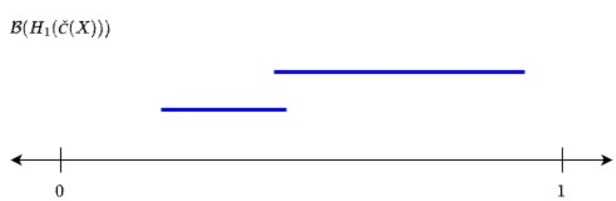


Figure 2: Barcode for H_1 of the example shown in figure 1

A single interval $[a, b)$ in the barcode represents a feature detected by the i-th homology that emerges at scale parameter a, persists until scale parameter b, and then dies. Longer

[1] intervals may also be of the form $[a, b], (a, b]$ or (a, b)

intervals represent stable features, whereas short intervals are typically considered to be noise.

The space of barcodes admits a pseudo-metric $d_{\mathcal{B}}$ called the *bottleneck distance*. This metric measures the distance between two barcodes by finding the best matching of intervals and taking the largest unmatched interval length as the distance. We will make this precise in the background section, but for now figure 3 illustrates the concept.

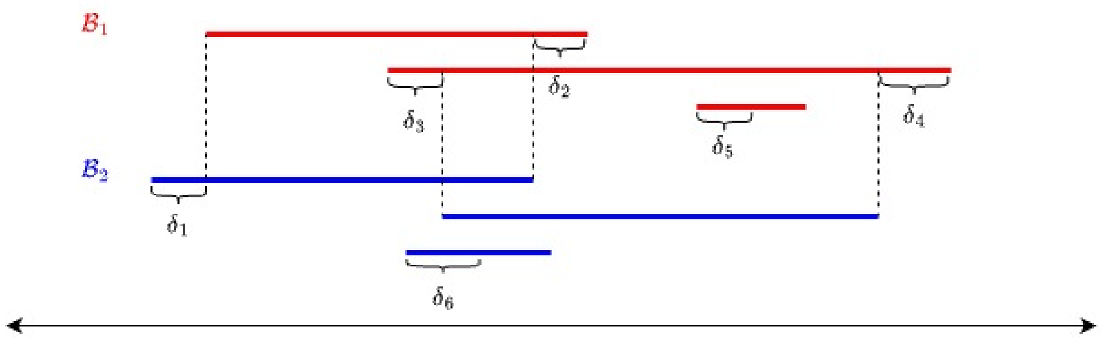


Figure 3: Optimal matching of $\mathcal{B}_1, \mathcal{B}_2$ with distance $\max_i \delta_i$

While the bottleneck distance on the space of persistence modules relies on the structure theorem, we can formulate an equivalent metric on the space of persistence modules known as the interleaving distance. We will expand on the interleaving distance and the isometry theorem which proves the equivalence of the two metrics in the background section.

In the above setting, we index our filtration over the poset category $\mathbb{R}$, which has objects $r \in \mathbb{R}$ and maps $r \to r'$ for $r < r'$. Work has been done to expand this methodology to replace $\mathbb{R}$ with an indexing category of a different shape. This arises naturally in many settings, for instance data where a second scale parameter is natural, or where we may require backwards maps as well as forward maps, such as $S_r(X) \leftarrow S_{r'}(X) \to S_{r''}(X)$. These generalizations are known as *multi-parameter persistence* and *zig-zag persistence* respectively. While certain results from standard persistence generalize easily to each of these settings, others become more complex. We will explore some of these results further in the background section.

The purpose of this thesis is to introduce a version of multi-parameter persistent homology, where one parameter is the usual $\mathbb{R}$, and the other is the zig-zag category $\mathcal{Z}$, which admits both forward and backward arrows:

$$\dots \longleftarrow \bullet \longrightarrow \bullet \longleftarrow \bullet \longrightarrow \bullet \longleftarrow \bullet \longrightarrow \dots$$

Definition 0.0.1. A $\mathcal{Z} \times \mathbb{R}$ *persistence module* over a field k is a functor,

$$F : \mathcal{Z} \times \mathbb{R} \to \mathbf{Vect}_k$$

This construction is motivated by the need for an additional scale parameter in many of the applications of zig-zag persistent homology. In particular, this thesis will establish some foundational definitions and results for persistence modules over the $\mathcal{Z} \times \mathbb{R}$ poset.

Chapter 1 of this thesis will provide relevant background on the field of persistence modules over the posets $\mathbb{R}, \mathbb{R}^d$ and $\mathcal{Z}$, aiming to lay the groundwork for the results we will adapt to the $\mathcal{Z} \times \mathbb{R}$ setting and the challenges that arise. Chapter 2 will then present the main definitions and construct a distance on the space of $\mathcal{Z} \times \mathbb{R}$ persistence modules by Kan extending the functor along an embedding into multi-parameter space. In chapter 3 we will verify that this construction satisfies some expected properties such as compatibility with the standard metrics on $\mathbb{R}$ and $\mathcal{Z}$, stability with respect to the Gromov Hausdorff distance on compact measure spaces, and convergence of randomly sampled data to the $\mathcal{Z} \times \mathbb{R}$ persistence of the underlying space.

Chapter 4 will develop the theory of invariants on the space of $\mathcal{Z} \times \mathbb{R}$ persistence modules, taking inspiration from approaches to invariants in the $\mathbb{R}^n$ setting such as the rank invariant and the fibered barcode. The final chapter, chapter 5, considers modules over $\mathcal{Z} \times \mathcal{P}$ where $\mathcal{P}$ is a multi-parameter poset, in order to explore applications of density sensitive bifiltrations to zig-zag diagrams of point clouds.

Section A of the appendix discusses the finiteness assumptions encountered in persistent homology. Although these conditions are automatically met when applying persistent homology to finite datasets and are often taken for granted, specifying the precise degree of finiteness is essential for stating theorems in the general form found in the literature. I have found it helpful to compile all these conditions in one place for easy reference. Section B provides a concise introduction to the key category theory concepts used in this thesis such as poset categories and left Kan extensions. Section C provides background on quiver representations that did not fit naturally into the main background section but is important to the study of persistent homology nonetheless.

Chapter 1: Background

1.1 Persistent homology

Persistent homology is an important tool in topological data analysis (TDA) that allows us to analyze the topological features of sampled data and how these features evolve with respect to a scale parameter. A primary application of this method is gaining information about the shape of point clouds, i.e. finite data sets in an ambient metric space such as $\mathbb{R}^n$. In the case where these point clouds arise from sampling data from an underlying space, we may hope to glean information about the topological features of the space through analyzing the features of the sampled data. Finite data sets are discrete and therefore have trivial homology, but if we "fatten" the points by taking the union of closed balls centered at points in X, $\bigcup_{x \in X} B_r(x)$, we may see features emerge that reflect the overall shape of the data. It is clear that $\bigcup_{x \in X} B_r(x) \subset \bigcup_{x \in X} B_{r'}(x)$ for $r < r'$, so this induces a functor $\mathbb{R}_{\geq 0} \to Top$ which we then compose with H_k. In practice, instead of working directly with this "union of balls" filtration, we construct filtrations of simplicial complexes that we hope have equivalent, or at least similar topology. Furthermore, some of these filtrations allow us to generalize to settings where our data sits inside a metric space other than $\mathbb{R}^n$, for instance genetic data equipped with the Levenshtein distance.

1.1.1 Filtrations

Most of the definitions in this section of this thesis can be found in many places throughout the literature, but I found [1] and the background section in [2] to be particularly excellent expositions and will be referencing them for many of these common definitions.

Viewing posets as categories, we define a filtration as follows:

Definition 1.1.1 (Filtration)**.** A *filtration* F indexed by a poset category $\mathcal{P}$ is a functor $F : \mathcal{P} \to \mathcal{S}$, where $\mathcal{S}$ is either the category **Top** of topological spaces or **Simp** of simplicial complexes, such that all $F(p \to p')$ are inclusion maps.

Standard 1-parameter persistent homology works over totally ordered posets, typically $\mathbb{R}, \mathbb{R}_{\geq 0}, \mathbb{Z}, \mathbb{N}$ or the corresponding opposite posets. When working with $\mathbb{R}$ or $\mathbb{R}_{\geq 0}$ we will assume the filtration is *essentially discrete*[1], which roughly speaking means that the filtration changes at discrete points. This is a condition clearly satisfied in the examples we will consider throughout this thesis. Filtrations indexed over opposite posets can be identified with standard order filtrations via negative indexing. For instance, a filtration $F : \mathbb{R}^{op} \to \mathcal{C}$ can be identified with $F' : \mathbb{R} \to \mathcal{C}$, $F'(r) = F(-r)$. Furthermore, filtrations over $\mathbb{R}_{\geq 0}$ can be extended to filtrations over $\mathbb{R}$ by taking $F(r) = \phi$ for $r < 0$, and filtrations over $\mathbb{N}$ or $\mathbb{Z}$ can be extended to $\mathbb{R}_{\geq 0}$ and $\mathbb{R}$ respectively by taking $F(r) = F(\lfloor r \rfloor)$. Thus results stated for $\mathbb{R}$ indexed persistence modules and filtrations extend to $\mathbb{Z}, \mathbb{N}, \mathbb{R}_{\geq 0}$ and opposite posets as well.

Given a finite data set in a metric space (X, d) and a scale parameter $r \in \mathbb{R}$, there are many ways to generate a filtration that reflects the topological features of the underlying space. These methods vary in terms of computability, their applicability to metric spaces beyond point clouds in $\mathbb{R}^n$ and the accuracy with which they represent the shape of the data. We will define several of the most commonly used filtrations here and discuss their strengths and weaknesses.

Definition 1.1.2 (Sublevelset filtration)**.** Let X be a topological space, and $f : X \to \mathbb{R}$ a function. The *sublevelset filtration* is the $\mathbb{R}$ indexed filtration defined by,

$$S_f^{\downarrow}(r) = f^{-1}((-\infty, r])$$

Remark 1.1.1. The union of balls filtration $r \mapsto \bigcup_{x \in X} B_r(x)$ mentioned earlier is a specific example of the sublevelset filtration with the function $f : \mathbb{R}^n \to \mathbb{R}_{\geq 0}, f(y) = d(y, X)$. Thus

[1] Formally defined in appendix section A.

$S^{\downarrow}_f(r) = f^{-1}((-\infty, r])$ consists of all points distance $\leq r$ from a point in X.

In TDA this union of balls filtration is the primary motivating example and the one often pictured when introducing the concept of persistent homology, so it serves as a natural starting point for the discussion of filtrations. The following theorem tells us that for sufficiently many data points and sufficiently small radius, the union of balls has equivalent homology to the underlying space with high confidence.

Theorem 1.1.1 (Niyogi, Smale, Weinberger [3])**.** *Let $\mathcal{M}$ be a compact submanifold of $\mathbb{R}^N$ with condition number τ.*[2] *Let $X = \{x_1, ..., x_n\}$ be a set of n points drawn in i.i.d. fashion according to the uniform probability measure on $\mathcal{M}$, and define $U = \bigcup_{x\in X} B_\varepsilon(x)$ for $0 < \varepsilon < \tau/2$. Then for all*

$$n > \beta_1(\log(\beta_2) + \log(\frac{1}{\delta}))$$

U deformation retracts to $\mathcal{M}$ with probability $> 1 - \delta$. [3]

In order to compute the homology of these spaces we will need to work with simplicial complexes rather than topological spaces. The next filtration we will discuss, the Čech complex, is a natural choice as the nerve theorem guarantees that the resulting simplicial complex is homotopy equivalent to $\bigcup_{x\in X} B_r(x)$. Before defining the Čech complex, we will need the definition of the nerve of an collection of sets.

Definition 1.1.3 (Nerve)**.** Given a collection of sets U, the *nerve of* U, denoted $\mathcal{N}(U)$, is the simplicial complex that has a k-simplex for every collection $\{U_0, ..., U_k\} \subset U$ with nonempty intersection. Specifically, $\mathcal{N}(U)_k = \{[U_0, ..., U_k] : \bigcap_{i=1}^k U_i \neq \emptyset\}$

Definition 1.1.4 (Čech filtration)**.** For $X \subset \mathbb{R}^n$ a finite point cloud, the *Čech filtration* is defined to be the $\mathbb{R}_{\geq 0}$ indexed filtration $C_r(X) = \mathcal{N}(\{B_r(x)\}_{x\in X})$

[2] τ is defined to be the largest number such that the open normal bundle around $\mathcal{M}$ of radius τ is embedded in $\mathbb{R}^N$.

[3] β_1 and β_2 are constants depending on the volume of $\mathcal{M}$, the volume of a $\dim(\mathcal{M})$-ball, ε and τ. The explicit formula can be found in Proposition 3.2 of [3].

Example 1.1.1. Figure 1.1.1 shows the Čech complex for a collection of points for increasing values of r

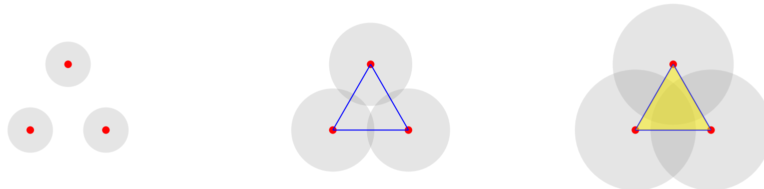


Figure 1.1: The Čech complex for 3 points with increasing values of r

The following foundational result known as the Nerve theorem is attributed to work of Borsuk, Weil and Leray in the 1940s and 50s.

Theorem 1.1.2 (Nerve Theorem)**.** *Let X be a paracompact Hausdorff*[4] *space and $\{U_\alpha\}_{\alpha\in A}$ be a cover of X. Assume that either,*

1. *$\{U_\alpha\}_{\alpha\in A}$ an open cover such that all finite intersections of elements of $\{U_\alpha\}_{\alpha\in A}$ are empty or contractible.*

2. *$X \subset \mathbb{R}^n$ with $\{U_\alpha\}_{\alpha\in A}$ a finite, closed, convex cover of X*

Then $\mathcal{N}(\{U_\alpha\}_{\alpha\in A})$ is homotopy equivalent to X.

This theorem tells us that the Čech complex results in a filtration that at each scale r is homotopy equivalent to $\bigcup_{x\in X} B_r(x)$. However from the perspective of persistent homology this is not enough to say that they are equivalent as filtrations, as this tells us nothing about the structure maps. A stronger version known as the *persistent nerve theorem* tells us that they are meaningfully topologically equivalent as filtrations, meaning that the resulting persistence modules obtained by taking homology of the filtrations are equivalent.

Before stating the persistent nerve theorem, we need to define a notion of equivalence for functors.

[4]The condition that the space X be *paracompact* and *Hausdorff* are mild conditions satisfied by all metrizable topological spaces.

Definition 1.1.5. Given a pair of functors $F, G : \mathcal{C} \to \mathbf{Top}$,

1. A natural transformation $\eta : F \Rightarrow G$ is an *object-wise homotopy equivalence* if $\eta_c : F(c) \to G(c)$ is a homotopy equivalence for all $c \in ob(\mathcal{C})$

2. F and G are *weakly equivalent* if there is a zig-zag sequence of object-wise homotopy equivalences connecting them:

$$F \leftarrow W_1 \to W_2 \leftarrow ... \to W_{n-1} \leftarrow W_n \to G$$

Remark 1.1.2. It can be shown that the diagram in part 2 of this definition can always be chosen to have only one intermediate map $F \leftarrow W \to G$.

Theorem 1.1.3 (Persistent Nerve Theorem)**.** *Suppose that a collection of functors* $\{U_\alpha : \mathcal{C} \to \textbf{\textit{Top}}\}_{\alpha \in A}$ *is a cover of a functor* $F : \mathcal{C} \to \textbf{\textit{Top}}$*, meaning that for every* $c \in \mathcal{C}$, $\{U_\alpha(c)\}_{\alpha \in A}$ *is a cover of* $F(c)$*, and the maps* $U_\alpha(c \to c')$ *are the restrictions of* $F(c \to c')$ *to* $U_\alpha(c)$*. Assume that for all* $c \in ob(\mathcal{C})$ *either,*

1. $F(c)$ *is paracompact and each* $\{U_\alpha(c)\}_{\alpha \in A}$ *is an open cover such that all finite intersections of elements in* $\{U_\alpha(c)\}_{\alpha \in A}$ *are empty or contractible.*

2. $F(c) \subset \mathbb{R}^n$ *with* $\{U_\alpha(c)\}_{\alpha \in A}$ *a finite, closed, convex cover of* X*.*

Then F *and* $\mathcal{N}(\{U_\alpha\}_{\alpha \in A})$ *are weakly equivalent.*

The case for open covers first appeared in [4] and [5], while a proof of the closed cover case, as well as a contained treatment of the persistent nerve theorem as a whole, can be found in [6].

For X a finite metric space, let $O(X)$ denote the filtration $O(X) : \mathbb{R}_{\geq 0} \to \mathbf{Top}$, $O(X)_r = \bigcup_{x \in X} B_r(x)$.

Corollary 1.1.1. $H_i C(X) \cong H_i O(X)$

While the persistent nerve theorem tells us that the homology of the Čech complex is equivalent to the homology of the union of balls filtration, due to its size it is computationally intractable. As such, for practical purposes, we often turn our attention to other filtrations of simplicial complexes that have equal or similar persistent homology.

The *alpha* filtration (otherwise known as the Delaunay filtration) is a smaller but equivalent filtration to $C(X)$. The key idea is to intersect the cover $\bigcup_{x \in X} B_r(x)$ with a cover of $\mathbb{R}^n$ known as the Voronoi diagram of X to obtain a new cover with fewer intersections.

Definition 1.1.6 (Voronoi cell/diagram)**.** For $X \subset \mathbb{R}^n$ finite, $x \in X$ the *Voronoi cell* of x is the set of all points in $\mathbb{R}^n$ that are closer to x than any other points in X. Specifically,

$$V(x) = \{y \in \mathbb{R}^n : d(x, y) \leq d(x', y) \forall x' \in X\}$$

The *Voronoi diagram* $V(X)$ is the closed convex cover of $\mathbb{R}^n$ consisting of all Voronoi cells of X.

See figure 1.2 for an example of the Voronoi diagram of points in $\mathbb{R}^2$.

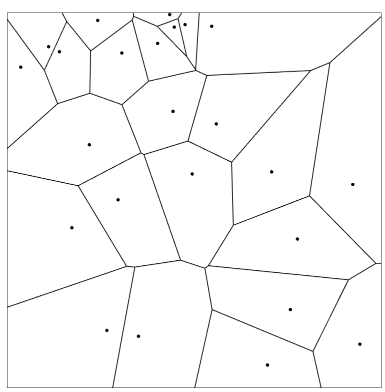


Figure 1.2: Voronoi diagram of 25 points in $\mathbb{R}^2$. Generated with `https://cfbrasz.github.io/Voronoi.html`

Definition 1.1.7 (α filtration)**.** The α *filtration* $\alpha(X) : \mathbb{R}_{\geq 0} \to \mathbf{Simp}$ is the nerve of the convex closed cover obtained by intersecting the closed union of balls filtration with the

Voronoi cells of X. Specifically,

$$\alpha_r(X) = \mathcal{N}(\{B_r(x) \cap V(x)\}_{x \in X})$$

The set $\{V(x) \cap B_r(x)\}_{x\in X}$ is a convex closed cover of $\bigcup_{x\in X} B_r(x)$, so by the persistent nerve theorem, $\alpha(X)$ is topologically equivalent to $\mathcal{O}(X)$ and thus $C(X)$. However, the cover $\{V(x) \cap B_r(x)\}_{x\in X}$ has fewer intersections than the cover $\{B_r(x)\}_{x\in X}$, resulting in a smaller complex when we take the nerve. We will be more specific about this later in this section when comparing the sizes of the filtrations.

Another approach to reducing the size of our filtration is the Witness complex introduced in [7]. The Witness complex uses a smaller set of points as "witnesses" to the existence of a simplex in order to reduce the size of the overall complex. Depending on the choice of points, this can provide a good representation of the underlying space with far fewer simplices.

Definition 1.1.8 (Witness complex/filtration)**.** For a finite data set $X \subset \mathbb{R}^n$ and a smaller set $L \subset X$ of landmark points , the *Witness complex* is defined inductively as follows:

- The vertex set $W(X;L)_0$ is L
- An edge $[l_{i_0}, l_{j_1}]$ is included in $W(X;L)_1$ if there exists a data point $x \in X$ such that l_{i_0}, l_{j_1} are the two closest points in L to x in L.
- A simplex $[l_{i_0}, ..., l_{i_k}]$ is included in $W(X;L)_k$ if all its faces are $k-1$ simplicies in $W(X;L)_{k-1}$ and there exists a data point x such that $l_{i_0}, ..., l_{i_k}$ are the $k+1$ closest points to x in L.

We say that the point x in the above definition *witnesses* the simplex σ. We obtain a $\mathbb{R}_{\geq 0}$-indexed filtration from this complex by setting $W(X;L,r)_k$ to be the subset of simplicies $\sigma \in W(X;L)_k$ such that all its vertices are distance $< r$ from the witness x.

Evidently this definition depends greatly on the choice of landmark set L. The leading methods for selecting the landmark set are random choice or the maxmin algorithm which greedily chooses the i-th landmark l_i to be the point $\max_{x\in X}\min\{d(x,l_1),...,d(x,l_{i-1})\}$. Empirically it appears that a ratio $|X|/|L| > 20$ provides a good approximation for data sampled from a 2-dimensional surface.

The final filtration we will discuss, the Vietoris-Rips complex, is the primary one used throughout this thesis. It's main advantage is that it does not require the data to be embedded in $\mathbb{R}^n$. Additionally, it can be simpler to compute for data embedded in high dimensions, as it depends only on pairwise distances.

Definition 1.1.9 (Vietoris-Rips filtration)**.** The Vietoris-Rips filtration of a metric space (X, d) is a $\mathbb{R}_{\geq 0}$-indexed filtration defined at scale r to be the simplicial complex with k-simplices given by,

$$R_r(X)^k = \{[x_0, ..., x_k] : d(x_i, x_j) \leq r\}$$

In other words, $R_r(X)$ has a k-simplex for every set of points $[x_0, ..., x_k]$ such that the points are all pairwise r-close. Figure 1.3 illustrates the difference between the Čech complex and the Vietoris-Rips complex for a simple collection of points in $\mathbb{R}^2$. Note we are comparing $C_r(X)$ with $R_{2r}(X)$.

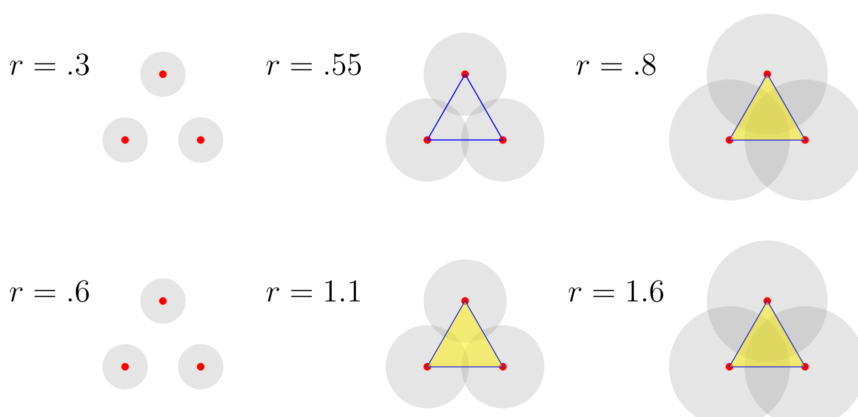


Figure 1.3: The Čech complex (top) compared to the Vietoris-Rips complex (bottom)

The Čech and Vietoris-Rips complexes are not equivalent, however they are closely related as illustrated by the following proposition,

Proposition 1.1.1. *For $X \subset (Z, d_Z)$,*

$$C_r(X) \subset R_{2r}(X) \subset C_{2r}(X)$$

Proof. Let $[x_0, ..., x_k]$ be a k simplex in $C_r(X)$. Then there exists a point y such that $y \in B_r(x_i)$ for all $i = 0, ..., k$. Thus by the triangle inequality, $d_Z(x_i, x_j) \leq d_Z(x_i, y) + d_Z(y, x_j) < 2r$, therefore $[x_0, ..., x_k]$ is a k simplex in $R_{2r}(X)$.

Now let $[x_0, ..., x_k] \in R_{2r}(X)^k$. We need to show that there exists a $y \in \bigcap_{i=1}^k B_{2r}(x_i)$. Take $y = x_0$. Since $d(x_0, x_i) < 2r$ for all i, this satisfies the condition. $\square$

A stronger version of this relationship for points in $\mathbb{R}^n$ appears in Theorem 2.5 of [8].

Computation We have mentioned trade-offs in size between these different filtrations in various settings. Now, we will provide a more detailed explanation of their relative computational advantages and disadvantages. Assuming that our filtration of simplicial complexes stabilizes to some $F_{\max}$ for sufficiently large R and only changes at a discrete set of scale parameters (conditions easily satisfied by applying any of the above constructions to finite point clouds) we can represent the data of F as $F_{\max}$ along with the birth time of each simplex in $F_{\max}$. As mentioned earlier, despite being perhaps the most intuitive, the Čech complex is not often used in practical settings where the Delaunay, Vietoris-Rips or Witness complexes are preferred.

The maximum Čech and Vietoris-Rips complexes have a simplex for every nonempty subset of X, therefore they both have size $2^{|X|} - 1$. By proposition 1.1.1 the Čech complex and Vietoris-Rips complex are comparable in size, however the Vietoris-Rips complex is easier to compute because it only depends on pairwise distances. When working with the α filtration we assume that the points $X \subset \mathbb{R}^n$ are in *spherical general position*, a condition stating that for all $1 \leq k < n$, no subset of $k + 3$ points in X lie on a k-dimensional sphere. This assumption can be met up to arbitrarily small perturbations of the data and dramatically improves the size of the complex by reducing the number of intersections of the

Voronoi cells. The maximum α complex for points X in $\mathbb{R}^n$ in spherical general position is $O(|X|^{\lceil n/2 \rceil})$[9].

1.1.2 Structure Theorem and the Barcode Metric

To obtain algebraic invariants from a filtration, we apply homology with coefficients in a field k. This results in a sequence of vector spaces, along with linear maps induced by inclusion, which together form what is known as a persistence module. We formalize this as follows:

Definition 1.1.10 (Persistence module)**.** A *persistence module* indexed by a poset $\mathcal{P}$ is a functor $F : \mathcal{P} \to \mathbf{Vect}_k$.

One of the most important results in the study of persistent homology is the structure theorem. The structure theorem tells us that a persistence module over a principle ideal domain (for example a persistence vector space) decomposes into a direct sum of interval persistence modules. This allows us to create useful visual representations of features in our data and define the bottleneck distance on persistence modules.

Theorem 1.1.4 (Structure Theorem for Persistence Modules)**.** *Let T be a totally ordered set, and $F : T \to \mathbf{Vect}_k$ a pointwise finite dimensional*[5] *persistence module. Then there exists a unique multiset of intervals $\mathcal{B}$ such that,*

$$F \cong \bigoplus_{I \in \mathcal{B}} k_I$$

where k_I is the T-indexed interval persistence module defined by,

$$k_I(t) = \begin{cases} k & t \in I \\ 0 & else \end{cases} \qquad k_I(t \to t') = \begin{cases} id & t, t' \in I \\ 0 & else \end{cases}$$

[5]defined in appendix section A.

In particular, any $\mathbb{R}$-indexed persistence vector space generated from a finite data set via taking homology of the filtrations defined above satisfies the structure theorem. When $T = \mathbb{N}$, the key idea of the proof is the isomorphism of categories between $Fun(\mathbb{N}, \mathbf{Vect}_k)$ and the category of $\mathbb{N}$-graded $k[t]$ modules. The correspondence is as follows: Given a $\mathbb{N}$-graded $k[t]$ module M, define $M(n) = M_n$ and $M(n \to n')$ by the action of $t^{n'-n}$. The remainder of the proof is then a graded version of the structure theorem for finitely generated modules over a PID. The first proof of this case can be found in [10]. The proof of the structure theorem for $\mathbb{R}$ originated in [11] and fully general case for any totally ordered poset T was proven in [12].

This theorem allows us to represent the data of the persistence module visually by graphing the multiset of intervals on the real line. Additionally, the space of barcodes $\mathcal{B}$ admits a distance known as the *bottleneck distance* which allows us to compare $\mathbb{R}$ persistence modules.

Definition 1.1.11 (Bottleneck distance)**.** Let $\mathcal{B}_1$, $\mathcal{B}_2$ be barcodes (i.e., multi-sets of intervals in $\mathbb{R}$). A matching $m : \mathcal{B}_1 \to \mathcal{B}_2$ is a bijection from a subset of $\mathcal{B}_1$ to a subset $\mathcal{B}_2$. We say m is a δ matching if

1. If $I \in \mathcal{B}_1 \cup \mathcal{B}_2$ is unmatched, then I has length $< 2\delta$
2. If $m([a, b)) = [c, d)$, then $|a - c|, |b - d| < \delta$ (the definition is analogous for intervals of the form $[a, b], (a, b)$ or $(a, b]$).

We define the distance $d_{\mathcal{B}}$ to be,

$$d_{\mathcal{B}}(\mathcal{B}_1, \mathcal{B}_2) = \inf\{\delta \geq 0 : \exists \text{ a } \delta \text{ matching between } \mathcal{B}_1 \text{ and } \mathcal{B}_2\}$$

Furthermore, for $F, G : \mathbb{R} \to \mathbf{Vect}_k$ pointwise finite dimensional persistence modules, we define $d_{\mathcal{B}}(F, G) = d_{\mathcal{B}}(\mathcal{B}_F, \mathcal{B}_G)$, for $\mathcal{B}_F, \mathcal{B}_G$ their respective barcode representations.

Remark 1.1.3. Technically, the bottleneck distance is an *extended pseudo-metric* on $\mathcal{B}$, since it can take on infinite values, for instance $d_{\mathcal{B}}(\phi, \{[0, \infty)\})$ and can be 0 for non-equal intervals

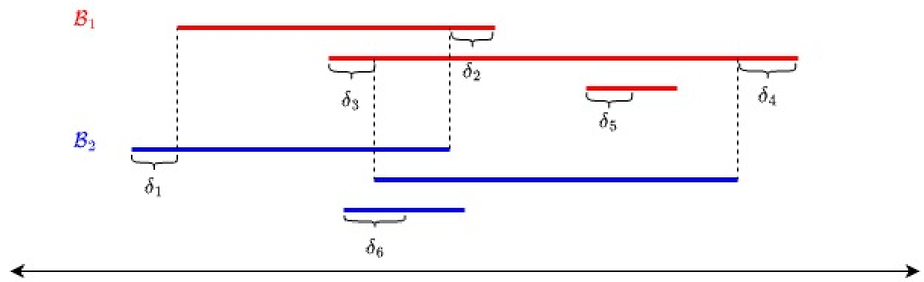


Figure 1.4: Optimal matching of $\mathcal{B}_1$ and $\mathcal{B}_2$ with distance $\max_i \delta_i$

such as $[a, b)$ and $(a, b]$. However it can be made into a proper metric by restricting to the space of finite multisets of bounded intervals of the form $[a, b)$, a reasonable assumption in practice.

The structure theorem plays a central role in the theory of persistent homology, prompting the question of how and when it can be generalized to other posets. The theory of quiver representations provides conclusive answers as to when such a structure theorem is possible for a general poset $\mathcal{P}$. We will provide a discussion of this in the appendix section C.

While the bottleneck distance relies on the existence of an interval decomposition, it is possible to construct an equivalent distance on this space known as the *interleaving distance*, which does not rely on a barcode representation. This is an important result, as it allows us to extend this distance to other settings in which the representation theory does not work out to be as nice and generalizes to categories other than $\mathbf{Vect}_k$ such as the category of spaces or simplicial complexes.

Definition 1.1.12 (Poset translation)**.** A translation of a poset $\mathcal{P}$ is a map $s : \mathcal{P} \to \mathcal{P}$ such that $s(p) \geq p$ for all $p \in \mathcal{P}$. Viewing our poset as a poset category, this is equivalent to a functor $s : \mathcal{P} \to \mathcal{P}$ that admits a natural transformation $\sigma : id_{\mathcal{P}} \Rightarrow s$.

Example 1.1.2. The poset $\mathbb{R}$ admits a translation s^δ for every $\delta \geq 0$ defined by $s^\delta(r) = r + \delta$.

Definition 1.1.13 (Interleaving distance)**.** Let $F, G : \mathbb{R} \to \mathcal{C}$, and for any $r \in \mathbb{R}, r > 0$ define $F^\delta : \mathbb{R} \to \mathcal{C}$ as $F^\delta = F \circ s^\delta$. Note that the natural transformation $id_{\mathbb{R}} \Rightarrow s^\delta$ induces a natural transformation $\sigma^\delta : F \Rightarrow F^\delta$. We say that F and G are δ-*interleaved* if there exists natural transformations $\alpha : F \Rightarrow G^\delta$ and $\beta : G \Rightarrow F^\delta$ such that the following diagrams commute,

$$\begin{tikzcd}
F \arrow[r, "\sigma_\delta"] \arrow[dr, "\alpha" near end] & F^\delta \arrow[r, "\sigma_\delta"] \arrow[dr, "\alpha" near start] & F^{2\delta} \\
G \arrow[r, "\sigma_\delta"] \arrow[ur, "\beta" near end] & G^\delta \arrow[r, "\sigma_\delta"] \arrow[ur, "\beta" near start] & G^{2\delta}
\end{tikzcd}$$

and we define the *interleaving distance* between F and G to be

$$d_I(F,G) = \inf\{\delta : F \text{ and } G \text{ are } \delta \text{ interleaved}\}$$

The isometry theorem now tells us that this distance is equivalent to the bottleneck distance $d_{\mathcal{B}}$ on the space of persistence modules.

Theorem 1.1.5 (Isometry theorem)**.** *Let $F, G : \mathbb{R} \to \textbf{Vect}_k$ be pointwise finite dimensional persistence modules, and let $\mathcal{B}_F, \mathcal{B}_G$ be their respective barcode representations. Then $d_{\mathcal{B}}(F,G) = d_I(F,G)$*

The proof of the $d_I(F,G) \le d_{\mathcal{B}}(F,G)$ direction is straightforward. First observe that If $F = k_{[a,b)}, G = k_{[c,d)}$ with $|a-c|, |b-d| < \delta$ then $\alpha : F \Rightarrow G$ defined by $\alpha_r = id_k$ if $r \in [a,b), r+\delta \in [c,d)$ and 0 else (and β defined analogously) define a δ interleaving. Additionally, $k_{[a,b)}$ for $|b-a| < 2\delta$ is δ interleaved with 0. Using these two facts we can construct a δ interleaving between $F = \bigoplus_{I \in \mathcal{B}_F} k_I$ and $G = \bigoplus_{J \in \mathcal{B}_G} k_J$ for $\mathcal{B}_F, \mathcal{B}_G$ δ matched by taking the direct sum of these interleavings for matched intervals and the 0 interleaving for unmatched intervals.

The converse direction is trickier. The original proof can be found in [13]. The induced matching theorem proved later in [14] provides a proof that constructs a matching of barcodes directly from a map $\alpha : F \Rightarrow G$. Additional important contributions include the work of Bjerkevik in [15], whose methods have proven instrumental in extending stability results to the multi-parameter setting (see, for example, [16]). Further generalizations are presented in [17], where Bjerkevik and Lesnick establish an approach that generalizes to ℓ^p-metrics.

The homotopy interleaving distance introduced in [18] is an example of how the interleaving distance can be generalized to other target categories.

Definition 1.1.14 (Homotopy interleaving distance)**.** For $\delta \geq 0$ and $F, G : \mathbb{R} \to \mathbf{Top}$ a filtration, we say F and G are δ *homotopy interleaved* if there exists $F', G' : \mathbb{R} \to \mathbf{Top}$ such that $F \cong F'$, $G \cong G'$ and F' and G' are δ-interleaved. Here $\cong$ denotes homotopy equivalence as persistence spaces (definition 1.1.5).

$$d_{HI}(F, G) = \inf\{\delta : F, G \text{ are } \delta \text{ interleaved}\}.$$

1.1.3 Stability results

We now review stability results that have been established for persistence modules. These results are fundamental to the theory of persistent homology, as they ensure that the resulting invariants change continuously under small perturbations of the input data.

Definition 1.1.15 (Gromov-Hausdorff distance)**.** For (M, d) a metric space, $X, Y \subset M$ nonempty subsets, the *Hausdorff distance* is defined to be,

$$d_H(X, Y) = \max\{\sup_{x \in X} d(x, Y), \sup_{y \in Y} d(y, X)\}$$

For general compact metric spaces $(X, d_X), (Y, d_Y)$, the *Gromov-Hausdorff distance* is defined to be,

$$d_{GH}((X, d_X), (Y, d_Y)) = \inf_{f, g, (Z, d_Z)} d_H(f(X), g(Y)$$

where (Z, d_Z) is a metric space and $f : X \to Z, g : Y \to Z$ are isometric embeddings.

Theorem 1.1.6. *Let (X, d_X), (Y, d_Y) be finite metric spaces. Then,*

$$d_{\mathcal{B}}(H_i(R(X)), H_i(R(Y))) \leq d_{GH}((X, d_X), (Y, d_Y))$$

A more recent version of this theorem was established in [18], which extends the result to the level of filtrations using the homotopy interleaving distance. Specifically, it shows that $d_{HI}(R(X), R(Y)) \leq d_{GH}((X, d_X), (Y, d_Y))$.

Additionally, we have a stability result for the sublevelset filtration with respect to the ℓ^∞ distance. This result first appeared in [19].

Theorem 1.1.7 (ℓ^∞ stability)**.** *For any topological space W, functions $f, g : W \to \mathbb{R}$, and k such that $H_k(S_f^\downarrow)$ and $H_k(S_g^\downarrow)$ are point-wise finite dimensional,*

$$d_{\mathcal{B}}(H_k(S_f^\downarrow), H_k(S_g^\downarrow)) \leq ||f - g||_\infty$$

Proof. The proof of this theorem is an immediate consequence of algebraic stability. If $||f - g||_\infty = \delta$, we obtain inclusions $S_f^\downarrow(r) \subset S_g^\downarrow(r + \delta)$ and $S_g^\downarrow(r) \subset S_f^\downarrow(r + \delta)$. These inclusions induce a δ-interleaving between $H_k(S_f^\downarrow)$ and $H_k(S_g^\downarrow)$. □

ℓ^∞ stability implies Hausdorff stability for the homology of the union of balls filtration via the identification described in remark 1.1.1. Hausdorff stability of the homology of the Čech and α-filtrations is an immediate consequence as they are weakly equivalent. The proof of Gromov-Hausdorff stability for homology of Vietoris-Rips complexes follows via the Kuratowski embedding, under which the Vietoris-Rips complex and the Čech complex of an embedded metric space are equal. Details of this proof can be found in the original paper [20].

Although stability results guarantee that persistent homology behaves continuously under small perturbations, real-world data often presents additional challenges. In many applications, we work with point clouds that are randomly sampled and may include noise or outliers. These issues can significantly affect the output. The Gromov-Hausdorff distance is particularly sensitive to outliers, as is the barcode distance on the resulting persistence modules.

There exist several approaches to this issue. One common strategy is to focus on only the densest points in your data set, as defined by a density function ρ. However, there are many possible choices for ρ, and it may not be clear from the data which to choose. Another approach is density sensitive bifiltrations which we will expand upon in the section on multi-parameter persistence.

Another approach introduced in [21] considers the distribution on the space of barcodes induced by taking persistent homology of randomly sampled point clouds from a metric measure space. We will state the main definition and result below.

Definition 1.1.16 (Gromov-Prokhorov distance)**.** Let (Z, d_Z) be a metric space and μ, ν measures on Z. The *Prokhorov distance* between μ and ν is defined to be,

$$d_{Pr}(\mu, \nu) = \sup_A \inf\{\delta \geq 0 : \mu(A) \leq \nu(B_\delta(A)) + \delta \text{ and } \nu(A) \leq \mu(B_\delta(A)) + \delta\}$$

Where A ranges over all closed subsets of Z. The *Gromov-Prokhorov distance* between two metric measure spaces $(X, d_X, \mu_X), (Y, d_Y, \mu_Y)$ is defined to be,

$$d_{GPr}((X, d_X, \mu_X), (Y, d_Y, \mu_Y)) = \inf_{\varphi, \psi} d_{Pr}(\varphi_*(\mu_X), \psi_*(\mu_Y))$$

where $\varphi : X \to Z$, $\psi : Y \to Z$ are isometric embeddings of X and Y into a metric space (Z, d_Z).

Definition 1.1.17 (Definition 1.1 in [21])**.** For a metric measure space (X, ∂_X, μ_X) and fixed $n, k \in \mathbb{N}$, the *n-sample kth persistent homology* is defined to be the probability distribution on the closure of the space of barcodes $\overline{\mathcal{B}}$ induced by the pushforward along the k-th persistent homology, PH_k from the product measure $\mu_X^{\otimes n}$ on X^n,

$$\Phi_k^n(X, \partial_X, \mu_X) = (PH_k)_*(\mu_X^{\otimes n})$$

Theorem 1.1.8 (Theorem 1.2 in [21])**.** *Let (X, ∂_X, μ_X) and (Y, ∂_Y, μ_Y) be compact metric* *measure spaces. Then we have the following inequality,*

$$d_{Pr}(\Phi_k^n(X, \partial_X, \mu_X), \Phi_k^n(Y, \partial_Y, \mu_Y)) \leq n d_{GPr}((X, \partial_X, \mu_X), (Y, \partial_Y, \mu_Y))$$

1.2 Multi-parameter Persistent Homology

In many cases, it is natural to consider filtrations that are indexed by posets other than $\mathbb{R}$, $\mathbb{Z}$ or $\mathbb{N}$. In the subsequent sections we will see a couple of examples of this, the first being *Multi-parameter persistence.* The primary sources we are referencing for this section are [2] and [22]

1.2.1 Definitions

Definition 1.2.1 (n-parameter filtration/multi-filtration)**.** An *n-parameter filtration or multi-filtration* is a filtration $F : \mathcal{P} \to \textbf{Top}$ where the poset $\mathcal{P}$ is the product poset of n totally ordered sets.

For the purposes of this section we will focus on $n = 2$, otherwise known as a *bifiltration.* Many examples of bifiltrations arise via real valued functions on the dataset in the following way:

Definition 1.2.2. Given a finite metric space X and a function $\gamma : X \to \mathbb{R}$, we define the superlevelset Vietoris-Rips bicomplex of X, γ to be,

$$R(\gamma)^{\uparrow} : \mathbb{R} \times \mathbb{R}^{op} \to \textbf{Simp}$$

$$R(\gamma)^{\uparrow}_{r,k} = R(\gamma^{-1}([k, \infty)))_r$$

and the sublevelset Vietoris-Rips bicomplex to be,

$$R(\gamma)^{\downarrow} : \mathbb{R} \times \mathbb{R} \to \textbf{Simp}$$

$$R(\gamma)^{\downarrow}_{r,k} = R(\gamma^{-1}((\infty, k]))_r$$

We can similarly define sub and superlevelset filtrations for the Čech, α and Witness complexes.

Example 1.2.1 (Eccentricity-Rips bifiltration)**.** Define $\gamma : X \to \mathbb{R}$ by,

$$\gamma(x) = \frac{1}{|X|} \sum_{y \in X} d(x, y)$$

the superlevelset filtration distinguishes clusters with high eccentricity, which can reveal tendril-like features in the data.

Example 1.2.2 (Time-varying data)**.** Suppose our metric space varies over time—specifically, for each time t, we are given a finite metric space X_t, with the assumption that $X_t \subset X_s$ whenever $t < s$. Under this assumption, we can construct a two-parameter filtration by applying a 1-parameter filtration to each X_t. The inclusion condition $X_t \subset X_s$ is necessary to ensure that the internal maps within the filtration are inclusion maps.

However, in many practical settings, the time-varying data may not satisfy this condition—that is, it may be the case that $X_t \not\subset X_s$. in which case we will need to consider other filtrations. This is discussed later in the thesis as an application of $\mathcal{Z} \times \mathbb{R}$ persistence, with $\mathcal{Z}$ the zig-zag filtration.

The following constructions demonstrate how we can generate multi-parameter filtrations that are density-sensitive. This is one approach to resolving the issue of 1-parameter persistence's high sensitivity to outliers.

Definition 1.2.3 (Multicover bifiltration [23, 24, 25])**.** The *unnormalized multicover bifiltration* $\mathcal{M}(X) : \mathbb{R}^{op}_{\geq 0} \times \mathbb{R}_{\geq 0} \to \textbf{Top}$ of a finite metric space $X \subset (Z, d)$ is defined to be

$$\mathcal{M}(X)_{r,s} = \{z \in Z : z \in \bigcap_{i=1,\dots,k} B_s(x_i) \text{ for some } \{x_i, \dots, x_k\} \subset X, r \leq k\}$$

In other words, $\mathcal{M}(X)_{r,s}$ consists of all points covered by at least r balls of radius s with centers in X. The *normalized multicover bifiltration* $\mathcal{M}^n(X) : \mathbb{R}_{\geq 0}^{op} \times \mathbb{R}_{\geq 0} \to \textbf{Top}$ is defined to be,

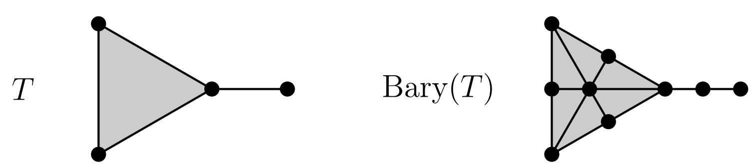
$$\mathcal{M}^n(X)_{r,s} = \{z \in Z : \nu_X(B_s(z)) \geq r\}$$

where ν_X is the pushforward of the uniform measure on X to Z. Specifically, $\nu_X(A) = |A \cap X|/|X|$. One can easily verify that $\mathcal{M}^u(X)_{r,s} = \mathcal{M}(X)_{r/|X|,s}$.

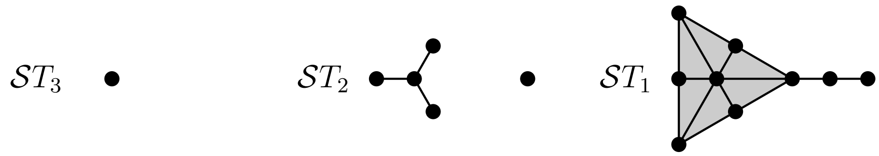
Definition 1.2.4 (Degree Rips/Čech filtration [26])**.** Let X be a finite metric space. The *unnormalized degree Rips* filtration $\mathcal{D}R(X) : \mathbb{R}_{\geq 0}^{op} \times \mathbb{R}_{\geq 0} \to \textbf{Simp}$ is defined on (r, s) to be the maximal subcomplex of $R(X)_s$ consisting of vertices x such that x has degree at least $r - 1$ in the 1-skeleton of $R(X)_s$. The *normalized degree rips filtration* is the $\mathbb{R}_{\geq 0}^{op} \times \mathbb{R}_{\geq 0}$ indexed filtration defined as $\mathcal{D}R^n(X)_{r,s} = \mathcal{D}R(X)_{r|X|,s}$. The *unnormalized and normalized degree Čech filtrations* $\mathcal{D}C(X), \mathcal{D}C^n(X)$ are defined analogously.

For the next filtration, we first need to define the barycentric subdivision of a simplicial complex.

Definition 1.2.5 (Barycentric subdivision)**.** For a given simplicial complex T, the barycentric subdivision $\mathrm{Bary}(T)$ is the simplicial complex with vertex set the simplicies of T and k-simplicies for all chains of simplicies $\sigma_0 \subset \sigma_1 \subset ... \subset \sigma_k$ in T. Visually, the vertex corresponding to a simplex $\sigma \in T$ can be identified with the point at the barycenter of σ. The geometric realization of $\mathrm{Bary}(T)$ is homeomorphic to the geometric realization of T.



Definition 1.2.6 (Subdivision bifiltration [23])**.** For a fixed simplicial complex T we can construct a $\mathbb{R}_{\geq 0}^{op}$ indexed filtration $\mathcal{S}T$ such that $\mathcal{S}T_r$ is the the sub-simplicial complex of $\mathrm{Bary}(T)$ using only vertices that correspond to simplicies in T of dimension at least $r - 1$, as illustrated in the following image:



Given a standard filtration of finite simplicial complexes $F : \mathcal{P} \to \mathbf{Simp}$ for $\mathcal{P}$ a totally ordered poset, the *subdivision bifiltration* $\mathcal{S}F : \mathbb{R}^{op}_{\geq 0} \times \mathcal{P} \to \mathbf{Simp}$ is defined to be $\mathcal{S}F_{r,s} = (\mathcal{S}F_s)_r$. For X a finite data set, the *normalized subdivision Rips bifiltration* $\mathcal{S}R^n(X)$ is defined to be $\mathcal{S}R^n(X)_{r,s} = \mathcal{S}R(X)_{r|X|,s}$ and the *normalized subdivision Čech bifiltration* is defined analogously.

Remark 1.2.1. By a multi-parameter version of the persistent nerve theorem found in [27], the subdivision Čech bifiltration is equivalent to the multi-cover bifiltration.

Defining a metric on a generalized persistence modules is essential for its practical applications. In particular, it enables numerical comparison of results obtained from applying a given method to different data samples. [28] provides a framework for approaching metrizing generalized persistence modules over a poset $\mathcal{P}$. Given either a sublinear projection or superlinear family of translations on our poset $\mathcal{P}$, we can use these to define a notion of interleaving.

Definition 1.2.7 (Sublinear projection/superlinear family [28]). Let $\mathcal{P}$ be a poset and $\mathbf{Trans}_{\mathcal{P}}$ the set of translations on $\mathcal{P}$, that is, the set of functors $T : \mathcal{P} \to \mathcal{P}$ such that $p \leq \Gamma(p)$.

- A sublinear projection of translations on $\mathcal{P}$ is a map $\omega : \mathbf{Trans}_{\mathcal{P}} \to [0, \infty]$ such that $\omega(Id_{\mathcal{P}}) = 0$ and $\omega(T_1 \circ T_2) \leq \omega(T_1) + \omega(T_2)$
- A superlinear family of translations is a map $\Omega : [0, \infty) \to \mathbf{Trans}_{\mathcal{P}}$ such that $T_{r_1} \circ T_{r_2}(p) \leq T_{r_1+r_2}(p)$ for all $p \in \mathcal{P}$

Given such a structure on our poset $\mathcal{P}$, we can define an interleaving distance between $F, G : \mathcal{P} \to \mathcal{C}$ by stating that F and G are ε interleaved if there are translations T, T'

and maps $\alpha : F \to T^*G$ and $\beta : G \to T'^*F$ with $\alpha \circ \beta = F \to (T' \circ T)^*F$ and $\beta \circ \alpha = G \to (T \circ T')^*G$ such that $\omega(T), \omega(T') \leq \varepsilon$ in the case of a sublinear projection, or $T = \Omega(r_1), T' = \Omega(r_2)$ for $r_1, r_2 \leq \varepsilon$ in the case of a superlinear family.

This framework gives rise to 2 natural possibilities for an interleaving on $\mathbb{R}^n$. The first is defined using the superlinear family of embeddings on $\mathbb{R}^n$ defined by $\delta \mapsto T^\delta$ where $T^\delta(x_1, ..., x_n) = (x_1, ..., x_n) + (\delta, ..., \delta)$. The second is defined using the superlinear projection on the family of translations $T^{\vec{u}} : \vec{x} \mapsto \vec{x} + \vec{u}$ defined by $\omega(T^{\vec{u}}) = ||\vec{u}||_2$. This choice does not meaningfully change the topology of the resulting space. Let d_ω and d_Ω denote the superlinear projection and sublinear family metrics on $\mathbb{R}^n$ persistence modules respectively. Since $||(\varepsilon, ..., \varepsilon)||_2 = \sqrt{n}\varepsilon$, we have $\frac{1}{\sqrt{n}} d_\omega(F, G) \leq d_\Omega(F, G)$. Conversely, if F and G are ε interleaved along maps $\alpha : F \Rightarrow G^{\vec{u}}, \beta : G \Rightarrow F^{\vec{v}}$ with $||\vec{u}||_2, ||\vec{v}||_2 < \varepsilon$, we may set $\varepsilon' = \max\{u_1, ..., u_n, v_1, ..., v_n\} \leq \varepsilon$ and compose α and β with the shift maps $\vec{x} + \vec{u} \to \vec{x} + (\varepsilon', ..., \varepsilon')$ and $\vec{x} + \vec{v} \to \vec{x} + (\varepsilon', ..., \varepsilon')$ to obtain ε' interleavings. Thus $d_\Omega(F, G) \leq d_\omega(F, G)$. While the sublinear family approach is more common in the literature as it is often easier to work with, we choose to use this more general definition later in this thesis as it allows us to consider the $\mathbb{R}$ and $\mathcal{Z}$ directions independently. However, it should be noted that most results presented here in the background section were proven with resect to the sublinear family perspective. For consistency with how these results appear in the literature we will define the multi-parameter interleaving distance with respect to the T^δ translations.

Definition 1.2.8 (Multi-parameter interleaving distance)**.** Let F, G be multi-parameter persistence modules, and define $F^\delta(x_1, ..., x_n) = (x_1 + \delta, ..., x_n + \delta)$. We say F and G are δ-interleaved if there exists morphism $\alpha : F \to G^\delta, \beta : G \to F^\delta$, such that the following diagram commutes,

$$\begin{tikzcd}
F \arrow[r] \arrow[dr, "\alpha"] & F^\delta \arrow[r] \arrow[dr, "\alpha"] & F^{2\delta} \\
G \arrow[r] \arrow[ur, "\beta"] & G^\delta \arrow[r] \arrow[ur, "\beta"] & G^{2\delta}
\end{tikzcd}$$

The interleaving distance $d_I(F, G)$ is defined to be $\inf\{\delta : F \text{ and } G \text{ are } \delta \text{ interleaved}\}$

The homotopy interleaving distance similarly generalizes to the multi-parameter setting as well,

Definition 1.2.9. Functors $F, G : \mathbb{R}^n \to \mathbf{Top}$ are *δ-homotopy interleaved* if there exists $F' \cong F$ and $G' \cong G$ such that F' and G' are δ-interleaved.

The following theorems of Blumberg and Lesnick [27] proves that the above density sensitive bifiltrations do satisfy stability results with respect to the Prokhorov distance

Definition 1.2.10. A metric space (Z, d_Z) is *good* if all finite intersections of open balls in X are empty or contractible.

Theorem 1.2.1 (Stability of multicover and subdivision bifiltrations, Theorem 1.6 in [27])**.**

1. *Let X, Y be nonempty finite subsets of a metric space Z, let μ_X, μ_Y be the uniform measures on X and Y respectively and ν_X, ν_Y their pushforwards onto Z. Then,*

$$d_I(\mathcal{M}^n(X), \mathcal{M}^n(Y)) \leq d_{Pr}(\nu_X, \nu_Y)$$

2. *If Z is good, then,*

$$d_{HI}(\mathcal{S}C^n(X), \mathcal{S}C^n(Y)) \leq d_{Pr}(\nu_X, \nu_Y)$$

3. *For any finite nonempty metric spaces X, Y,*

$$d_{HI}(\mathcal{S}R^n(X), \mathcal{S}R^n(Y) \leq d_{GPr}(\mu_X, \mu_Y)$$

Theorem 1.2.2 (Stability of degree bifiltrations, Theorem 1.7 in [27])**.** *Let γ^δ be the shift on $\mathbb{R}^{op}_{\geq 0} \times \mathbb{R}_{\geq 0}$ defined by $\gamma^\delta(x, y) = (x - \delta, 3y + \delta)$.*

1. *If X, Y are non-empty finite subsets of a good metric space, then $\mathcal{D}C^n(X)$ and $\mathcal{D}C^n(Y)$ are γ^δ-homotopy interleaved for all $\delta > d_{Pr}(\nu_X, \nu_Y)$.*

2. If X an Y are non-empty finite metric space, then $\mathcal{D}R^n(X)$ and $\mathcal{D}R^n(Y)$ are γ^δ-homotopy interleaved for all $\delta > d_{GPr}(\mu_X, \mu_Y)$.

Remark 1.2.2. There are subtleties in defining these interleavings with respect to the poset $\mathbb{R}^{op}_{\geq 0} \times \mathbb{R}_{\geq 0}$, as $s^\delta(x, y) = (x - \delta, y + \delta)$ is not a well defined poset translation $\mathbb{R}^{op}_{\geq 0} \times \mathbb{R}_{\geq 0} \to \mathbb{R}^{op}_{\geq 0} \times \mathbb{R}$. Specifically, for any $0 \leq x < \delta$, $s^\delta(x, y) = (x - \delta, y + \delta) \notin \mathbb{R}^{op}_{\geq 0} \times \mathbb{R}_{\geq 0}$. For a detailed treatment of interleavings in this setting, see Section 2.5 of [27].

1.2.2 Invariants

The theory of quiver representations allows us to precisely characterize which posets admit a structure theorem analogous to Theorem 1.1.4. A discussion of the relevant theorems can be found in section C of the appendix of this thesis. Unfortunately multi-parameter posets are what is called "wild" type, meaning that classifying the indecomposable representations is intractable.

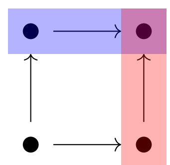
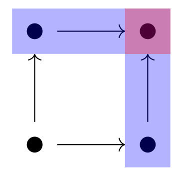
Despite this setback, we may still attempt to understand the outputs of multi-parameter persistence through weaker invariants. An invariant is considered *complete* if it fully determines the isomorphism class of the object it represents. For example, the barcode is a complete invariant, as two 1-parameter persistence modules have equivalent barcodes if and only if they are isomorphic. While multi-parameter persistence modules do not have a complete invariant, there are still several incomplete invariants that provide useful summaries of the structure of the module. We will outline some of them here.

Definition 1.2.11 (Rank invariant)**.** The rank invariant of a persistence module $F : \mathcal{P} \to \mathbf{Vect}_k$ assigns to each morphism $x \to y$ the rank $rk(F(x \to y))$.

In the case of 1 parameter persistence modules the rank invariant is equivalent to the barcode, and is therefore complete. However the following example demonstrates that it fails to be complete in multi-parameter setting.

Example 1.2.3. The following are two interval decomposable persistence modules with their interval decompositions illustrated beneath them. The left side module is the direct sum $k_I \oplus k_J$ for $I = \langle (1,0),(1,1) \rangle$ and $J = \langle (0,1),(1,1) \rangle$[6] and the right side module is $k_{I \cup J} \oplus k_{I \cap J}$

$$\begin{tikzcd} k \arrow[r, "{[1,0]^T}"] & k^2 \\ 0 \arrow[u] \arrow[r] & k \arrow[u, "{[0,1]^T}"'] \end{tikzcd} \qquad \begin{tikzcd} k \arrow[r, "{[1,0]^T}"] & k^2 \\ 0 \arrow[u] \arrow[r] & k \arrow[u, "{[1,0]^T}"'] \end{tikzcd}$$

One can easily verify that these modules have equivalent rank invariants, however they are not equivalent modules as their decomposition into indecomposable modules is different.

The identification of the rank invariant with the barcode in the one parameter case is the following: for any $x \leq y \in \mathbb{R}$, $rk(F(x \to y))$ is equal to the number of intervals in the barcode $\mathcal{B}$ representing F that contain both x and y. One might ask if we can do something analogous in the 2-parameter setting.

Definition 1.2.12 (Interval in $\mathcal{P}$)**.** An interval in a poset $\mathcal{P}$ is a nonempty subset $I \subset \mathcal{P}$ such that,

1. If $a, b \in I$, and $a \leq c \leq b$, then $c \in I$.

2. I is connected, meaning every for every $a, b \in I$, there is a sequence $a = p_1, ..., p_n = b$ with p_i, p_{i+1} comparable.

Can we find a collection of intervals for a multi-parameter persistence module such that the rank of $F(x \to y)$ is the number of intervals containing both x and y? The answer is in general no, as seen in the following example:

[6]In a general poset $\mathcal{P}$, $\langle p, p' \rangle$ denotes all elements q such that $p \leq q \leq p'$.

Example 1.2.4. This example is taken from [29]

$$\begin{tikzcd}
k \arrow[r, "id"] & k \arrow[r] & 0 \\
k \arrow[u, "id"] \arrow[r, "{[1,0]^T}"] & k^2 \arrow[u, "{[1,0]}"] \arrow[r, "{[1,1]}"] & k \arrow[u] \\
0 \arrow[u] \arrow[r] & k \arrow[u, "{[0,1]^T}"] \arrow[r, "id"] & k \arrow[u, "id"]
\end{tikzcd}$$

This two-parameter persistence module does not admit an interval decomposition. To see this, consider the following observations: there must be an interval that covers all elements in the upper-left square, another that covers the lower-right square, and a third that spans the middle horizontal line.

Note that the vector spaces at coordinates $(0,1)$ and $(2,1)$ both have rank 1, implying that they cannot be covered by more than 1 interval. This means that there must be a single interval covering all of the upper left square and lower right square. But then, this interval would necessarily also include $(1,0)$ and $(1,2)$. This implies that the map from $(1,0) \to (1,2)$ must have rank at least 1. However, in the given module, this map is zero, yielding a contradiction. Therefore, no such interval decomposition exists.

Still, while we cannot always express the rank of $F(x \to y)$ as the number of ordinary intervals containing both x and y, we can obtain such a description if we allow intervals to contribute with signs—some adding to the rank, others subtracting from it. This yields what is known as a *signed barcode*: a multiset of intervals, each with a positive or negative sign, such that the rank of $F(x \to y)$ equals the net number of signed intervals containing both x and y.

Definition 1.2.13 (Signed barcode)**.** A *signed barcode* for a persistence module $F : \mathcal{P} \to \mathbf{Vect}_k$ is a pair $(\mathcal{R}, \mathcal{S})$ of multisets of intervals in $\mathcal{P}$ such that for every $x \leq y \in \mathcal{P}$,

$$rk(F(x \to y)) = rk(\bigoplus_{I \in \mathcal{R}} k_I(x \to y)) - rk(\bigoplus_{J \in \mathcal{S}} k_J(x \to y))$$

In other words,

$$rk(F(x \to y)) = \#\{I \in \mathcal{R} : x, y \in I\} - \#\{J \in \mathcal{S} : x, y \in J\}$$

Such a decomposition always exists for a multi-parameter persistence module, though it is not unique. In figure 1.5 we can see 2 different decompositions of the persistence module in example 1.2.4:

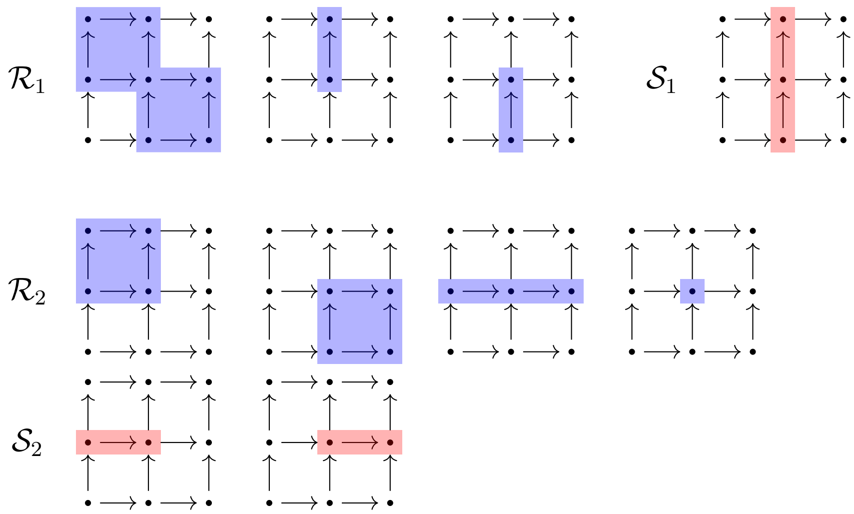


Figure 1.5: 2 examples of signed barcodes for example 1.2.4

However, by restricting to a specific type of intervals, we can define a unique minimal signed barcode.

Definition 1.2.14. Let $\mathcal{P}$ be a product poset of n totally ordered sets $T_1 \times ... \times T_n$. A *rectangle* in $\mathcal{P}$ is a subset $I = I_1 \times ... \times I_n$ such that each $I_j \subset T_j$ is an interval.

Theorem 1.2.3 ([29])**.** *Let $\mathcal{P}$ be a product poset of n totally ordered sets $T_1 \times ... \times T_n$ such that each $T_i \subseteq \mathbb{R}$ and let F be a finitely presented $\mathcal{P}$ persistence module. Then exists a minimal signed barcode consisting of rectangles, $(\mathcal{R}, \mathcal{S})$, meaning that for any other signed barcode $(\mathcal{R}', \mathcal{S}')$ consisting of rectangles we have $\mathcal{R} \subset \mathcal{R}'$ and $\mathcal{S} \subset \mathcal{S}'$*

Remark 1.2.3. While this rectangular signed barcode may be minimal among all other rectangular signed barcodes, it may not be the simplest overall. In the examples given in figure 1.5, the second signed barcode is the minimal rectangular barcode, but the first representation requires fewer intervals.

Another invariant that is equivalent to the rank invariant for a persistence module F indexed by $\mathbb{R}^n$ is the *fibered barcode*, which assigns to each positively sloped line $\ell \subset \mathbb{R}^n$ the barcode of the 1-parameter persistence module $F|_\ell$. This concept was first introduced in [30], although the term "fibered barcode" appeared in the literature somewhat later.

Definition 1.2.15 (Fibered barcode). Let $\mathcal{L}$ be the set of all lines that admit a parameterization $\vec{a}t + \vec{b}$ such that all coordinates of $\vec{a}$ are ≥ 0, and let $F : \mathbb{R}^n \to \mathbf{Vect}_k$ be p.f.d. The *fibered barcode* of F is the function,

$$\mathcal{F}_F : \mathcal{L} \to \mathcal{B}$$

$$\ell \mapsto \mathcal{B}_{F|_\ell}$$

To see that this invariant is equivalent to the rank invariant, observe that $rk(x \to y)$ can be determined by the number of intervals containing x and y in the barcode assigned to the line through the points x and y. Conversely the barcode is equivalent to the rank invariant for 1-parameter persistence, and the rank invariant of $F|_\ell$ is given by the restriction of the rank invariant of F to ℓ.

This invariant satisfies a nice internal and external stability result. The results appear in their modern form in [31], however they were proven earlier in a different form in [30] and [32]. For the following theorem, we'll restrict our attention to the subset $\mathcal{L}^\circ \subset \mathcal{L}$ consisting of all lines which admit a parametrization $\vec{a}t + \vec{b}$ with all entries of $\vec{a}$ strictly positive. By choosing such a parametrization with $||\vec{a}|| = 1$, all $a_i > 0$ and $||\vec{b}||$ to be the distance from ℓ to the origin, can identify $\mathcal{L}^\circ$ with a subset of $\mathbb{R}^{2n}$ and topologize it with the subset topology.

Theorem 1.2.4 (Internal stability). *For F a finitely presented persistence module, the fibered barcode map $\mathcal{F}_F : \mathcal{L}^\circ \to \mathcal{B}$ is continuous with respect to the topology induced on $\mathcal{B}$ by the bottleneck distance.*

Theorem 1.2.5 (External stability). *Let F, G be a pointwise finite-dimensional persistence modules, $\ell \in \mathcal{L}^\circ$. Then,*

$$\min_i a_i d_{\mathcal{B}}(\mathcal{B}_{F|_\ell}, \mathcal{B}_{G|_\ell}) \leq d_I(F, G)$$

and therefore the map $F \mapsto \mathcal{B}_{F|_\ell}$ is continuous with respect to the interleaving distance on the space of multi-parameter persistence modules and the bottleneck distance on the space of barcodes.

1.3 Zig-Zag Persistence

In this section, we review the theory of zigzag persistence. Originally introduced for analyzing level sets of real-valued functions in [33] and [34], zigzag persistence has since been developed and applied in various contexts. For a more comprehensive background, [35] and [36] offer valuable insights.

Zig-zag persistence replaces the poset $\mathbb{R}$ with a zig-zag poset $\mathcal{Z} \subset \mathbb{Z} \times \mathbb{Z}$ consisting of objects (i, j) with $i \in \mathbb{Z}, j \in \{-i, -i+1\}$

$$\begin{tikzcd}
\cdots & (0,1) & & (1,0) & & (2,-1) & \cdots \\
(-1,1) \arrow[ur] & & (0,0) \arrow[ul] \arrow[ur] & & (1,-1) \arrow[ul] \arrow[ur] & & (2,-2) \arrow[ul]
\end{tikzcd}$$

This formulation allows us to compare features of simplicial complexes that are related in important ways but do not naturally include into one another. I will briefly review some of these use cases here, but for a more comprehensive treatment see [35].

Remark 1.3.1. In some treatments of zig-zag persistence, including [35], the zig-zag poset is defined to have more general shape by allowing each arrow to be either forward or backward. For consistency with later definitions we will stick with this more restrictive version, which is sufficient for most applications, as we can obtain the general case by taking some arrows to be the identity.

Remark 1.3.2. We can equivalently formalize the zig-zag poset as subsets of $\mathbb{Z}$ of the form $\{i\}$ or $\{i, i+1\}$ with the inclusion order. This perspective is more natural for indexing $\mathcal{Z}$ filtrations and will often be used in this context.

1.3.1 Motivating Examples

The first, and most relevant use case for this thesis, is topological bootstrapping. Given a very large sample of points X, *topological bootstrapping* is the method of subsampling smaller sets $X_1, ..., X_n$ with replacement to understand the probability of various features within the underlying space. However, this requires that we are able to correlate features across subsamples in order to understand if a repeatedly observed feature represents the same feature sampled repeatedly or different features of the underlying space sampled with lower probability. Since these subsamples do not naturally include into one another, zig-zag persistence allows us to join these subsamples together via the maps $X_1 \to X_1 \cup X_2 \leftarrow X_2$. If the features coincide along these maps then they are the same, but if not they are different.

Another natural application of zigzag persistence is to compare witness complexes across different landmark sets. For distinct landmark sets L_1, L_2, there is not a natural inclusion between $W(X; L_1)$ and $W(X; L_2)$. However, we can construct a *witness bicomplex*

Definition 1.3.1 (Witness bicomplex [35])**.** For landmark sets L_1, L_2, the *witness bicomplex* $W(X; L_1, L_2)$ has simplices $\sigma \in L_1 \times L_2$ such that $\pi_{L_1}(\sigma)$ is a simplex in $W(X; L_1)$, $\pi_{L_2}(\sigma)$ is a simplex in $W(X; L_2)$ and there is an $x \in X$ that witnesses both $\pi_{L_1}(\sigma)$ and $\pi_{L_2}(\sigma)$.

Clearly $W(X; L_1, L_2)$ comes equipped with maps,

$$W(X; L_1) \leftarrow W(X; L_1, L_2) \to W(X; L_2)$$

This allows us to understand how the structure of a witness complex of a data set depends on the choice of landmark set.

The last example we'll mention in this section is comparing behavior of data across different density parameters. When examining the structure of our data we may consider only the densest points according to some density function ρ. It is often the case that ρ depends on a choice of parameter, for instance $\rho_r(x)$ may be defined as the number of points within distance r of x for some distance r. However, there is not necessarily a clear relationship between the densest points across different parameters. To compare the behavior of the resulting persistence modules across different choices of r we can examine the union and intersection complexes of the densest points for different parameters to understand how features coincide.

In each of these cases, to analyze the topological features of the data a choice of scale parameter is needed as in the original persistent homology setup. However, replacing $\mathbb{R}$ with $\mathscr{Z}$ means that we are no longer able to vary the scale parameter and are rather forced to make a choice. The purpose of this thesis is to introduce a version of multi-parameter persistence that extends the zig-zag poset with an additional $\mathbb{R}$ dimension in order to analyze zig-zag spaces at varying scale parameters.

1.3.2 Interval Decomposition

Unlike the multi-parameter persistence modules, zig-zag persistence modules have nice representation theory and therefore admit interval decompositions similar to the 1 parameter setting. By Gabriel's theorem (Theorem C.0.1 in the appendix), we see that because the underlying undirected graph for a finite zig-zag poset is of the type A_n, it is of finite type. Furthermore, [37] proved the decomposition result for the infinite case,

Theorem 1.3.1 (Theorem 1.7 in [37]). *Any non-zero pointwise finite dimensional zig-zag persistence module M decomposes into a direct sum of interval persistence modules.*

Thus zig-zag persistence objects admit a barcode invariant as well, where each interval represents an indecomposable interval module summand. Section 4 of [35] provides a stand-alone proof of the interval decomposition theorem for zig-zag persistence modules which gives rise to an algorithm for computing the decomposition.

Remark 1.3.3. In the regular one parameter persistence case, if an interval module k_I is detected as a submodule of M, then there necessarily exists an interval $J \supset I$ such that k_J is a component in the interval decomposition of M. In other words, detecting an interval submodule k_I corresponds to a feature that persists for at least the duration of the interval I. This is however not the case with zig-zag persistence, as illustrated in the following example sourced from [35],

$$\begin{tikzcd} k & k^2 \arrow[l] \arrow[r] & k \end{tikzcd}$$

$$\begin{tikzcd} (x,0) & (x,y) \arrow[l, mapsto] \arrow[r, mapsto] & (0,y) \end{tikzcd}$$

The interval decomposition is $K_{[0,1]} \oplus K_{[1,2]}$ however $K_{[0,2]}$ is a submodule. This demonstrates that in zigzag persistence, interval submodules may not correspond directly to the intervals in the decomposition. Therefore, care must be taken when interpreting detected features—one must consider the entire decomposition rather than relying on submodules alone.

1.3.3 Metric

The primary references for this section are [35], [36] and [38].

In section 1.2.1 we discussed how we may obtain an interleaving distance on the space of $\mathcal{P}$ indexed persistence modules given a sufficiently nice family of poset translations on $\mathcal{P}$ (see Definition 1.2.7). Unfortunately this method fails when we attempt to apply it to $\mathcal{Z}$, as $\mathcal{Z}$ does not admit any nontrivial translations. It is easy to see why by considering the simple example of a 3 element zig-zag as outlined in [38]:

$$\begin{tikzcd} A & & C \\ & B \arrow[ul] \arrow[ur] & \end{tikzcd}$$

If we attempt to define a translation T on this poset, we immediately see that the condition $p \leq T(p)$ forces $T(A) = A$ and $T(C) = C$. If $T(B) = A$ then we would have no choice for $T(B \to C)$, and likewise $T(B) = C$ fails, forcing $T(B) = B$.

To resolve this, we embed the space of $\mathcal{Z}$-modules into the space of $\mathbb{R}^2$-modules by Kan extending along the embedding $i : \mathcal{Z} \to \mathbb{R}^2$ as pictured in figure 1.6.

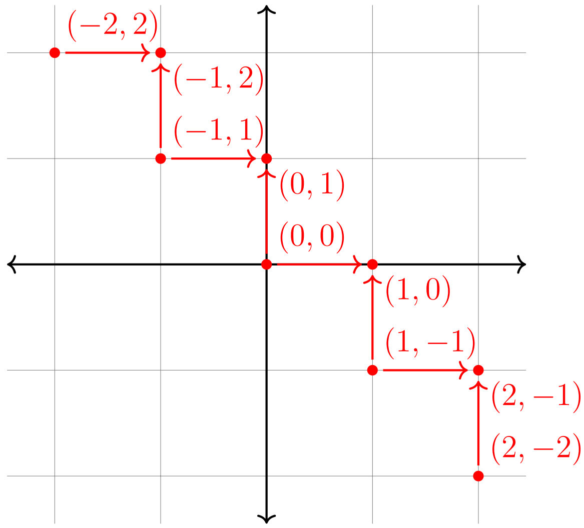


Figure 1.6: Embedding of $\mathcal{Z}$ into $\mathbb{R}^2$

For F a $\mathcal{Z}$ persistence module, let L_iF denote the left Kan extension of F along i (defined in Appendix B). For G a $\mathbb{R}^2$ indexed persistence module, define $i^*G = G \circ i : \mathcal{Z} \to \mathbf{Vect}_k$.

Lemma 1.3.1. *If $i : \mathcal{P} \to \mathcal{Q}$ is a fully faithful embedding of posets, then i^* is right adjoint to L_i and the unit of the adjunction is a natural isomorphism.*

Proof. Let $F : \mathcal{P} \to \mathcal{C}, G : \mathcal{Q} \to \mathcal{C}$. The data of a map $L_iF \Rightarrow G$ is equivalent to maps $L_iF(q) = \operatorname{colim}_{(i \downarrow q)} F \to G(q)$ that commute with the corresponding $q \to q'$ maps. In particular, this gives us maps

$$L_iF(i(p)) = \operatorname{colim}_{(i \downarrow i(p))} F = F(p) \to G(i(p)) = i^*G(p)$$

Note the identification $\operatorname{colim}_{(i \downarrow i(p))} F = F(p)$ follows from i being a fully faithful embedding of posets, as $i(p') \to i(p) \iff p' \to p$, therefore p is the terminal object in the diagram $(i \downarrow i(p))$. This gives us a map,

$$Hom_{Fun(\mathcal{Q},\mathcal{C})}(L_iF, G) \to Hom_{Fun(\mathcal{P},\mathcal{C})}(F, i^*G)$$

To obtain the inverse, observe that a map $F \Rightarrow i^*G$ determines a map $L_iF \Rightarrow G$ by taking the data of the maps $L_iF(q) = \mathrm{colim}_{(i\downarrow q)}F \to G(q)$ to be determined by the maps $F(p) \to G(i(p)) \to G(q)$ for all $i(p) \to q$ in the diagram $(i \downarrow q)$. It is easy to verify that this assembles into a natural transformation $L_iF \Rightarrow G$ and the resulting map,

$$Hom_{Fun(\mathcal{P},\mathcal{C})}(F, i^*G) \to Hom_{Fun(\mathcal{Q},\mathcal{C})}(L_iF, G)$$

is inverse to the one defined above.

The unit of this adjunction is the natural transformation $F \Rightarrow L_iF \circ i = i^*L_iF$ given by the universal property of the left Kan extension[7]. Proposition B.0.1 in the appendix proves an explicit formula for this natural transformation as being the data of the canonical inclusions $F(p) \to \mathrm{colim}_{(i\downarrow i(p))}F$ which is the identity under the identification $\mathrm{colim}_{(i\downarrow i(p))}F \cong F(p)$ discussed above. $\square$

A version of the following result appears in Proposition 4.3 of [36]:

Corollary 1.3.1. *The map*

$$L_i : Fun(\mathcal{Z}, \boldsymbol{Vect_k}) \to Fun(\mathbb{R}^2, \boldsymbol{Vect_k})$$

is a fully faithful embedding.

Proof. One can easily verify that the embedding $i : \mathcal{Z} \to \mathbb{R}^2$ is fully faithful. Therefore, from the above adjunction and the fact that $i^*L_iF \cong F$ we have,

$$Hom_{Fun(\mathcal{Z},\mathbf{Vect}_k)}(F, G) \cong Hom_{Fun(\mathcal{Z},\mathbf{Vect}_k)}(F, i^*L_iG) \cong Hom_{Fun(\mathbb{R}^2,\mathbf{Vect}_k)}(L_iF, L_iG)$$

$\square$

Thus we can consider the space of $\mathcal{Z}$-modules as a subspace of $\mathbb{R}^2$-modules and restrict

[7]See definition B.0.3 in the appendix.

the metric defined above in the multi-parameter persistence section to $Fun(\mathcal{Z}, \mathbf{Vect}_k)$. This embedding was first explored in [33] and [34] in the context of interlevel sets of Morse functions, and then expanded upon in [36]. A more general formulation, extending the approach to arbitrary posets $\mathcal{P}$ via the language of sheaves, is presented in [38].

Since the space of zigzag persistence modules admits a barcode decomposition, we can also define a bottleneck distance on this space. In [36], it is shown that this distance satisfies an analogue of the isometry theorem. A refinement of this result, featuring improved constants, is provided in [15].

1.4 Summary

In summary, we have examined 3 theories of persistence so far: the original $\mathbb{R}$ persistence, multi-parameter persistence over $\mathbb{R}^n$, with special attention paid to the case of $\mathbb{R}^2$, and zigzag persistence over $\mathcal{Z}$. $\mathbb{R}$ persistence serves as the blueprint, with its main advantage being the barcode invariant which provides a clear and complete visual summary of the data of any $\mathbb{R}$-persistence module. This barcode allows us construct a distance on the space of $\mathbb{R}$ persistence modules, which was later proved equivalent to the interleaving distance, a purely algebraic construction that readily generalizes to other posets. The primary limitations of $\mathbb{R}$ persistence are its failure of robustness to outliers and its inability to compare spaces that do not naturally include into one another.

Multi-parameter persistence arises naturally in cases where an additional parameter is natural such as a time or density parameter. A distance is put on the space of multi-parameter persistence via interleaving along the family of poset shifts given by $(x_1, ..., x_n) \mapsto (x_1+\varepsilon, ..., x_n+\varepsilon)$. The addition of a density parameter makes our outputs robust as illustrated in Theorem 1.2.1 (Theorem 1.6 in [27]). Unfortunately the poset $\mathbb{R}^n$ does not have nice representation theory meaning we are not able to find an invariant that is both simple and complete like the barcode for $\mathbb{R}$ persistence. However the rank invariant and the fibered barcode offer useful summaries of the data.

Zig-zag persistence allows us to apply the tools of persistence to spaces that do not include into one another. The primary motivating example is the case of smaller subsamples from a larger dataset. Because the zig-zag poset does not admit a nice family of shifts we Kan extend the $\mathcal{Z}$ persistence modules along the embedding $i : \mathcal{Z} \to \mathbb{R}^2$ and interleave along the ε shifts in $\mathbb{R}^2$. Like $\mathbb{R}$, the poset $\mathcal{Z}$ does have nice representation theory, meaning every $\mathcal{Z}$-module decomposes into a direct sum of interval modules. The largest drawback to $\mathcal{Z}$ persistence is the need to choose a scale parameter to generate a simplicial complex from the data. This is the problem this thesis will address, by constructing a theory of $\mathcal{Z} \times \mathbb{R}$ persistence.

Chapter 2: Definitions

We will now present the main definitions of this thesis.

Definition 2.0.1 ($\mathcal{Z} \times \mathbb{R}$ persistence object)**.** A $\mathcal{Z} \times \mathbb{R}$ *persistence object* in a category $\mathcal{C}$ is a functor,

$$F : \mathcal{Z} \times \mathbb{R} \to \mathcal{C}$$

or, equivalently by adjunction, a functor,

$$F : \mathcal{Z} \to Fun(\mathbb{R}, \mathcal{C})$$

The most common way in which 2 parameter zig-zag persistence modules arise is via applying filtrations to a zigzag sequence of pairwise unions or intersections of finite subsets of a metric space, and then taking homology with coefficients in a field. In particular, given a sequence of finite subsets $X_1, ..., X_n$ of a larger metric space (X, d), we may form a 2 parameter zig-zag persistence object by applying a 1 parameter filtration methods to the union zig-zag:

$$X_1 \to X_1 \cup X_2 \leftarrow X_2 \to ... \leftarrow X_{n-1} \to X_{n-1} \cup X_n \leftarrow X_n$$

or the intersection zig-zag:

$$X_1 \leftarrow X_1 \cap X_2 \to X_2 \leftarrow ... \to X_n \leftarrow X_{n-1} \cap X_n \to X_n$$

The examples outlined in Section 1.3.1 such as topological bootstrapping, comparing different density parameters and the Witness bicomplex would all benefit from the use of an additional scale parameter. Furthermore, time series data is a natural application of $\mathcal{Z} \times \mathbb{R}$ persistence.

Example 2.0.1 (Time series data)**.** A time series is a family of point clouds $(X_t)_{t\geq 0}$ that evolves with respect to time parameter t. If X_t does not naturally include into $X_{t'}$ for $t \leq t'$, we cannot apply 2 parameter persistence to this data as the time parameter does not define a filtration. Instead, we can construct a zig-zag sequence to compare data across different time steps. Given a sequence of times $t_0, t_1, ...$ we form a zig-zag using unions or intersections and then apply a one-parameter filtration to obtain a $\mathcal{Z} \times \mathbb{R}$ persistence module.

In some cases there may be a natural choice of distance in the zig-zag direction. Thus one may want to consider a weighted version of this construction that reflects the added structure of the edge weights.

Definition 2.0.2 (Weighted (2 parameter) zig-zag persistence object)**.** A *weighted zig-zag persistence object* in a category $\mathcal{C}$ is a functor

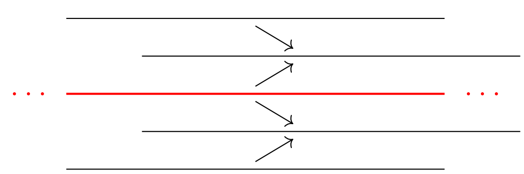
$$F : \mathcal{Z} \to \mathcal{C}$$

and a function $w : hom(\mathcal{Z}) \to \mathbb{R}_{\geq 0}$ assigning weights to edges. A 2-parameter weighted zig zag persistent object in $\mathcal{C}$ is therefore a functor $F : \mathcal{Z} \to Fun(\mathbb{R}, \mathcal{C})$ equipped with a weight function $w : hom(\mathcal{Z}) \to \mathbb{R}_{\geq 0}$.

For the purposes of this thesis we will be primarily focusing on the unweighted case, which can be thought of as a weighted object with uniform weights 1. Many of the theorems and definitions will easily generalize to the case where the weight function is uniform, i.e. $w(z_1 \to z_2) = c$.

Our first task will be to define a metric on the space of $\mathcal{Z} \times \mathbb{R}$ persistence modules. [28] provides a method for generalizing the interleaving distance to modules over posets that admit a sub or super linear family of translations. $\mathcal{Z} \times \mathbb{R}$ admits such a family in the $\mathbb{R}$ direction, however these translations are necessarily trivial in the $\mathcal{Z}$ direction, which will not fully capture the structure of the $\mathcal{Z} \times \mathbb{R}$ modules. If we are not able to shift in the $\mathcal{Z}$ parameter direction, the metric will not be able to detect what happens to features along maps in the zig-zag direction. For instance, define M in the following way,

$$M(z,t) = \begin{cases} \mathbb{R} & z = (0,0) \\ 0 & \text{else} \end{cases} \qquad M((z,t) \to (z',t')) = \begin{cases} id & z = z' = (0,0) \\ 0 & \text{else} \end{cases}$$



We might expect M to be distance 1 from the 0 module because every point $(z,t) \in \mathcal{Z} \times \mathbb{R}$ such that $M(z,t) \neq 0$ is one zig zag step away from 0. However, if we are only able to move in the $\mathbb{R}$ direction, this module is infinite distance from 0. To address this issue, we'll generalize the methods introduced in [38] and [36] to $\mathcal{Z} \times \mathbb{R}$.

Recall that we can fully and faithfully embed $\mathcal{Z}$ into the poset category $\mathbb{R} \times \mathbb{R}$ as in figure 1.6. This can be extended to a full and faithful embedding of $\mathcal{Z} \times \mathbb{R}$ by adding a 3rd dimension as in figure 2.1.

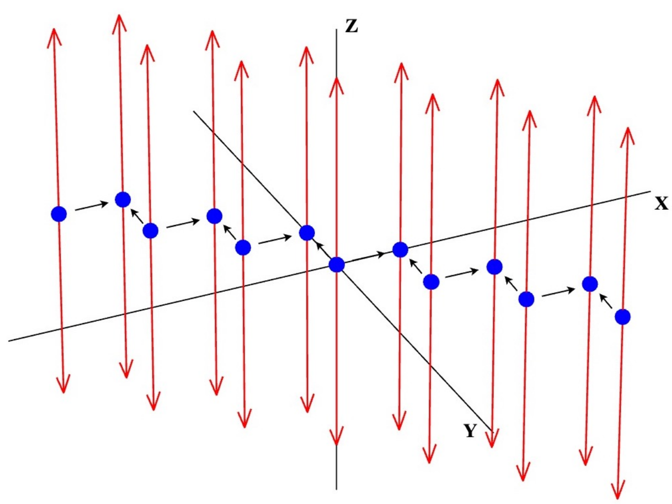


Figure 2.1: Embedding of $\mathcal{Z} \times \mathbb{R}$ to $\mathbb{R}^3$

In the weighted case, we modify the embedding $i : \mathcal{Z} \to \mathbb{R}^2$ such that $||i(z) - i(z')|| = w(z \to z')$.

To extend a functor $F : \mathcal{Z} \times \mathbb{R} \to \mathcal{C}$ to all of $\mathbb{R}^3$ we take the left Kan extension along this inclusion. This gives us an extended version of the original functor F that admits shifts that may be nontrivial in the zig-zag direction. Specifically, we consider the family of poset translations on $\mathbb{R}^3$ given by $T^{\vec{u}}(x, y, z) = (x, y, z) + \vec{u}$ for $\vec{u} = (u_1, u_2, u_3)$ with $u_i \geq 0$.

Definition 2.0.3. Let F and G be $\mathcal{Z}\times\mathbb{R}$ persistence modules. For vectors $\vec{u} = (u_1, u_2, u_3), \vec{v} = (v_1, v_2, v_3)$ such that $u_1, u_2, u_3, v_1, v_2, v_3 \geq 0$ we say F and G are $\vec{u}, \vec{v}$-interleaved if there exists morphisms,

$$\alpha : L_\iota F \to L_\iota G^{\vec{u}} \qquad \beta : L_\iota G \to L_\iota F^{\vec{v}}$$

such that the following diagrams commute

$$\begin{tikzcd}
L_\iota F \arrow[r] \arrow[dr, "\alpha" near end] & L_\iota F^{\vec{v}} \arrow[r] \arrow[dr, "\alpha" near end] & L_\iota F^{\vec{u}+\vec{v}} \\
L_\iota G \arrow[r] \arrow[ur, "\beta" near end] & L_\iota G^{\vec{u}} \arrow[r] \arrow[ur, "\beta" near end] & L_\iota G^{\vec{u}+\vec{v}}
\end{tikzcd}$$

We say F and G are $\vec{u}$-interleaved if they are $\vec{u}, \vec{u}$-interleaved. Furthermore, we define the distance $d_{\mathcal{Z}\times\mathbb{R}}$ on the space of $\mathcal{Z} \times \mathbb{R}$ persistence modules to be,

$$d_{\mathcal{Z}\times\mathbb{R}}(F, G) = \inf\{\varepsilon \text{ such that } F \text{ and } G \text{ are } \vec{u}, \vec{v}\text{-interleaved with } ||\vec{u}||_2, ||\vec{v}||_2 \leq \varepsilon\}$$

We say F is ε-trivial if $d_{\mathcal{Z}\times\mathbb{R}}(F, 0) < \varepsilon$.

As stated earlier, the $\mathcal{Z} \times \mathbb{R}$ poset does have a linear family of shifts in the $\mathbb{R}$ direction. It will be useful later on to prove some lemmas demonstrating how these shifts interact with the Kan extension to $\mathbb{R}^3$. For $F \in Fun(\mathcal{Z} \times \mathbb{R}, \mathcal{C})$, let F_ε denote the $\mathbb{R}$-shifted functor $F_\varepsilon(z, t) = F(z, t + \varepsilon)$.

Lemma 2.0.1. *Let σ_ε denote the shift map $F \to F_\varepsilon$. Then $L_\iota F_\varepsilon \cong L_\iota F^{(0,0,\varepsilon)}$ and $L_\iota(\sigma_\varepsilon)$:*

$L_\iota F \to L_\iota F_\varepsilon \cong L_\iota F^{(0,0,\varepsilon)}$ is the regular shift function induced by the inclusion of the comma category $(\iota \downarrow (a,b,c)) \hookrightarrow (\iota \downarrow (a,b,c) + (0,0,\varepsilon))$.

Proof. First we want to show $L_\iota F_\varepsilon \cong L_\iota F^{(0,0,\varepsilon)}$.

$$\begin{aligned} L_\iota F_\varepsilon(a,b,c) &= \mathrm{colim}_{\iota(z,t)\to(a,b,c)} F_\varepsilon(z,t) \\ &= \mathrm{colim}_{\iota(z,t)\to(a,b,c)} F(z,t+\varepsilon) \\ &= \mathrm{colim}_{\iota(z,t)\to(a,b,c+\varepsilon)} F(z,t) \\ &= L_\iota F(z,t)^{(0,0,\varepsilon)} \end{aligned}$$

Both functors behave the same on maps as well as they are both induced by the inclusion maps of one diagram into another.

Next we need to verify that the Kan extension of the shift map σ_ε is equivalent to the shift map $L_\iota F \Rightarrow L_\iota F^{(0,0,\varepsilon)}$ induced by the inclusion of the comma category. The map $L_i(\sigma_\varepsilon)$ is characterized on (a,b,c) by the compositions,

$$F(z,t) \xrightarrow{F((z,t)\to(z,t+\varepsilon))} F(z,t+\varepsilon) \longrightarrow \mathrm{colim}_{\iota(z',t')\to(a,b,c)} F_\varepsilon(z',t') = L_i(F_\varepsilon)(a,b,c) \cong L_i F^{(0,0,\varepsilon)}(a,b,c)$$

for all $\iota(z,t) \to (a,b,c)$. On the other hand, the shift map $L_i F \to L_i F^{(0,0,\varepsilon)}$ is characterized on (a,b,c) by the compositions,

$$F(z,t) \xrightarrow{id} F(z,t) \longrightarrow \mathrm{colim}_{\iota(z',t')\to(a,b,c+\varepsilon)} F(z',t') = L_i F^{(0,0,\varepsilon)}(a,b,c)$$

for all $\iota(z,t) \to (a,b,c)$ Note that by the universal property of the colimit, the following diagram commutes,

$$\begin{tikzcd} F(z,t) \arrow[r, "id"] & F(z,t) \arrow[r, "f_{(z,t)}"] \arrow[dr, "F((z,t)\to(z,t+\varepsilon))"'] & \mathrm{colim}_{\iota(z',t')\to(a,b,c)+(0,0,\varepsilon)} F(z',t') \\ & & F(z,t+\varepsilon) \arrow[u, "f_{(z,t+\varepsilon)}"] \end{tikzcd}$$

which proves the two functions are characterized by equivalent maps, and are therefore the same. $\square$

Combining this lemma with proposition B.0.2 we can easily show that an interleaving in the $\mathbb{R}$ direction of a pair of 2 parameter zig zag persistence modules induces an interleaving on the Kan extended objects.

Lemma 2.0.2. *Let $F, G \in Fun(\mathcal{Z} \times \mathbb{R}, \boldsymbol{Vect_k})$. If there exists an interleaving,*

$$\alpha : F \to G_\varepsilon \qquad \beta : G \to F_\varepsilon$$

then $L_\iota\alpha$ and $L_\iota\beta$ form a $\vec{u} = (0, 0, \varepsilon)$ interleaving of $L_\iota F$ and $L_\iota G$.

Proof. Apply the above lemmas to see that the following diagram is well defined and commutes:

$$\begin{tikzcd} L_iF \arrow[r, "L_i\alpha"] \arrow[rr, bend right] & L_i(G_\varepsilon) \cong L_iG^{(0,0,\varepsilon)} \arrow[r, "L_i\beta"] & L_i(F_{2\varepsilon}) \cong L_iF^{(0,0,2\varepsilon)} \end{tikzcd}$$

Specifically by lemma 2.0.1 and proposition B.0.2 we have,

$$L_i(\alpha) \circ L_i(\beta) = L_i(\alpha \circ \beta) = L_i(F \to F_{2\varepsilon}) = L_iF \to L_iF^{(2\varepsilon,2\varepsilon)}$$

Checking the interleaving condition for G is equivalent. $\square$

Chapter 3: Stability and Convergence

Now that we have defined a metric on the space of $\mathcal{Z} \times \mathbb{R}$ persistent objects, we want to verify that it meets some fundamental "sanity checks". Specifically, we expect this metric to satisfy some form of stability and demonstrate a convergence property for randomly sampled data.

To begin, we can naturally consider the spaces of regular persistence modules and zig-zag persistence modules as sitting inside the space of $\mathcal{Z} \times \mathbb{R}$ persistence modules via pulling back along the projections,

$$\pi_{\mathbb{R}} : \mathcal{Z} \times \mathbb{R} \to \mathbb{R} \qquad \pi_{\mathcal{Z}} : \mathcal{Z} \times \mathbb{R} \to \mathcal{Z}$$

$$\pi_{\mathbb{R}}^* : Fun(\mathbb{R}, \mathbf{Vect}_k) \to Fun(\mathcal{Z} \times \mathbb{R}, \mathbf{Vect}_k)$$

$$M \mapsto M \circ \pi_{\mathbb{R}}$$

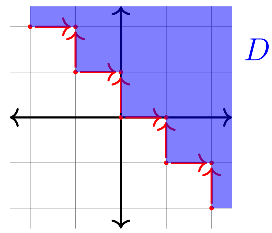
$$\pi_{\mathcal{Z}}^* : Fun(\mathcal{Z}, \mathbf{Vect}_k) \to Fun(\mathcal{Z} \times \mathbb{R}, \mathbf{Vect}_k)$$

$$M \mapsto M \circ \pi_{\mathcal{Z}}$$

Let $d_{\mathbb{R}}$ be the interleaving distance on $Fun(\mathbb{R}, \mathbf{Vect}_k)$, and let $d_{\mathcal{Z}}$ be the distance on $Fun(\mathcal{Z}, \mathbf{Vect}_k)$ obtained by Kan extending along the embedding $i : \mathcal{Z} \to \mathbb{R}^2$ and interleaving in $Fun(\mathbb{R}^2, \mathbf{Vect}_k)$ with shifts given by $(x, y) \mapsto (x, y) + (a, b)$ for $a, b \geq 0$.

Theorem 3.0.1. *$\pi_{\mathbb{R}}^*$ and $\pi_{\mathcal{Z}}^*$ are isometric embeddings.*

Proof. We'll begin by showing $\pi_{\mathbb{R}}^*$ is an isometry. Let $D \subset \mathbb{R}^2$ be the set of (x, y) such that $i(z) \to (x, y)$ for some $z \in \mathcal{Z}$ as in the following image,



First we'll show that $L_\iota \pi_{\mathbb{R}}^* F(a,b,c) = F(c)$ and $L_\iota \pi_{\mathbb{R}}^* F((a,b,c) \to (a',b',c')) = F(c \to c')$ for $(a,b),(a',b') \in D$ and 0 otherwise. This then implies that shifting in the x,y-direction does not "help", meaning an (x,y,z)-shift is equivalent to a $(0,0,z)$ shift. Therefore it is sufficient to consider only ε interleavings in the $(0,0,\varepsilon)$ direction, which is equivalent to a ε interleaving of F and G. To show $L_\iota \pi_{\mathbb{R}}^* F(a,b,c) = F(c)$ for $(a,b) \in D$ we'll examine the colimit formula for $L_\iota \pi_{\mathbb{R}}^* F$:

$$\begin{aligned} L_\iota \pi_{\mathbb{R}}^* F(a,b,c) &= \mathrm{colim}_{\iota(z,t)\to(a,b,c)} \pi_{\mathbb{R}}^* F(z,t) \\ &= \mathrm{colim}_{\iota(z,t)\to(a,b,c)} F(t) \\ &= F(c) \end{aligned}$$

The final inequality we verify by considering the universal property of this colimit. For every $\iota(z,t) \to (a,b,c)$ it factors uniquely through $\iota(z,t) \to \iota(z,c) \to (a,b,c)$ thus a map out of $F(c)$ determines a map out of $F(t)$ for every such $\iota(z,t) \to (a,b,c)$. Furthermore, these maps are necessarily commutative since we working over a poset category. Thus a map out of the diagram $\mathrm{colim}_{\iota(z,t)\to(a,b,c)} F(t)$ is uniquely determined by a map out of $F(c)$. Additionally, any map out of this diagram must behave the same on all $F(c)$ terms in order to commute because $\pi_{\mathbb{R}}^* F(z,c) \to \pi_{\mathbb{R}}^* F(z',c) = id_{F(c)}$ for all $z \to z'$. $L_\iota \pi_{\mathbb{R}}^* F(a,b,c)$ is 0 when $(a,b) \notin D$ because it is the colimit of an empty diagram.

Next we verify the morphisms. A map $L_\iota \pi_{\mathbb{R}}^* F(a,b,c) \to L_\iota \pi_{\mathbb{R}}^* F(a',b',c')$ is determined by the data,

$$\{\pi_{\mathbb{R}}^* F(z,t) \xrightarrow{id} \pi_{\mathbb{R}}^* F(z,t) \xrightarrow{f_{(z,t)}} \mathrm{colim}_{\iota(z',t')\to(a',b'c')} \pi_{\mathbb{R}}^* F(z',t')\}_{\iota(z,t)\to(a,b,c)}$$

However these maps are all determined by what happens to the $\pi_{\mathbb{R}}^*F(z,c) = F(c)$ terms. The canonical maps $f(z,t)$ are just the $F(t \to c')$, so the $F(c)$ terms are mapped to the colimit $F(c')$ by $F(c \to c')$.

Now that we have established that $L_\iota \pi_{\mathbb{R}}^* F$ behaves as F in the z direction and is constant or 0 in the x and y-direction we can examine the interleavings. Clearly an ε-interleaving of F and G extends to a $(0,0,\varepsilon)$ interleaving between $L_\iota \pi_{\mathbb{R}}^* F$ and $L_\iota \pi_{\mathbb{R}}^* G$ implying that $d_{\mathcal{Z}\times\mathbb{R}}(\pi_{\mathbb{R}}^*F, \pi_{\mathbb{R}}^*G) \leq d_{\mathbb{R}}(F,G)$. As mentioned earlier, any (x,y,z)-interleaving between $L_\iota \pi_{\mathbb{R}}^* F$ and $L_\iota \pi_{\mathbb{R}}^* G$ can be improved to a $(0,0,z)$ interleaving which is equivalent a z-interleaving on F and G, implying $d_{\mathbb{R}}(F,G) \leq d_{\mathcal{Z}\times\mathbb{R}}(L_\iota \pi_{\mathbb{R}}^* F, L_\iota \pi_{\mathbb{R}}^* G)$.

We'll similarly verify that $\pi_{\mathcal{Z}}$ is an isometry by showing $L_\iota \pi_{\mathcal{Z}}^* F(a,b,c) = L_i F(a,b)$ and $L_\iota \pi_{\mathcal{Z}}^* F((a,b,c) \to (a',b',c')) = L_i F(a,b) \to L_i F(a',b')$ where $i : \mathcal{Z} \to \mathbb{R}^2$ is the embedding discussed in section 2.4.

$$\begin{aligned} L_\iota \pi_{\mathcal{Z}}^* F(a,b,c) &= \mathrm{colim}_{\iota(z,t)\to(a,b,c)} \pi_{\mathcal{Z}}^* F(z,t) \\ &= \mathrm{colim}_{\iota(z,t)\to(a,b,c)} F(z) \\ &= \mathrm{colim}_{i(z)\to(a,b)} F(z) \end{aligned}$$

Again the last equality is obtained by considering the universal property of both colimits. Clearly a map out of the first diagram implies a map out of the second as the second diagram is a sub-diagram of the first (considering the terms corresponding to $t = c$). Furthermore a map out of the first diagram is fully determined by what it does on the $t = c$ terms as all other terms must commute with the identity maps $F(z) \to F(z)$ corresponding to $(z,t) \to (z,c)$.

Next we need to verify that the maps $L_\iota F((a,b,c) \to (a',b',c')) = L_i F((a,b) \to (a',b'))$. The map $L_\iota F((a,b,c) \to (a',b',c'))$ is determined by the data,

$$\{\pi_{\mathcal{Z}}^* F(z,t) \xrightarrow{id} \pi_{\mathcal{Z}}^* F(z,t) \xrightarrow{f_{(z,t)}} \mathrm{colim}_{\iota(z',t')\to(a',b',c')} \pi_{\mathcal{Z}}^* F(z',t')\}_{\iota(z,t)\to(a,b,c)}$$

and the map $L_iF((a,b) \to (a',b'))$ is determined by the data,

$$\{F(z) \xrightarrow{id} F(z) \xrightarrow{f_{(z)}} \operatorname{colim}_{i(z')\to(a',b')} F(z')\}_{i(z)\to(a,b)}$$

Under the above correspondence it is clear these maps coincide.

Now again we note that a (x,y,z) interleaving of $L_\iota \pi^*_{\mathcal{Z}} F$ and $L_\iota \pi^*_{\mathcal{Z}} G$ is equivalent to a $(x,y,0)$ interleaving, as both functors are constant in the z direction. Such an interleaving is then equivalent to a (x,y) interleaving of L_iF and L_iG demonstrating that $d_{\mathcal{Z}\times\mathbb{R}}(L_\iota \pi^*_{\mathcal{Z}} F, L_\iota \pi^*_{\mathcal{Z}} G) = d_{\mathcal{Z}}(L_iF, L_iG)$ □

Next we want to demonstrate that this metric respects some kind of stability with respect to the Gromov-Hausdorff distance on compact metric spaces. To prove this we will use the formalism of multivalued maps, which was introduced in section 3 of [39].

Definition 3.0.1 (Simplicial multivalued map, subordinate, contiguous)**.**

1. A *multi-valued map (MVM)* $F : X \to Y$ is a function from X to the power set of Y.

2. If X and Y are vertex sets of simplicial complexes, we say an MVM is simplicial if the image of a simplex, $F([x_{i_0}, ..., x_{i_k}]) = \bigcup F(x_{i_j})$ satisfies the property that every finite subset forms a simplex in Y.

3. If $f : X \to Y$ is a regular function such that $f(x) \in F(x)$ for all $x \in X$, we say f is *subordinate* to F.

4. If f and g are both subordinate to the same simplicial multivalued map, we say f and g are *contiguous*.

Lemma 3.0.1. *If f and g are contiguous, then f and g are homotopic.*

Proof. Let f, g be subordinate to a simplicial multivalued map F. Then we have that $f(x), g(x) \in F(x)$, thus $f(x)$ and $g(x)$ form a simplex in Y and therefore lie on a convex set in the geometric realization of Y. Thus we can construct a homotopy between the maps induced by f and g on the geometric realizations of X and Y by defining $H(x,t) = (1-t)f(x)+tg(x)$ on the vertices and extending to the rest of $||X||$. This is well defined on all of $||X||$ as $f([x_0,...,x_k]), g([x_0,...,x_k]) \subset F([x_0,...,x_k])$ implies that the images $f([x_0,...,x_k])$ and $g([x_0,...,x_k])$ form a simplex in Y as well. $\square$

Now the statement of Gromov-Hausdorff stability for $\mathcal{Z} \times \mathbb{R}$ persistence modules:

Theorem 3.0.2. *Let*

$$Z_X = X_1 \hookrightarrow X_{1,2} \hookleftarrow X_2 \hookrightarrow \ldots \hookleftarrow X_{n-1} \hookrightarrow X_{n-1,n} \hookrightarrow X_n$$

$$Z_Y = Y_1 \hookrightarrow Y_{1,2} \hookleftarrow Y_2 \hookrightarrow \ldots \hookleftarrow Y_{n-1} \hookrightarrow Y_{n-1,n} \hookrightarrow Y_n$$

Be zig-zag sequences of compact subsets of metric spaces (X, d_X) and (Y, d_Y) respectively. Assume that there exists isometric embeddings $i_X : \bigcup_{z\in\mathcal{Z}} X_z \to (W, d_W)$, $i_Y : \bigcup_{x\in\mathcal{Z}} Y_z \to (W, d_W)$ with $d_H(i_X(X_z), i_Y(Y_z)) < \varepsilon$ for all $z \in \mathcal{Z}$. Let M_X and M_Y denote the $\mathcal{Z} \times \mathbb{R}$ persistence modules $H_i \circ R(Z_X)$ and $H_i \circ R(Z_Y)$. Then $d_{\mathcal{Z}\times\mathbb{R}}(M_X, M_Y) < 2\varepsilon$

Proof. The proof will use the embeddings i_X and i_Y to construct an ε interleaving between the 2 zigzags.

Let F^z_ε be the multi-valued map from X_z to Y_z defined by,

$$F^z_\varepsilon(x) = \{y : d(i_X(x), i_Y(y)) < \varepsilon\}$$

and let G_ε be the similarly defined map from Y_z to X_z. The assumption that the Hausdorff distance $d_H(i_X(X_z), i_Y(Y_z)) < \varepsilon$ ensures that $F_\varepsilon(x), G_\varepsilon(y)$ are nonempty for all x and y and therefore that F_ε and G_ε are well defined.

Next we will check that these F^z_ε are simplicial with respect the simplicial complexes $R^t(X_z)$ and $R^{t+2\varepsilon}(Y_z)$. Let $[x_{i_0}, ..., x_{i_k}]$ be a simplex in $R^t(X_z)$, meaning $d_X(x_{i_m}, x_{i_n}) \leq t$, or equivalently, $d_W(i_X(x_{i_m}), i_X(x_{i_n})) \leq t$. Let $[y_{j_1}, ..., y_{j_l}]$ be a subset of $F^z_\varepsilon([x_{i_0}, ..., x_{i_k}])$. By the triangle inequality and the fact that each y_{j_l} is in the image of some x_{i_l}, we have

$$\begin{aligned} d_Y(y_{j_r}, y_{j_s}) &= d_W(i_Y(y_{j_r}), i_Y(y_{j_s})) \\ &\leq d_W(i_Y(y_{j_r}), i_X(x_{i_m})) + d_W(i_X(x_{i_m}), i_X(x_{i_n})) + d_W(i_X(x_{i_n}), i_Y(y_{j_s})) \\ &\leq t + 2\varepsilon \end{aligned}$$

Therefore $[y_{j_1}, ..., y_{j_l}]$ is a simplex in $R^{t+2\varepsilon}(Y_z)$.

Thus the F_ε^z induce maps $H_i(R^t(X_z)) \to H_i(R^{t+2\varepsilon}(Y_z))$. We want to show that these maps form a map $F_{\varepsilon *} : M_X \Rightarrow M_Y^{2\varepsilon}$ by showing they commute with the internal maps in $\mathcal{Z} \times \mathbb{R}$.

For the shift maps in the $\mathbb{R}$ parameter, the maps on homology are induced by the inclusions $R^t(X) \to R^{t+s}(X)$, which correspond to the identity function $id : X \to X$ at the level of the vertex set. Therefore, since $F_\varepsilon^z \circ id_{X_z} = id_{Y_z} \circ F_\varepsilon^z$ as multi-valued maps, the induced maps on homology commute as well.

To verify that the zigzag direction commutes, we need to consider diagrams of the form,

$$\begin{tikzcd} H_i(R^t(X_j)) \arrow[r, "F_{\varepsilon *}^j"] \arrow[dr, "i_*"] & H_i(R^{t+\varepsilon}(Y_j)) \arrow[dr, "i_*"] & \\ & H_i(R^t(X_{j,j+1})) \arrow[r, "F_{\varepsilon *}^{j,j+1}"] & H_i(R^{t+\varepsilon}(Y_{j,j+1})) \\ H_i(R^t(X_{j+1})) \arrow[r, "F_{\varepsilon *}^{j+1}"] \arrow[ur, "i_*"] & H_i(R^{t+\varepsilon}(Y_{j+1})) \arrow[ur, "i_*"] & \end{tikzcd}$$

Note that it may not be possible to choose subordinate maps of F_ε for $j, j+1$ and $(j, j+1)$ such that this diagram commutes on the level of the simplicial complexes, however F_ε^j is subordinate to $F_\varepsilon^{j,j+1}$ restricted to X_j (it may not be strictly equal as elements of Y_{j+1} may end up in the image of an element in X_j). Therefore it is possible to choose a subordinate function such that the the squares commute. This tells us that the diagram of simplicial complexes commutes up to homotopy, so at the level of homology the diagram commutes strictly. Thus this multi-valued map F_ε induces a well defined map $F_{\varepsilon *} : M_X \Rightarrow M_Y^{2\varepsilon}$. Equivalently, $G_{\varepsilon *}$ forms a well defined map $M_Y \Rightarrow M_X^{2\varepsilon}$

All that remains to show is that $F_{\varepsilon *}$ and $G_{\varepsilon *}$ form an interleaving. Observe that the identity on X is subordinate to $G_\varepsilon \circ F_\varepsilon$ as $G_\varepsilon^z \circ F_\varepsilon^z(x) = \{x' \in X_z : \exists y \in Y_z \text{ s.t. } d_W(i_X(x), i_Y(y)) < \varepsilon, d_W(i_X(x'), i_Y(y)) < \varepsilon\} \ni x$, and recall that the map that shifts in the $\mathbb{R}$ direction by 4ε is induced by the identity on the vertex set. Therefore since the identity is subordinate to $G_\varepsilon \circ F_\varepsilon$, the induced map $(G_\varepsilon \circ F_\varepsilon)_* = G_{\varepsilon *} \circ F_{\varepsilon *}$ is equal to the shift map $M_X \Rightarrow M_X^{4\varepsilon}$. A similar argument shows $F_{\varepsilon *} \circ G_{\varepsilon *}$ is the shift $M_Y \Rightarrow M_Y^{4\varepsilon}$, thus proving that these maps form a 2ε interleaving in the $\mathbb{R}$ direction. Applying lemma 2.0.2 completes the proof. □

The next result we will show verifies that this metric behaves as expected on sampled data. Specifically, if we sample data from a space X and form a $\mathcal{Z} \times \mathbb{R}$-persistence module from it according to a typical methodology of taking the homology of a filtration obtained from a union-zigzag of subsamples, we want to show that this converges to some $\mathcal{Z} \times \mathbb{R}$-persistence module representing the true homology of the space X. In this case the "ground truth" module will be defined to be $M_X(z,t) = H_k(R_t(X))$

Theorem 3.0.3. *Let (X, d, μ) be a compact metric measure space such that $Supp(\mu) = X$. Suppose A is a set of N points sampled from X according to μ, and $A_1, ..., A_k$ a sequence of uniform subsamples of A (sampled with replacement), each with $|A_i| = n$. Let M be the $\mathcal{Z} \times \mathbb{R}$-persistence module obtained by applying $H_k \circ R$ to the union zigzag,*

$$A_1 \rightarrow A_1 \cup A_2 \leftarrow A_2 \rightarrow \ldots \leftarrow A_{k-1} \rightarrow A_{k-1} \cup A_k \leftarrow A_k$$

Note that we are taking $M(z,t) = 0$ for $z \neq 1, ..., k$ or $i, i+1$ for $i = 1, ..., k-1$. Furthermore, let M_X denote the $\mathcal{Z} \times \mathbb{R}$-persistence module defined by

$$M_X(z,t) = H_k(R_t(X))$$

For $z = 1, ..., k$ or $i, i+1$ for $i = 1, ..., k-1$ and 0 else. Then for any $\varepsilon > 0$,

$$\lim_{n,N\to\infty} \mathbb{P}(d_{\mathcal{Z}\times\mathbb{R}}(M, M_X) > \varepsilon) = 0$$

For the proof of this theorem, we will need the following lemma:

Lemma 3.0.2. *If (X, d, μ) is compact with Supp$(\mu) = X$, then for every ε there exists a $p_\varepsilon > 0$ such that $x \in X \implies \mu(B_\varepsilon(x)) > p_\varepsilon$*

Proof. Define $m_\varepsilon : X \to [0,1]$ to be $m_\varepsilon(x) = \mu(B_\varepsilon(x))$. First I want to show that m_ε is lower semi-continuous. Let $\{x_n\}_n \to x$ in X. For every $\delta > 0$ there is an N such that $m > N \implies B_{\varepsilon-\delta}(x) \subset B_\varepsilon(x_m)$, therefore $\mu(B_{\varepsilon-\delta}(x)) \leq \mu(B_\varepsilon(x_m))$. Thus $\liminf_{x_n\to x} m_\varepsilon(x_n) \geq B_{\varepsilon-\delta}(x)$. since the choice of δ was arbitrary and measures are continuous from below for measurable sets, this implies $m_\varepsilon(x) \leq \liminf_{x_n\to x} m_\varepsilon(x_n)$.

Now I want to show that this implies m_ε achieves a minimum. let $a = \inf m_\varepsilon(X)$. Construct a sequence $\{x_n\}$ in X such that $m_\varepsilon(x_n) < a + 1/n$ (note this assumes a is a limit point of $m_\varepsilon(X)$, but if this is not the case it must be true that $a \in m_\varepsilon(X)$). Because X is compact there exists a convergent subsequence $\{x_{n_i}\} \to x$. By lower semi-continuity we then have $a \leq m_\varepsilon(x) \leq \liminf_{x_{n_i}\to x} m_\varepsilon x_{n_i} = a$ so $a = m_\varepsilon(x)$.

Because $x \in X$, we know that $a = \mu(B_\varepsilon(x)) = \min_{x'\in X} B_\varepsilon(x') > 0$ $\square$

Now we'll return to the proof of theorem 3.0.3.

Proof. By theorem 3.0.2,

$$\begin{aligned}
\mathbb{P}(d_{\mathcal{Z}\times\mathbb{R}}(M_X, M) > \varepsilon) &\leq \mathbb{P}(\bigcup_{i=1}^{k}\{d_{GH}(A_i, X) > \varepsilon\}) \\
&\leq \sum_{i=1}^{k} \mathbb{P}(d_{GH}(A_i, X) > \varepsilon) && \text{Union Sum Bound} \\
&= k\mathbb{P}(d_{GH}(A_i, X) > \varepsilon) && A_i \text{ are identically distributed}
\end{aligned}$$

Let $B_{\varepsilon/2}(x_1), ..., B_{\varepsilon/2}(x_m)$ be a finite cover of X by $\varepsilon/2$ balls. Note that if $A_i \cap B_{\varepsilon/2}(x_j) \neq \phi$ for all j, then for all $x \in X$ we have $x \in B_{\varepsilon/2}(x_j)$ for some j therefore $d(x, A_i) \leq d(x, x_j) + d(x_j, a) \leq \varepsilon$ for some $a \in B_{\varepsilon/2}(x_j)$, which implies $d(X, A_i) \leq \varepsilon$. Thus we have,

$$\begin{aligned}\mathbb{P}(d_{GH}(A_i, X) > \varepsilon) &\leq \mathbb{P}(\bigcup_{i=1}^{m}\{|A_i \cap B_{\varepsilon/2}(x_j)| = 0\}) \\ &\leq \sum_{i=1}^{m} \mathbb{P}(|A_i \cap B_{\varepsilon/2}(x_j)| = 0) \qquad \text{Union sum bound}\end{aligned}$$

To analyze $\mathbb{P}(|A_i \cap B_{\varepsilon/2}(x_j)| = 0)$ we'll condition on the number of points sampled from $B_{\varepsilon/2}(x_j)$ in A. Furthermore, since we are upper bounding this probability and the event is clearly most likely when $\mu(B_{\varepsilon/2}(x_j))$ has minimal measure, we may obtain an upper bound by taking $\mu(B_{\varepsilon/2}(x_j)) = p$ where $p = p_{\varepsilon/2} = \min_{x\in X} B_{\varepsilon/2}(x)$, which by lemma 3.0.2 is > 0.

$$\begin{aligned}\mathbb{P}(|A_i \cap B_{\varepsilon/2}(x_j)| = 0) &= \sum_{l=0}^{N} \mathbb{P}(|A_i \cap B_{\varepsilon/2}(x_j)| = 0 || A \cap B_{\varepsilon/2}(x_j)| = l)\mathbb{P}(|A \cap B_{\varepsilon/2}(x_j)| = l) \\ &\leq \sum_{l=0}^{N} (\frac{N-l}{N})^n \binom{N}{l} p^l (1-p)^{N-l}\end{aligned}$$

At this point we can see that the last expression is $\mathbb{E}[(Y_N/N)^n]$ for Y_N a binomial $(N, 1-p)$ random variable. Combining all our bounds together so far we have,

$$\mathbb{P}(d_{\mathcal{Z}\times\mathbb{R}}(M, M_X) > \varepsilon) \leq km\mathbb{E}[(\frac{Y_N}{N})^n]$$

I want to show that the RHS of the inequality goes to 0 as n and N go to ∞. Since k is fixed and m is a constant with respect to ε, all we need to show is that $\mathbb{E}[(Y_N/N)^n]$ converges to 0 as $n, N \to \infty$. To do this we'll use indicator random variables to separate the expectation into 2 cases. Let $1-p<q<1$.

$$\begin{aligned}
\mathbb{E}[(\frac{Y_N}{N})^n] &= \mathbb{E}[(\frac{Y_N}{N})^n \mathbb{1}[\frac{Y_N}{N} \le q]] + \mathbb{E}[(\frac{Y_N}{N})^n \mathbb{1}[\frac{Y_N}{N} > q]] \\
&\le q^n + \mathbb{E}[(\frac{Y_N}{N})^n \mathbb{1}[\frac{Y_N}{N} > q]] && (\frac{Y_N}{N})^n \mathbb{1}[\frac{Y_N}{N} \le q] \le q^n \\
&\le q^n + \mathbb{E}[\mathbb{1}[\frac{Y_N}{N} > q]] && (\frac{Y_N}{N})^n \le 1 \\
&= q^n + \mathbb{P}(\frac{Y_N}{N} > q)
\end{aligned}$$

Observe that Y_N is equivalent to the sum of N i.i.d. Bernoulli $1-p$ random variables. Therefore by the law of large numbers, $Y_N/N \xrightarrow{N\to\infty} 1-p$ in probability, thus $\mathbb{P}(Y_N/N > q) \xrightarrow{N\to\infty} 0$ and because $q<1$ we have $q^n \xrightarrow{n\to\infty} 0$, thus concluding the proof. $\square$

Remark 3.0.1. The proof of this theorem also provides us with information about the rate of convergence of $\mathbb{P}(d_{\mathcal{Z}\times\mathbb{R}}(M, M_X) > \varepsilon) \to 0$. Write $q = 1-p+\delta$ where $\delta < p = \min_{x\in X} B_{\varepsilon/2}(x)$, then Hoeffding's inequality tells us that,

$$\mathbb{P}(\frac{Y_N}{N} > q) \le e^{-2N\delta^2} = (e^{-2\delta^2})^N$$

so we have,

$$\mathbb{P}(d_{\mathcal{Z}\times\mathbb{R}}(M, M_X) > \varepsilon) = \mathcal{O}((1-p+\delta)^n + (e^{-2\delta^2})^N)$$

for any $0<\delta<p$.

Chapter 4: Invariants

Invariants are a critical component of any persistent homology theory. As is the case with multi-parameter persistence, the $\mathcal{Z} \times \mathbb{R}$ poset does not have nice representation theory and therefore does not admit a nice barcode like invariant. However we can attempt to adapt some of the approaches used for multi-parameter persistence such as the rank invariant, the signed barcode and the fibered barcode.

To begin, we'll discuss the rank invariant and the generalized rank invariant. Recall the definitions for a general poset $\mathcal{P}$:

Definition 4.0.1 (Rank Invariant)**.** The *rank invariant* of $M : \mathcal{P} \to \mathbf{Vect}_k$ is a map $Rk_M : mor(\mathcal{P}) \to \mathbb{Z}$ that sends $p \to p'$ to

$$Rk_M(p \to p') = rk(M(p \to p'))$$

Definition 4.0.2 (Generalized Rank Invariant)**.** The *generalized rank invariant* of $M : \mathcal{P} \to \mathbf{Vect}_k$ is a map $GRk : \mathcal{I} \to \mathbb{Z}$, that sends I an interval in $\mathcal{P}$ to,

$$Grk_M(I) = rk(\lim M|_I \to \mathrm{colim} M|_I)$$

The generalized rank invariant extends the regular rank invariant. If $I = [p, p'] = \{q \in \mathcal{P} : p \leq q \leq p'\}$ then $\lim M|_I \to \mathrm{colim} M|_I = M(p) \to M(p')$, so $Grk_M(I) = rkM(p \to p')$. The rank invariant allows us to see more information across different types of intervals in $\mathcal{P}$, which is critical when zig-zags are involved as many elements are not comparable.

As is the case for multi-parameter persistence, these invariants are not complete. In the case of the normal rank invariant, it is not even complete on modules that are block decomposable as we'll see later in example 4.0.1. The generalized rank invariant however is complete on block decomposable modules by proposition 2.4 in [29]. However both invariants are still useful for understanding the outputs of $\mathcal{Z} \times \mathbb{R}$ persistent homology.

A natural question is whether there is an equivalent representation of the (generalized) rank invariant on $\mathcal{Z} \times \mathbb{R}$ to the signed barcode or the fibered barcode. To begin with the signed barcode, the answer is yes. This is a consequence of a general theorem proven in [29]:

Theorem 4.0.1 (Theorem 2.12 in [29])**.** *Let $\mathcal{I}$ be a locally finite collection of intervals in $\mathcal{P}$. Then any function $r : \mathcal{I} \to \mathbb{Z}$ with upward finite support*[1] *can uniquely be written as a (possibly infinite, but point-wise finite) $\mathbb{Z}$-linear combination of the functions Grk_{k_I} on $\mathcal{I}$, with $I \in \mathcal{I}$.*

Proof. The proof of this theorem is constructive so I will include a sketch of it here. Let $(\mathcal{I}, \subseteq)$ be a locally finite poset (in this context, it is the poset of intervals in $\mathcal{P}$ with the subset partial order). The *incidence algebra* on $(\mathcal{I}, \subseteq)$ is the space of functions from nonempty intervals in $\mathcal{I}$ to $\mathbb{Z}$, with products defined by convolution,

$$(f * g)(I, J) = \sum_{I \subseteq K \subseteq J} f(I, K) g(K, J)$$

This algebra has a unit given by $\mathbb{1}_{I=J}$. Let ζ be the element defined by $\zeta(I, J) = 1$ for all nonempty intervals (I, J). This element has an inverse μ defined inductively by,

$$\mu(I, J) = \begin{cases} 1 & I = J \\ -\sum_{I \subsetneq K \subseteq J} \mu(K, J) & \text{else} \end{cases}$$

The space of functions $\mathcal{I} \to \mathbb{Z}$ with upward finite support is a right module over this algebra

[1] defined in appendix section A.

with multiplication defined by,

$$(r * f)(I) = \sum_{I \subseteq J} r(J) f(I, J)$$

Note that the requirement that r have upward finite support implies that this sum is finite. Thus we have the relation $r = s * \zeta \iff r * \mu = s$

Define $\alpha_I = (r * \mu)(I)$. Then,

$$\begin{aligned} r(I) = ((r * \mu) * \zeta)(I) &= \sum_{I \subseteq J} (r * \mu)(J) \zeta(I, J) \\ &= \sum_{I \subseteq J} (r * \mu)(J) && \zeta(I, J) = 1 \text{ if } I \subseteq J \\ &= \sum_{I \subseteq J} \alpha_J \\ &= \sum_{I \in \mathcal{I}} \alpha_J \mathbb{1}_{I \subseteq J} \\ &= \sum_{I \in \mathcal{I}} \alpha_J Grk_{k_J}(I) && Grk_{k_J}(I) = \mathbb{1}_{I \subseteq J} \end{aligned}$$

Therefore $r = \sum_{I \in \mathcal{I}} \alpha_J Grk_{k_J}(I)$ $\square$

Corollary 4.0.1. *Let M be an essentially discrete, point-wise finite dimensional $\mathcal{Z} \times \mathbb{R}$ persistence module. Then there exists a signed barcode for the (generalized) rank invariant, that is a pair $(\mathcal{R}, \mathcal{S})$ of collections of intervals in $\mathcal{Z} \times \mathbb{R}$ such that,*

$$Grk_M = (\bigoplus_{R \in \mathcal{R}} Grk_{k_R}) - (\bigoplus_{S \in \mathcal{S}} Grk_{k_S})$$

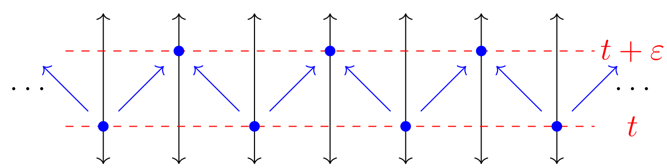
$$Rk_M = (\bigoplus_{R \in \mathcal{R}} Rk_{k_R}) - (\bigoplus_{S \in \mathcal{S}} Rk_{k_S})$$

Proof. To construct $\mathcal{R}$ and $\mathcal{S}$ apply the theorem to write $GRk_M = \sum_{I \in \mathcal{I}} \alpha_I k_I$ where $\alpha_I \in \mathbb{Z}$. Then $\mathcal{R} = \{I : \alpha_I > 0\}, \mathcal{S} = \{I : \alpha_I < 0\}$ with multiplicities given by the $|\alpha_I|$. Observe that because the generalized rank invariant extends the rank invariant, a signed barcode for the generalized rank invariant is signed barcode for the rank invariant as well. $\square$

Furthermore, this in combination with theorem 2.6 of [29] tells us that a minimal signed barcode is characterized by $\mathcal{R} \cap \mathcal{S} = \phi$ and is unique.

The next invariant that we will adapt is the fibered barcode. Recall that the fibered barcode of a multi-parameter persistence module F is the invariant which assigns to each line $\ell = \vec{a}t + \vec{b} \in \mathbb{R}^n$ the barcode $\mathcal{B}_{F|_\ell}$. This invariant is equivalent to the rank invariant and satisfies an internal and external stability condition.

Definition 4.0.3 (Fibered zig-zag barcode)**.** Let M be a $\mathcal{Z} \times \mathbb{R}$ persistence module. For every $(t, \varepsilon) \in \mathbb{R} \times \mathbb{R}_{\geq 0}$ let $\mathcal{Z}_{t,\varepsilon} \subset \mathcal{Z} \times \mathbb{R}$ be the zig-zag slice obtained by taking the "valleys" at points with $\mathbb{R}$-value t and the "peaks" at points with $\mathbb{R}$-value $t + \varepsilon$ as in the following image,



The fibered zig-zag barcode is the function,

$$\mathcal{F} : \mathbb{R} \times \mathbb{R}_{\geq 0} \to \mathcal{B}$$

$$(t, \varepsilon) \mapsto \mathcal{B}_{M|_{\mathcal{Z}_{t,\varepsilon}}}$$

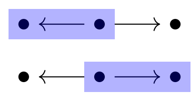
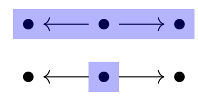
where $\mathcal{B}_{M|_{\mathcal{Z}_{t,\varepsilon}}}$ is the multiset of intervals associated to the interval decomposition of $M|_{\mathcal{Z}_{t,\varepsilon}}$

Observe that the fibered zig-zag determines the rank invariant because any map $(z, t) \to (z', t')$ is contained in some $\mathcal{Z}_{t,t'-t}$ and the rank of $M((z, t) \to (z', t')) = M|_{\mathcal{Z}_{t,t'-t}}((z, t) \to (z', t'))$ is determined by the interval decomposition. However, the converse is not true. This is because, unlike in the case of a totally ordered set, we cannot recover the interval decomposition of a zig-zag persistence module from the rank invariant alone. This is illustrated by the following example of 2 zig-zag persistence modules that have the same rank invariant but different interval decompositions:

Example 4.0.1.

$$M_1 \quad k \xleftarrow{[1,0]} k^2 \xrightarrow{[0,1]} k \qquad\qquad M_2 \quad k \xleftarrow{[1,0]} k^2 \xrightarrow{[1,0]} k$$

One can easily verify that these have equivalent rank invariants, however the interval decompositions are as follows:

M_1  M_2 

We might wonder if the fibered zig-zag is equivalent to the generalized rank invariant on $\mathcal{Z} \times \mathbb{R}$. This unfortunately does not quite work out either. The generalized rank invariant does not restrict so nicely to sub-posets as you loose the information of intervals spanning the larger poset. Thus the fibered zig-zag appears to provide slightly more information than the rank invariant, but not quite as much as the entirety of the generalized rank invariant.

The primary flaw of the fibered zig-zag is that it fails to satisfy an internal or external stability result. This is because the resulting zig-zags $M|_{Z_{t,\varepsilon}}$ are completely blind to the $\mathbb{R}$ direction parameter. In the case of the regular fibered barcode, shifting the line $\ell = \vec{a}t + \vec{b}$ by a factor of ε is akin to shifting by $\varepsilon\vec{a}$ in $\mathbb{R}^n$. This shift sees all directions of $\mathbb{R}^n$. The interleavings of $M|_{\mathcal{Z}_{t,\varepsilon}}$ will see only the $\mathcal{Z}$ direction and not the $\mathbb{R}$ direction. We can see this illustrated in the following particularly bad example:

Example 4.0.2. Consider the following $\mathcal{Z} \times \mathbb{R}$ persistence module,

$$M = \pi^*_{\mathbb{R}}\Big(\bigoplus_{n\in\mathbb{Z}} k_{[n\delta,(n+1)\delta)}\Big)$$

This module is $\delta/2$ trivial, as all of the intervals in $\bigoplus_{n\in\mathbb{Z}} k_{[n\delta,(n+1)\delta)}$ have length δ. However the zig-zag slices can be characterized as,

$$M|_{\mathcal{Z}_{t,\varepsilon}}(z) = k \qquad M|_{\mathcal{Z}_{t,\varepsilon}}(z \to z') = \begin{cases} id & t, t+\varepsilon \in [n\delta, (n+1)\delta) \text{ for some } n \in \mathbb{Z} \\ 0 & \text{else} \end{cases}$$

Both cases are infinite distance from 0, demonstrating that external stability can fail arbitrarily badly. Furthermore, the two cases are infinite distance from one another, demonstrating that internal stability can fail arbitrarily badly as well.

Thus despite the fact that the fibered barcode is equivalent to the rank invariant, it may not prove to be the best approach to understanding a $\mathcal{Z} \times \mathbb{R}$ persistence module in practice as the zig-zag slices provide limited information about the $\mathbb{R}$ direction and are therefore very unstable with respect to perturbations in the $\mathbb{R}$ direction. One possible approach to this problem would be to consider the data of the fibered barcodes $M(z, -)$ in addition to the zig-zag slices.

Chapter 5: Density Sensitive Bifiltrations

As stated in the background section, robustness to outliers is a common concern in the study of persistent homology. Several of the examples of $\mathcal{Z} \times \mathbb{R}$ persistence modules discussed in section 2 serve to address this concern, such as bootstrap sampling and comparing different density parameters. In other settings, it may be useful to consider constructions of multi-parameter zig-zag modules that are robust. Specifically, we will consider extending our metric to the space of $\mathcal{Z} \times \mathcal{P}$ persistence modules where $\mathcal{P}$ is a product poset of totally ordered sets. This allows us to apply density sensitive bifiltrations to zig-zag sequences of point clouds. We will then show that the proof Gromov-Prokhorov stability for density sensitive bifiltrations in [27] readily generalizes to the zig-zag case.

The metric we defined on the space of $\mathcal{Z} \times \mathbb{R}$ persistence modules extends easily to $\mathcal{Z} \times \mathcal{P}$ persistence modules with $\mathcal{P}$ a multi-parameter poset by extending the $\mathcal{Z} \to \mathbb{R}^2$ embedding in the obvious way to,

$$\iota : \mathcal{Z} \times \mathcal{P} \to \mathbb{R}^2 \times \mathcal{P}$$

and taking the left Kan extension to obtain a fully faithful embedding,

$$Fun(\mathcal{Z} \times \mathcal{P}, \mathbf{Vect}_k) \to Fun(\mathbb{R}^2 \times \mathcal{P}, \mathbf{Vect}_k)$$

We then obtain an interleaving distance on the space of $\mathcal{Z} \times \mathcal{P}$ modules by restricting the interleaving distance on the space of $\mathbb{R}^2 \times \mathcal{P}$ modules.

Recall the definitions of the density sensitive bifiltrations stated in the multi-parameter background section. For the purposes of this discussion we will use the unnormalized versions, as the normalized versions do not naturally include into one another along the zig-zag direction.

Definition 5.0.1 (Density sensitive bifiltrations)**.** Let (X, d) be a finite measure space. For all but the Rips complexes, we must assume that (X, d) is embedded in a larger ambient metric space (Z, d).

1. (Multicover) The *multicover bifiltration* $\mathcal{M}(X) : \mathbb{R}^{op}_{\geq 0} \times \mathbb{R}_{\geq 0} \to \mathbf{Top}$ is defined to be,

$$\mathcal{M}(X)_{k,r} = \{z \in Z : |B_r(z) \cap X| \geq k\}$$

2. (Degree Rips/Čech) The *degree Rips* bifiltration $\mathcal{D}R(X) : \mathbb{R}^{op}_{\geq 0} \times \mathbb{R}_{\geq 0} \to \mathbf{Simp}$ is defined such that $\mathcal{DR}(X)_{k,r}$ is the maximal subcomplex of $R(X)_r$ consisting of vertices $x \in X$ such that x has degree at least $k-1$ in the 1-skeleton of $R(X)_r$. Similarly, the *degree Čech* complex $\mathcal{D}C(X) : \mathbb{R}^{op}_{\geq 0} \times \mathbb{R}_{\geq 0} \to \mathbf{Simp}$ is defined in the same way using the Čech complex instead of the Vietoris-Rips complex.

3. (Subdivision Rips/Čech) The *subdivision Rips* bifiltration $\mathcal{S}R(X) : \mathbb{R}^{op}_{\geq 0} \times \mathbb{R}_{\geq 0} \to \mathbf{Simp}$ is defined such that $\mathcal{S}R(X)_{k,r}$ is the subcomplex of $\mathrm{Bary}(R(X)_r)$ consisting only of vertices corresponding to simplices with dimension at least $k-1$ in $R(X)_r$. The *subdivision Čech* bifiltration $\mathcal{S}C : \mathbb{R}^{op}_{\geq 0} \times \mathbb{R}_{\geq 0} \to \mathbf{Simp}$ is defined analogously.

Applying any of these bifiltrations to a zig-zag sequence of point clouds,

$$X_1 \to X_{1,2} \leftarrow X_2 \to \ldots \leftarrow X_{n-1} \to X_{n-1,n} \leftarrow X_n$$

results in a $\mathcal{Z} \times \mathbb{R}^{op}_{\geq 0} \times \mathbb{R}_{\geq 0}$ filtration as the inclusions $X_i \hookrightarrow X_{i,i+1} \hookleftarrow X_{i+1}$ induce inclusions $\mathcal{M}(X_i) \hookrightarrow \mathcal{M}(X_{i,i+1}) \hookleftarrow \mathcal{M}(X_{i+1})$, and likewise for $\mathcal{S}C, \mathcal{S}R, \mathcal{D}C$ and $\mathcal{D}R$. However, this is not true for the normalized versions. For instance, in the case of $\mathcal{M}$, we have that $z \in \mathcal{M}(X_i)_{r,k}$ if $|B_r(z) \cap X_i| \geq k$. This implies that $|B_r(z) \cap X_{i,i+1}| \geq k$ as $X_i \subseteq X_{i,i+1}$ resulting in an inclusion map $\mathcal{M}(X_i) \hookrightarrow \mathcal{M}(X_{i,i+1})$, but it is not the case that $|B_r(z) \cap X_i|/|X_i| \geq s \implies |B_r(z) \cap X_{i,i+1}|/|X_{i,i+1}| \geq s$.

Theorems 1.6 and 1.7 in [27] proved that these density sensitive filtrations are stable with respect to the Prokhorov/Gromov-Prokhorov distance on metric measure spaces. We can adapt the proof of these theorems to the zig-zag case by observing that the interleavings constructed in [27] commute with the maps induced by the inclusions in the zig-zag direction. We will state and prove these theorems below for $\mathcal{Z} \times \mathbb{R}^{op}_{\geq 0} \times \mathbb{R}_{\geq 0}$ following the same method as in [27], with adjustments made for the unnormalized versions of the filtrations.

Theorem 5.0.1 (Stability of multicover and subdivision bifiltrations)**.** *Let*

$$Z_X = X_1 \to X_{1,2} \leftarrow X_2 \to \ldots \leftarrow X_{n-1} \to X_{n-1,n} \leftarrow X_n$$

$$Z_Y = Y_1 \to Y_{1,2} \leftarrow Y_2 \to \ldots \leftarrow Y_{n-1} \to Y_{n-1,n} \leftarrow Y_n$$

be zigzag sequences of subsets of a metric space (W, d). Let ν^z_X be the pushforward of the counting measure on X_z for $z = i$ or $i + 1$, specifically $\nu^z_X(A) = |A \cap X_z|$ and define ν^z_Y analogously.

1. *$\mathcal{M}(Z_X)$ and $\mathcal{M}(Z_Y)$ are $s^\delta(z, k, r) = (z, k - \delta, r + \delta)$ interleaved for all δ such that $\delta > \max\{d_{Pr}(\nu^z_X, \nu^z_Y) : z = i$ or $i, i + 1\}$*

2. *If we assume that W is good, then $\mathcal{SC}(Z_X)$ and $\mathcal{SC}(Z_Y)$ are s^δ-homotopy interleaved.*

3. *If we remove the assumption that $X = \bigcup_{z \in \mathcal{Z}} X_z$ and $Y = \bigcup_{z \in \mathcal{Z}} Y_z$ are subsets of an ambient metric space, and let μ^z_X, μ^z_Y be the counting measures for $z = i$ or $i, i + 1$, then $\mathcal{SR}(Z_X)$ and $\mathcal{SR}(Z_Y)$ are s^δ interleaved for all δ such that,*

$$\delta > \inf_{\varphi, \psi}\{\max\{d_{Pr}(\varphi_* \mu^z_X, \psi_* \mu^z_Y) : z = i \text{ or } i, i + 1\}\}$$

where $\varphi : (X, d_X) \to (W, d)$ and $\psi : (Y, d_Y) \to (W, d)$ range over all isometric embeddings into metric spaces (W, d).

Proof. We'll begin with the proof of 1. Let $\delta > \max\{d_{Pr}(\nu^z_X, \nu^z_Y) : z = i$ or $i, i + 1\}$. Then for any $w \in W$, we have that $\nu_X^z(\overline{B}_r(w)^\delta) \geq \nu_Y^z(B_r(w)) - \delta$ for all $z = i$ or $i, i+1$. One can easily verify using the triangle inequality that $\overline{B}_r(w)^\delta \subseteq B_{r+\delta}(w)$ therefore

$$\nu_X^z(B_{r+\delta}(w)) \geq \nu_Y^z(B_r(w)) - \delta$$

Since $\nu_X^z(A) = |A \cap X_z|$ and likewise for Y, we have inclusions,

$$\mathcal{M}(Y_z)_{k,r} \hookrightarrow \mathcal{M}(X_z)_{k-\delta,r+\delta}$$

Since these maps are inclusions, they commute with the inclusions $\mathcal{M}(X_i) \hookrightarrow \mathcal{M}(X_{i,i+1}) \hookleftarrow \mathcal{M}(X_{i+1})$, $\mathcal{M}(Y_i) \hookrightarrow \mathcal{M}(Y_{i,i+1}) \hookleftarrow \mathcal{M}(Y_{i+1})$ and therefore assemble into an inclusion map $\mathcal{M}(Z_Y) \Rightarrow \mathcal{M}(Z_X) \circ s^\delta$ where s^δ is the poset translation $s^\delta(z,k,r) = (z, k-\delta, r+\delta)$. The inverse map $\mathcal{M}(Z_X) \Rightarrow \mathcal{M}(Z_Y) \circ s^\delta$ is defined analogously.

Part 2 of the theorem follows from 1 by application of a version of the multicover nerve theorem for general diagrams of spaces which can be found as theorem 4.13 of [27]. Specifically, this theorem states that if U is a weakly good open cover of a functor $F : \mathcal{C} \to \mathbf{Top}$, then $\mathcal{M}(U)$ is weakly equivalent to to $\mathcal{S}(\mathcal{N}(U))$. Taking $\mathcal{C} = \mathcal{Z} \times \mathbb{R}$ and $U_{z,r} = \{B_r(x)\}_{x \in X_z}$, then we have $\mathcal{M}(U) = \mathcal{M}(Z_X)$ and $\mathcal{S}(\mathcal{N}(U)) = \mathcal{SC}(Z_X)$. Thus combining the interleaving from part 1 with this weak equivalence provides a homotopy interleaving of $\mathcal{SC}(Z_X)$ with $\mathcal{SC}(Z_Y)$.

Part 3 follows from part 2 using the following construction. First, given δ such that

$$\delta > \inf_{\varphi,\psi}\{\max\{d_{Pr}(\varphi_*\mu_X^z, \psi_*\mu_Y^z) : z = i \text{ or } i, i+1\}\}$$

Let $\varphi : X \to W$ and $\psi : Y \to W$ be isometric embeddings to a finite metric space (W, d_W) such that

$$\delta > \max\{d_{Pr}(\varphi_*\mu_X^z, \psi_*\mu_Y^z) : z = i \text{ or } i, i+1\}$$

Now consider the *Kuratowski embedding*,

$$K : (W, d_W) \to (\mathbb{R}^W, || - ||_\infty)$$

$$K(w) = d_W(w, -)$$

This is an isometric embedding, and furthermore, $R(A) = R(K(A)) = C(K(A))$ for any $A \subset Z$. Thus taking $X' = K(\varphi(X))$ and $Y' = K(\psi(Y))$ we have $R(X) = C(X')$ and $R(Y) = C(Y')$, therefore $\mathcal{S}R(X) = \mathcal{S}C(X')$ and $\mathcal{S}R(Y) = \mathcal{S}C(Y')$. Combining this with the s^δ homotopy interleaving from 2 completes the proof. □

Theorem 5.0.2 (Stability of degree bifiltrations). *Let γ be the $\mathcal{Z} \times \mathbb{R}^{op}_{\geq 0} \times \mathbb{R}_{\geq 0}$ poset translation $\gamma(z, k, r) = (z, k, 3r)$.*

1. *If Z_X and Z_Y are zig-zag sequences of finite nonempty subsets of a good metric space W, then $\mathcal{D}C(Z_X)$ and $\mathcal{D}C(Z_Y)$ are $s^\delta \circ \gamma$ interleaved for $\delta > \max\{d_{Pr}(\nu^z_X, \nu^z_Y) : z = i$ or $i, i+1\}$*

2. *If $X = \bigcup_{z \in \mathcal{Z}} X_z$ and $Y = \bigcup_{z \in \mathcal{Z}} Y_z$ are finite nonempty metric spaces than $\mathcal{D}R(Z_X)$ and $\mathcal{D}R(Z_Y)$ are $s^\delta \circ \gamma$ homotopy interleaved for all $\delta > \inf_{\varphi,\psi}\{\max\{d_{Pr}(\varphi_*\mu^z_X, \psi_*\mu^z_Y) : z = i$ or $i, i+1\}\}$ where $\varphi : X \to W$ and $\psi : Y \to W$ range over all isometric embeddings into metric spaces (W, d_W).*

Proof. Part 1 follows from theorem 5.0.1 part 2 and a variation on proposition 3.4 from [27] proving that $\mathcal{D}C(\mathcal{Z}_X)$ and $\mathcal{S}C(\mathcal{Z}_X)$ are (γ, Id) interleaved. Part 2 then follows by the same logic as part 3 of theorem 20.

The zig-zag version of proposition 3.4 of [27] is essentially the same as appears in the paper. First we construct the following filtration:

$$\mathcal{D}\mathcal{O}(Z_X) : \mathcal{Z} \times \mathbb{R}^{op}_{\geq 0} \times \mathbb{R}_{\geq 0} \to \mathbf{Top}$$

$$\mathcal{D}\mathcal{O}(Z_X)_{z,k,r} = \{w \in W : d(x, w) < r \text{ for some } x \text{ a vertex in } \mathcal{D}C(Z_X)_{z,k,r}\}$$

Clearly we have $\mathcal{N}(\mathcal{DO}(Z_X)) = \mathcal{D}C(Z_X)$, therefore by the persistent nerve theorem for general diagrams (theorem 4.11 in [27]), $\mathcal{N}(\mathcal{DO}(Z_X))$ is weakly equivalent to $\mathcal{D}C(Z_X)$. Thus it is sufficient to show that $\mathcal{DO}(Z_X)$ and $\mathcal{M}(Z_X)$ are (γ, Id)-interleaved.

First we will construct an inclusion,

$$\mathcal{DO}(Z_X)_{z,k,r} \hookrightarrow (\mathcal{M}(Z_X) \circ \gamma)_{z,k,r} = \mathcal{M}(Z_X)_{z,k,3r}$$

Let $w \in \mathcal{DO}(Z_X)_{z,k,r}$. Then there must exist an $x \in X_z$ such that $d(x, w) < r$ and x has degree at least $k-1$ in the 1-skeleton of $C(X_z)_r$. This means there must exist a set $A \subset X_z$ with $|A| \geq k-1$ and $d(x, a) < 2r$ for all $a \in A$. By the triangle inequality this implies $d(w, a) < 3r$ for all $a \in A$, therefore $w \in \mathcal{M}(Z_X)_{z,k,3r}$.

Next we will construct an inclusion.

$$\mathcal{M}(Z_X)_{z,k,r} \hookrightarrow \mathcal{DO}(Z_X)_{z,k,r}$$

Let $w \in \mathcal{M}(Z_X)_{z,k,r}$. Then there exists a subset $A \subset X_z$ with $|A| \geq k$ and $d(w, a) < r$ for all $a \in A$. By the triangle inequality, for any $a, a' \in A$, we have $d(a, a') < 2r$, therefore any $a \in A$ has degree at least $k-1$ in the 1-skeleton of $C(X_z)_r$. Thus $w \in \mathcal{DO}(Z_X)_{z,k,r}$.

Since this interleaving is constructed via inclusions it commutes with the inclusion maps in the zig-zag direction, thus $\mathcal{DO}(Z_X)$ and $\mathcal{M}(Z_X)$ are interleaved, concluding the proof. $\square$

Conclusion

In conclusion, the main contribution of this thesis is a construction of a variant of persistent homology where the indexing poset is taken to be $\mathcal{Z} \times \mathbb{R}$. This construction was motivated by the need for an additional scale parameter in various applications of zig-zag persistent homology. We established fundamental results for $\mathcal{Z} \times \mathbb{R}$ persistence such as Gromov-Hausdorff stability and convergence under a sampling regime. Additionally we explored potential invariants on the space of $\mathcal{Z} \times \mathbb{R}$ modules, such as a variant on the signed barcode and fibered barcode. Finally, we extended stability results for density sensitive bi-filtrations filtrations with respect to the Prokhorov distance. We hope these contributions provide groundwork for further exploration and applications of $\mathcal{Z} \times \mathbb{R}$ persistence.

An important open question is whether there exists an efficient algorithm for computing $\mathcal{Z} \times \mathbb{R}$ persistence modules. Developing such an algorithm would be critical for practical applications. Strategies from existing computational methods in multi-parameter and zig-zag persistence may offer valuable insights for approaching the problem in the $\mathcal{Z} \times \mathbb{R}$ case.

Another key consideration for practical applications is vector representations of $\mathcal{Z} \times \mathbb{R}$ persistence modules. In standard $\mathbb{R}$ persistence, methods such as the persistence landscape and persistence images enable the use of statistical and machine learning methods to data analysis using $\mathbb{R}$ persistence modules. It is worth while to explore whether similar approaches can be done for $\mathcal{Z} \times \mathbb{R}$ persistence modules.

References

[1] M. Lesnick, *Notes on multiparameter persistence*, `https://www.albany.edu/~ML644186/840_2022/Math840_Notes_22.pdf`, Accessed: 2024-08-23, 2023.

[2] M. B. Botnan and M. Lesnick, "An introduction to multiparameter persistence," in *Representations of Algebras and Related Structures: International Conference on Representations of Algebras, ICRA 2020*, EMS Press, 2023, pp. 77–150.

[3] P. Niyogi, S. Smale, and S. Weinberger, "Finding the homology of submanifolds with high confidence from random samples," *Discrete Comput. Geom.*, vol. 39, no. 1-3, pp. 419–441, 2008.

[4] M. B. Botnan and G. Spreemann, "Approximating persistent homology in euclidean space through collapses," *Applicable Algebra in Engineering, Communication and Computing*, vol. 26, no. 1–2, 73–101, Jan. 2015.

[5] F. Chazal and S. Y. Oudot, "Towards persistence-based reconstruction in euclidean spaces," in *Proceedings of the Twenty-Fourth Annual Symposium on Computational Geometry*, ser. SCG '08, College Park, MD, USA: Association for Computing Machinery, 2008, 232–241, ISBN: 9781605580715.

[6] U. Bauer, M. Kerber, F. Roll, and A. Rolle, "A unified view on the functorial nerve theorem and its variations," *Expositiones Mathematicae*, vol. 41, no. 4, p. 125 503, Dec. 2023.

[7] V. Silva and G. Carlsson, "Topological estimation using witness complexes," *Proc. Sympos. Point-Based Graphics*, Jun. 2004.

[8] V. de Silva and R. Ghrist, "Coverage in sensor networks via persistent homology," *Algebraic Geometric Topology*, vol. 7, no. 1, pp. 339 –358, 2007.

[9] R. Seidel, "On the number of faces in higher-dimensional voronoi diagrams," in *Proceedings of the Third Annual Symposium on Computational Geometry*, ser. SCG '87, Waterloo, Ontario, Canada: Association for Computing Machinery, 1987, 181–185, ISBN: 0897912314.

[10] C. Webb, "Decomposition of graded modules," *Proceedings of the American Mathematical Society*, vol. 94, no. 4, pp. 565–571, 1985.

[11] W. Crawley-Boevey, "Decomposition of pointwise finite-dimensional persistence modules," *Journal of Algebra and Its Applications*, vol. 14, no. 05, p. 1 550 066, 2015. eprint: `https://doi.org/10.1142/S0219498815500668`.

[12] M. B. Botnan and W. Crawley-Boevey, "Decomposition of persistence modules," *Proceedings of the American Mathematical Society*, vol. 148, no. 11, pp. 4581–4596, 2020.

[13] F. Chazal, V. De Silva, M. Glisse, and S. Oudot, *The structure and stability of persistence modules*. Springer, 2016, vol. 10.

[14] U. Bauer and M. Lesnick, "Induced matchings and the algebraic stability of persistence barcodes," *Journal of Computational Geometry*, Vol. 6 No. 2 (2015): Special issue of Selected Papers from SoCG 2014, 2015.

[15] H. Bjerkevik, "On the stability of interval decomposable persistence modules," *Discrete Computational Geometry*, vol. 66, Jul. 2021.

[16] M. B. Botnan, S. Oppermann, S. Oudot, and L. Scoccola, "On the bottleneck stability of rank decompositions of multi-parameter persistence modules," *Advances in Mathematics*, vol. 451, p. 109 780, 2024.

[17] H. B. Bjerkevik and M. Lesnick, *ℓ^p-distances on multiparameter persistence modules*, 2021. arXiv: `2106.13589 [math.AT]`.

[18] A. J. Blumberg and M. Lesnick, "Universality of the homotopy interleaving distance," *Transactions of the American Mathematical Society*, vol. 376, no. 12, pp. 8269–8307, 2023.

[19] D. Cohen-Steiner, H. Edelsbrunner, and J. Harer, "Stability of persistence diagrams," *Discrete & Computational Geometry*, vol. 37, no. 1, pp. 103–120, 2007.

[20] F. Chazal, D. Cohen-Steiner, L. J. Guibas, F. Mémoli, and S. Y. Oudot, "Gromov-hausdorff stable signatures for shapes using persistence," in *Proceedings of the Symposium on Geometry Processing*, ser. SGP '09, Berlin, Germany: Eurographics Association, 2009, 1393–1403.

[21] A. J. Blumberg, I. Gal, M. A. Mandell, and M. Pancia, "Robust statistics, hypothesis testing, and confidence intervals for persistent homology on metric measure spaces," *Foundations of Computational Mathematics*, vol. 14, pp. 745 –789, 2012.

[22] G. Carlsson and A. Zomorodian, "The theory of multidimensional persistence," *Discrete & Computational Geometry*, vol. 42, no. 1, pp. 71–93, 2009.

[23] D. R. Sheehy, "A multicover nerve for geometric inference," in *Proceedings of the Canadian Conference on Computational Geometry (CCCG)*, 2012, pp. 309–314.

[24] H. Edelsbrunner and G. Osang, “The Multi-cover Persistence of Euclidean Balls,” in *34th International Symposium on Computational Geometry (SoCG 2018)*, B. Speckmann and C. D. Tóth, Eds., ser. Leibniz International Proceedings in Informatics (LIPIcs), vol. 99, Dagstuhl, Germany: Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2018, 34:1–34:14, ISBN: 978-3-95977-066-8.

[25] F. Chazal, D. Cohen-Steiner, and Q. Mérigot, “Geometric inference for probability measures,” *Foundations of Computational Mathematics*, vol. 11, pp. 733–751, Dec. 2011.

[26] M. Lesnick and M. Wright, *Interactive visualization of 2-d persistence modules*, 2015. arXiv: `1512.00180 [math.AT]`.

[27] A. Blumberg and M. Lesnick, “Stability of 2-parameter persistent homology,” *Foundations of Computational Mathematics*, Oct. 2022.

[28] P. Bubenik, V. de Silva, and J. Scott, “Metrics for generalized persistence modules,” *Foundations of Computational Mathematics*, vol. 15, no. 6, 1501–1531, Oct. 2014.

[29] M. B. Botnan, S. Oppermann, and S. Oudot, “Signed barcodes for multi-parameter persistence via rank decompositions and rank-exact resolutions,” *Foundations of Computational Mathematics*, 2024.

[30] A. Cerri, B. D. Fabio, M. Ferri, P. Frosini, and C. Landi, “Betti numbers in multidimensional persistent homology are stable functions,” *Mathematical Methods in the Applied Sciences*, vol. 36, no. 12, pp. 1543–1557, Aug. 2013.

[31] C. Landi, “The rank invariant stability via interleavings,” in *Research in Computational Topology*, E. W. Chambers, B. T. Fasy, and L. Ziegelmeier, Eds. Cham: Springer International Publishing, 2018, pp. 1–10, ISBN: 978-3-319-89593-2.

[32] S. Biasotti, A. Cerri, P. Frosini, and D. Giorgi, “A new algorithm for computing the 2-dimensional matching distance between size functions,” *Pattern Recognition Letters*, vol. 32, no. 14, pp. 1735–1746, 2011.

[33] G. Carlsson, V. de Silva, and D. Morozov, “Zigzag persistent homology and real-valued functions,” in *Proceedings of the Twenty-Fifth Annual Symposium on Computational Geometry*, ser. SCG ’09, Aarhus, Denmark: Association for Computing Machinery, 2009, 247–256, ISBN: 9781605585017.

[34] P. Bendich, H. Edelsbrunner, D. Morozov, and A. Patel, “Homology and robustness of level and interlevel sets,” *Homology, Homotopy and Applications*, vol. 15, no. 1, pp. 51 –72, 2013.

[35] G. Carlsson and V. de Silva, "Zigzag persistence," *Foundations of Computational Mathematics*, vol. 10, no. 4, pp. 367–405, 2010.

[36] M. Botnan and M. Lesnick, "Algebraic stability of zigzag persistence modules," *Algebraic & Geometric Topology*, vol. 18, no. 6, 3133–3204, Oct. 2018.

[37] M. Botnan, "Interval decomposition of infinite zigzag persistence modules," *Proceedings of the American Mathematical Society*, vol. 145, Jul. 2015.

[38] M. B. Botnan, J. Curry, and E. Munch, *A relative theory of interleavings*, 2020. arXiv: `2004.14286 [math.CT]`.

[39] F. Chazal, V. de Silva, and S. Oudot, "Persistence stability for geometric complexes," *Geometriae Dedicata*, vol. 173, no. 1, pp. 193–214, 2014.

[40] E. Riehl, *Category theory in context* (Aurora: Dover modern math originals). Dover Publications, 2017, ISBN: 978-0-486-82080-4.

[41] S. Lane, *Categories for the Working Mathematician* (Graduate Texts in Mathematics). Springer, 1998, ISBN: 9780387984032.

[42] H. Derksen and J. Weyman, *An Introduction to Quiver Representations* (Graduate Studies in Mathematics). American Mathematical Society, 2017, ISBN: 9781470425562.

[43] A. Kirillov, *Quiver Representations and Quiver Varieties* (Graduate Studies in Mathematics). American Mathematical Society, 2016, ISBN: 9781470423070.

[44] J. A. Drozd, "Tame and wild matrix problems," in *Representation Theory II*, V. Dlab and P. Gabriel, Eds., Berlin, Heidelberg: Springer Berlin Heidelberg, 1980, pp. 242–258, ISBN: 978-3-540-38387-1.

[45] P. Gabriel, "Unzerlegbare Darstellungen. I," *Manuscripta Math.*, vol. 6, 71–103; correction, ibid. 6 (1972), 309, 1972.

[46] L. A. Nazarova, "Representations of quivers of infinite type," *Mathematics of the USSR-Izvestiya*, vol. 7, no. 4, p. 749, 1973.

Appendix A: Finiteness Conditions

Throughout the body of work on persistence modules, various finiteness conditions are assumed in different contexts and for different reasons. All such conditions are very reasonable assumptions when considering persistent homology from the perspective of data analysis. Any persistent modules obtained from the usual methods applied to finite data sets have nice finite/discrete properties. In this section I will organize some of the more commonly seen assumptions as a reference to refer back to throughout this thesis.

Definition A.0.1 (Pointwise finite dimensional). For any poset $\mathcal{P}$ we say a $\mathcal{P}$ persistence module $M : \mathcal{P} \to \mathbf{Vect}_k$ is *pointwise finite dimensional* (or p.f.d. for short) if for every $p \in \mathcal{P}$, $\dim(M(p)) < \infty$.

Occasionally the stronger condition of finitely presented is required:

Definition A.0.2 (Finitely presented). A $\mathcal{P}$ persistence module M is *finitely presented* if there exists an exact sequence,

$$F_1 \to F_0 \to M \to 0$$

where F_1 and F_0 are finitely generated free $\mathcal{P}$ persistence modules.

Many results require that the poset $\mathcal{P}$ be discrete, such as in the following definition:

Definition A.0.3 (Locally finite poset). A poset $\mathcal{P}$ is *locally finite* if any interval $[p, q]$ in $\mathcal{P}$ is finite.

Clearly $\mathbb{Z}$ and $\mathbb{N}$ are locally finite, as well as $\mathbb{Z}^d$ for any d. Note that the most important poset in the study of persistent homology, $\mathbb{R}$, is not locally finite. However any poset can essentially be assumed to be locally finite if it satisfies the next definition.

The following definition is meant to capture the idea of a $\mathbb{R}$ persistence module that changes only at discrete points. This allows us to apply theorems proven for $\mathbb{Z}$ persistence modules to filtrations that are discrete but more natural over r such as the Čech or Vietoris Rips filtrations applied to finite metric spaces.

Definition A.0.4 (Essentially discrete)**.** A $\mathbb{R}$-persistence module M is *essentially discrete* if there is a monotonic injection $j : \mathbb{Z} \to \mathbb{R}$ such that,

1. $\lim_{z \to \pm\infty} \pm\infty$
2. For all $z \in \mathbb{Z}$ and $r \leq s \in [j(z), j(z+1))$, $M(r \to s)$ is an isomorphism.

A $\mathcal{Z} \times \mathbb{R}$ persistence module is essentially discrete if for every $z \in \mathcal{Z}$, the $\mathbb{R}$ module given by $M(z, -)$ is essentially discrete.

This definition is always satisfied in the setting of applying a filtration and then homology to a finite set of data points as the filtration only changes at a finite number of possible thresholds given by the set of pairwise distances of the points.

Another approach to discretizing a persistence module is to pixilate it along a locally finite sub-poset $\mathcal{P}' \subset \mathcal{P}$ by pulling back and then Kan extending along the inclusion $i : \mathcal{P}' \to \mathcal{P}$, $M \to L_i M \circ i$. For a further discussion of pixilation see [38].

The following definition arises in the context of the rank invariant. It is a necessary assumption for the construction of the signed barcode using the Mobius inversion formula.

Definition A.0.5 (Upward finite support)**.** For $(\mathcal{I}, \subseteq)$ the poset of intervals in $\mathcal{P}$, a function $r : \mathcal{I} \to \mathbb{Z}$ has *upward finite support* if the set $\{J : I \subseteq J, r(J) \neq 0\}$ is finite.

In the context of the generalized rank invariant, this amounts to the condition that for sufficiently large intervals, the map $J \lim M|_J \to \mathrm{colim} M|_J$ has rank 0, which is clearly true in all practical settings.

Appendix B: Category Theory Background

In this section I will give a brief overview of the category theory definitions and theorems used in this thesis. For a more complete treatment, see [40] or [41].

To begin, recall the definition of a partially ordered set (poset).

Definition B.0.1 (Poset)**.** A poset $(X, \leq)$ is a set X equipped with a comparison $\leq$ such that,

1. $x \leq x$ for all $x \in X$
2. $x \leq y, y \leq z \implies x \leq z$
3. $x \leq y, y \leq x \implies x = y$

In other words, a poset is a weakened version of an ordered set where we do not require that all objects be comparable. Any fully order set is also a partially ordered set, for instance $\mathbb{R}$ or $\mathbb{N}$. The following two examples will also be important to us later on:

Example B.0.1. Let A be a set. Then the set of all subsets (or power set) of A, denoted 2^A is a poset with order given by $E \leq E'$ if $E \subseteq E'$.

Example B.0.2 (Product Poset)**.** Let A and B be any posets. Then $A \times B$ is a poset with order relation given by $(a, b) \leq (a', b')$ if and only if $a \leq a', b \leq b'$.

This definition is equivalent to the following, which will be used in this thesis:

Definition B.0.2 (Poset category)**.** A poset category $\mathcal{P}$ is a category with at most 1 morphism between any 2 objects.

The equivalence is given by stating that $p_1 \le p_2$ if there exists a morphism $p_1 \to p_2$ in $\mathcal{P}$. The composition and identity axioms of a category ensure that this relation satisfies the axioms of a poset.

The metric we will define on the space of 2 parameter zigzag persistence vector spaces depends on the use of left Kan extensions. Informally, the left Kan extension of a functor $F : \mathcal{C} \to \mathcal{D}$ along an inclusion of a subcategory $i : \mathcal{C} \to \mathcal{C}'$ is a way of extending F to $\mathcal{C}'$ via gluing together (taking the colimit) $F(c)$ for all objects $c \in \mathcal{C}$ that map into a given c' to form a natural choice for $F(c')$. Concretely, we begin by defining the Kan extension in terms of its universal property:

Definition B.0.3 (Left Kan Extension)**.** Given functors $F : \mathcal{C} \to \mathcal{D}$, $G : \mathcal{C} \to \mathcal{E}$, a left Kan extension of F along G is a functor $L_G F : \mathcal{D} \to \mathcal{E}$ along with a natural transformation $\eta : F \Rightarrow L_G F \circ G$ such that for any other pair $(H : \mathcal{D} \to \mathcal{E}, \epsilon : F \Rightarrow H \circ G)$, ϵ factors uniquely as in the following diagram:

$$\begin{tikzcd}
\mathcal{C} \arrow[rr, "F"] \arrow[dr, "G"'] & & \mathcal{E} \\
& \mathcal{D} \arrow[u, phantom, "\eta\Downarrow"] \arrow[ur, bend left, "L_G F"] \arrow[ur, bend right, "H"'] \arrow[ur, phantom, "\exists!\Rightarrow"] &
\end{tikzcd}$$

This universal property states that $L_G F$ is initial among the functors extending F to $\mathcal{D}$, that is, any other such extension H comes equipped with a natural transformation $L_G F \Rightarrow H$. This provides some intuition for why we choose this extension to define the metric. In a sense, it is the extension that is the “closest" to the original functor F. However, for this thesis we will find the following concrete characterization of the left Kan extension more useful.

To begin with, for a given $d \in \mathcal{D}$, we need to define the comma category of a functor G over d.

Definition B.0.4 (Comma category)**.** Let $G : \mathcal{C} \to \mathcal{D}$ be a functor, $d \in ob(\mathcal{D})$. The comma category $(G \downarrow d)$ has objects $\{(f, c) : c \in ob(\mathcal{C}), f : G(c) \to d\}$ and morphisms $g : (f, c) \to (f', c')$ given by morphisms $g : c \to c'$ in $\mathcal{C}$ that satisfy the following commutative triangle:

$$\begin{tikzcd} G(c) \arrow[rr, "G(g)"] \arrow[dr, "f"] & & G(c') \arrow[dl, "f'"] \\ & d & \end{tikzcd}$$

This category comes equipped with a projection $\pi_d : (G \downarrow d) \to \mathcal{C}$ sending a pair (f, c) to c and a map $g : (f, c) \to (f', c')$ to the corresponding $g : c \to c'$. Composing this projection with F gives a functor $\pi_d \circ F : (G \downarrow d) \to \mathcal{C} \to \mathcal{E}$. Additionally, for any $h : d \to d'$, there exists a functor $h_* : (G \downarrow d) \to (G \downarrow d')$ sending (f, c) to $(h \circ f, c)$.

Proposition B.0.1. *If for every $d \in \mathcal{D}$, the colimit $colim(F \circ \pi_d)$ exists, then the functor defined on objects as,*

$$L_G F(d) = colim(F \circ \pi_d)$$

and on morphisms h via the map $colim(F \circ \pi_d) \to colim(F \circ \pi_{d'})$ induced by h_* is a left Kan extension with natural transformation η induced by the canonical inclusion maps $F(c) \to colim(F \circ \pi_{G(c)})$ as the object $(id_c, c) \in (G \downarrow G(c))$

Before beginning the proof we'll provide some intuition for what this theorem is telling us. For now assume that G is an inclusion of categories. The comma category should be thought of as the diagram of all $c \in \mathcal{C} \subset \mathcal{D}$ that map into d with arrows between them that commute with these maps. We then take the colimit to glue together the corresponding $F(c)$ along these maps to construct a natural choice for how this functor should behave on d. The proof of this theorem involves a lot of diagram chasing, but the important point is that this "glued" object ends up being the natural choice as it is initial among all others.

Proof. To prove this theorem we must show that this functor and natural transformation satisfy the universal property. Suppose we have a functor $H : \mathcal{D} \to \mathcal{E}$ and a natural transformation $\varepsilon : F \Rightarrow H \circ G$. We want to show that ε factors uniquely through a natural transformation $F \xrightarrow{\eta} L_G F \circ G \xrightarrow{\varepsilon'} H \circ G$.

We begin by defining $\varepsilon'(c) : L_G F(c) = \operatorname{colim}(F \circ \pi_{G(c)}) \to H \circ G(c)$. By the universal property of $\operatorname{colim}(F \circ \pi_{G(c)})$, such a map is specified by the data of maps $F(c') \to H(G(c))$ for every $f : G(c') \to G(c)$ such that for every morphism $g : (f, c') \to (f', c'')$ in $(G \downarrow G(c))$, the following diagram commutes,

$$\begin{tikzcd} F(c') \arrow[rr, "F(g)"] \arrow[dr] & & F(c'') \arrow[dl] \\ & H(G(c)) & \end{tikzcd}$$

These maps arise naturally from the natural transformation $\varepsilon : F \Rightarrow H \circ G$. Specifically,

$$F(c') \xrightarrow{\varepsilon(c')} H(G(c')) \xrightarrow{H(f)} H(G(c))$$

The commutativity follows from the naturality of ε and the functoriality of H. To check that these maps assemble into a natural transformation ε' we must verify that the following diagram commutes for every $g : c \to c'$,

$$\begin{tikzcd} \operatorname{colim}(F \circ \pi_{G(c)}) \arrow[r, "g_*"] \arrow[d, "\varepsilon'(c)"] & \operatorname{colim}(F \circ \pi_{G(c)}) \arrow[d, "\varepsilon'(c')"] \\ H(G(c)) \arrow[r, "H(G(g))"] & H(G(c')) \end{tikzcd}$$

Note that by the universal property of $\operatorname{colim}(F \circ \pi_{G(c)})$, any map is uniquely characterized by the data of the maps obtained by precomposing with the canonical inclusions. That is, it is sufficient to show that for any $(f, c'') \in (G \downarrow G(c))$, the following outer diagram commutes,

$$\begin{tikzcd}
F(c'') \arrow[dr, "i_{F(c'')}"] \arrow[drr, bend left, "g_* \circ i_{F(c'')}"] \arrow[ddr, bend right, "\varepsilon'(c) \circ i_{F(c'')}"'] & & \\
& \mathrm{colim}(F \circ \pi_{G(c)}) \arrow[r, "g_*"] \arrow[d, "\varepsilon'(c)"] & \mathrm{colim}(F \circ \pi_{G(c')}) \arrow[d, "\varepsilon'(c')"] \\
& H(G(c)) \arrow[r, "H(G(g))"] & H(G(c'))
\end{tikzcd}$$

The top arrow includes $F(c'')$ as the object $(G(g) \circ f, c'')$ in $(G \downarrow G(c'))$, and the right-side down arrow is then defined by $H(G(g) \circ f) \circ \varepsilon(c'')$. The left-side down arrow is defined by $H(f) \circ \varepsilon(c'')$ so the left-bottom composition gives $H(G(g)) \circ H(f) \circ \varepsilon(c'')$ which is equivalent to the top right composition $H(G(g) \circ f) \circ \varepsilon(c'')$ via functoriality of H.

Now we have a natural transformation $\varepsilon' : L_G F \Rightarrow H$. All that remains to be shown is that $\varepsilon = \varepsilon' \circ \eta$. $\eta(c)$ is the inclusion map $F(c) \to \mathrm{colim}(F \circ \pi_{G(c)})$ arising from the object $(id_c, c) \in (G \downarrow G(c))$, thus $\varepsilon'(c) \circ \eta(c) = H(id_c) \circ \varepsilon(c) = \varepsilon(c)$. □

Remark B.0.1. This characterization of the left Kan extension is particularly useful in the context of this thesis, as it gives a more concrete description of how the extended functor behaves on objects. It also provides a way of understanding maps out of the Kan extended functors, as the universal property of the colimit states that maps out of the colimit of a diagram $\mathrm{colim}_I F$ are uniquely characterized by maps out of each $F(i)$ that commute with the $F(i \to i')$ for $i \to i' \in I$.

Proposition B.0.2. *If the left Kan extension along $G : \mathcal{C} \to \mathcal{D}$ exists for all $F : \mathcal{C} \to \mathcal{E}$, then it is functorial on the category $Fun(\mathcal{C}, \mathcal{E})$. Specifically, for every $\alpha : F \Rightarrow F'$ there exists a $L_G\alpha : L_G F \to L_G F'$ such that $L_G id_F = id_{L_G F}$ and $L_G(\alpha \circ \beta) = L_G\alpha \circ L_G\beta$.*

Proof. Using the above formula for $L_G F$, we define the natural transformation $L_G\alpha$ using the universal property of the colimit. A map out of $L_G F(d) = \mathrm{colim}_{(G \downarrow d)} F \circ \pi_d$ is equivalent to the data of a map out of $F(c)$ for every $G(c) \to d$ that commutes with all maps in the diagram. Therefore define $L_G\alpha(d) : L_GF(d) \to L_GF'(d)$ to be the data of,

$$F(c) \xrightarrow{\alpha(c)} F'(c) \xrightarrow{i_{F'(c)}} \mathrm{colim}_{(G\downarrow d)} F' \circ \pi_d = L_GF'(d)$$

for every $(c, f : G(c) \to d) \in (G \downarrow d)$, where the $i_{F'(c)}$ are the maps of the universal cocone over $\mathrm{colim}_{(G\downarrow d)} F' \circ \pi_d$. If we consider the following diagram, both vertical arrows are are induced by the map of the comma categories $(G \downarrow d) \to (G \downarrow d')$ given by post composition with $d \to d'$ and the horizontal arrows are both induced by the data of the cocone maps composed with $\alpha(c)$. Therefore the diagram commutes, meaning $L_G\alpha$ is natural.

$$\begin{tikzcd}
L_GF(d) \arrow[r, "L_G\alpha(d)"] \arrow[d, "L_GF(d\to d')"] & L_GF'(d) \arrow[d, "L_GF'(d\to d')"] \\
L_GF(d') \arrow[r, "L_G\alpha(d')"] & L_GF'(d')
\end{tikzcd}$$

To see that $L_Gid_F = id_{L_GF}$ we observe that the data of $i_{F(c)} \circ id_{F(c)}$ induces the identity map on $\mathrm{colim}_{(G\downarrow d)} F \circ \pi_d$. The last point we need verify is that $L_G\alpha \circ \beta = L_G\alpha \circ L_G\beta$. The data of $L_G(\beta \circ \alpha)(d)$ is defined by the compositions,

$$\{F(c) \xrightarrow{\alpha(c)} G(c) \xrightarrow{\beta(c)} H(c) \xrightarrow{i_{H(c)}} \mathrm{colim}_{(G\downarrow d)} H \circ \pi_d\}_{(c, G(c)\to d) \in (G\downarrow d)}$$

The data of $L_G(\beta) \circ L_G(\alpha)$ is determined by the compositions,

$$\begin{tikzcd}
F(c) \arrow[r, "\alpha(c)"] & G(c) \arrow[r, "i_{G(c)}"] \arrow[dr, bend right, "\beta(c)"] & \mathrm{colim}_{(G\downarrow d)} G \circ \pi_d & \\
& & H(c) \arrow[r, "i_{H(c)}"] & \mathrm{colim}_{(G\downarrow d)} H \circ \pi_d
\end{tikzcd}$$

for all $(c, G(c) \to d) \in (G \downarrow d)$. In both cases the characteristic map $F(c) \to \mathrm{colim}_{(G\downarrow d)} H \circ \pi_d$ is given by $\beta(c) \circ \alpha(c)$ thus they are the same map. $\square$

Appendix C: Quiver Representations

In this section we will provide a brief background on the theory of quiver representations. The primary references for this section are [42] and [43].

Definition C.0.1 (Quiver representation)**.** A *quiver* $\mathcal{Q}$ is a directed graph. A representation of $\mathcal{Q}$ is an assignment of a vector space to each node of the graph and a linear map to each edge.

It is not generally required that the resulting diagram commute, but for our purposes we will consider only those quiver representations that do. Therefore we may view a quiver representation to be a functor $F : \mathcal{Q} \to \mathbf{Vect}_k$, where $\mathcal{Q}$ is the category with an object for each node and a morphism $e : v \to w$ for each edge. In particular, any essentially discrete persistence modules is an example of a quiver representation.

A quiver representation is said to be *indecomposable* if it cannot be written as $F = G \oplus H$ for G and H both nonzero. Every quiver has a unique decomposition into indecomposable representations. In the case of 1-parameter persistent modules, these indecomposable representations can be easily classified as the interval persistence modules. However, this is one of the rare cases where it is feasible to classify the indecomposable representations of the graph. The following theorems provide a complete classification of the complexity of indecomposable representations of quivers.

Definition C.0.2 (finite, tame and wild type)**.** Let Q be a quiver, and consider $d : obQ \to \mathbb{Z}_+$ an assignment of a dimensions to each node of Q. Then Q is

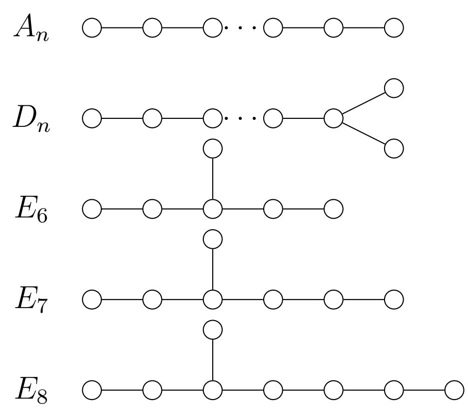
1. Finite type if the number of isomorphism classes of indecomposable representations with dimension d is finite

2. Tame type if the set of isomorphism classes of indecomposable representations with dimension d is a union of a finite number of one-parameter families and a finite number of isolated points

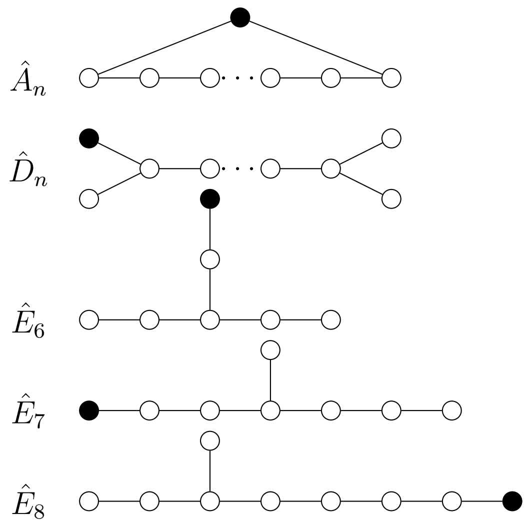
3. Wild type if there exists a fully faithful functor from the category of representations of any finitely generated k-algebra to the category of representations of $\mathcal{Q}$. In other words, the representation theory of Q is at least as complicated as the representation theory of any finite-dimensional associative algebra.

A theorem of Drozd's states that every quiver is either of finite, tame or wild type [44].

Theorem C.0.1. *[Gabriel's theorem 1 [45]] A connected quiver Q is of finite type if and only if the underlying undirected graph is one of the following forms, known as Dynkin graphs:*



Theorem C.0.2 ([46])**.** *A connected quiver Q which is not of finite type is of tame type if* *and only if the underlying undirected graph is of one of the following extended Dynkin graphs:*



Thus from these theorems we can see that $\mathbb{Z}, \mathbb{N}$ and $\mathcal{Z}$ persistence modules are of finite type, while $\mathbb{R}^n$ and $\mathcal{Z} \times \mathbb{R}$ persistence modules are of wild type, meaning classifying the space of indecomposable modules is hopelessly complex.